REPRESENTATIONS OF LIE ALGEBRAS, WITH APPLICATIONS TO PARTICLE PHYSICS

JAMES MARRONE
UNIVERSITY OF CHICAGO MATHEMATICS REU, AUGUST 2007

Abstract. The structure of Lie groups and the classification of their representations are subjects undertaken by many an author of mathematics textbooks; Lie algebras are always considered as an indispensable component of such a study. Yet every author approaches Lie algebras differently - some begin axiomatically, some derive the algebra from prior principles, and some barely connect them to Lie groups at all. The ensuing problem for the student is that the importance of the Lie algebra can only be deduced by "reading between the lines," so to speak; the relationship between a Lie algebra and a Lie group is not always illuminated (in fact, there isn't always a relationship in the first place) and so considering the representations of Lie algebras may appear redundant or at best cloudy. The approach to Lie algebras in this paper should clarify such problems. Lie algebras are historically a consequence of Lie groups (namely, they are the tangent space of a Lie group at the identity) which may be axiomatized so that they stand alone as an algebraic concept. The important correspondence between representations of Lie algebras and Lie groups, however, makes Lie algebras indispensable to the study of Lie groups; one example of this is the Eight-Fold Way of particle physics, which is actually just an 8-dimensional representation of $\mathfrak{sl}_3(\mathbb{C})$ but which has the enlightening property of corresponding to the strong nuclear interaction, giving a physical manifestation of the action of a representation and showing just one of the many interesting ramifications of mathematics on science within the past 50 years.

1. Lie Groups and Representations

We begin with

Definition 1. A *Lie group* is a group such that the map $\omega : G \times G \to G$ sending $(x, y) \mapsto xy^{-1}$ is C^∞.

The important consequence of this is that a Lie group is endowed with a manifold structure. Without looking very far, we can easily find Lie groups - just take groups of non-singular $n \times n$ matrices over a complete metric space, such as $\mathbb{R}$ or $\mathbb{C}$. Then this is locally equivalent to working in the finite-dimensional space $\mathbb{R}^m$ or $\mathbb{C}^m$, depending on the number of degrees of freedom characterizing the matrix group. For example, $SL_n(\mathbb{R})$, the group of non-singular real matrices with determinant 1, has dimension $n^2 - 1$. These are elementary facts from linear algebra, and we might create a complete list of all the classical groups $SO_n(\mathbb{R})$, $SU_n(\mathbb{R})$, $Sp_{p,q}(\mathbb{R})$ and the corresponding complex Lie groups with their corresponding dimensions. To see how these might have a manifold structure, consider the map above as a continous map on a vector space. For example, consider any matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{R})$. Since $ad-cb = 1$, we have 3 free parameters but the last element is then fixed by the other three in order to get determinant 1. This group consists of many 'components,' one of which is the set of matrices $\begin{pmatrix} a & b \\ c & \frac{1+cb}{a} \end{pmatrix}$ such that $a \neq 0$. We can map such a matrix to the element $\left(a, b, c, \frac{1+cb}{a}\right) \in \mathbb{R}^4$. This defines a subspace of $\mathbb{R}^4$ that gives the local structure that defines a manifold; to get differentiability, the map $\omega : (x, y) \mapsto xy^{-1}$ can be defined by

$$\left(\begin{pmatrix} a & b \\ c & \frac{1+cb}{a} \end{pmatrix}, \begin{pmatrix} e & f \\ g & \frac{1+fg}{e} \end{pmatrix}\right) \mapsto \begin{pmatrix} a & b \\ c & \frac{1+cb}{a} \end{pmatrix} \begin{pmatrix} \frac{1+fg}{e} & -f \\ -g & e \end{pmatrix} = \begin{pmatrix} \frac{a+afg}{e} - bg & \frac{c+cfg}{e} - \frac{g+gcb}{a} \\ be - af & \frac{e+ecb}{a} - cf \end{pmatrix}$$

corresponding to the map

$$\omega_{\mathbb{R}} : \left((a, b, c, \frac{1+cb}{a}), (e, f, g, \frac{1+fg}{e})\right) \mapsto (\frac{a+afg}{e} - bg, \frac{c+cfg}{e} - \frac{g+gcb}{a}, be - af, \frac{e+ecb}{a} - cf).$$

Since any of the six parameters a, b, c, e, f, g can be continuously varied in the reals, we can take $\frac{\partial \omega_{\mathbb{R}}}{\partial x_i}$ for $x_i = a, b, c$, etc. Hence $\frac{\partial \omega_{\mathbb{R}}}{\partial c} = (0, \frac{1+fg}{e} - \frac{gb}{a}, 0, \frac{eb}{a} - f)$ and so forth, and we have a smooth map on the group. This map gives a manifold structure because it provides a local coordinatization for each element by multiplying The next important concept is

Definition 2. A *representation* of a group G is a smooth group homomorphism $\varphi : G \to GL_n(V)$ for some finite-dimensional vector space V.

Hence, a representation assigns to each $g \in G$ a linear map $\varphi_g : V \to V$, and one can say for short that "$V$is a representation of G." Given V a finite-dimensional vector space over a field F, it is clear that if V is an FG-module, V is also a representation of G since we can define $\varphi_g(v) = g \cdot v$. Conversely, if V is a representation of G we can turn it into an FG-module by defining

$$(\sum \alpha_i g_i) \cdot v = \sum \alpha_i \varphi_{g_i}(v)$$

so that there is a one-to-one correspondence between finite-dimensional representations of G and FG-modules. This is useful in creating a conceptual basis for the idea of a representation, especially since finite-dimensional representations are the only ones in which we will be interested. Without going into too many details, two important examples are stated, as these will reappear later:

Example 1. (The Standard Representation): Let G be a permutation group, i.e. $G \cong S_n$ for some n. Then $g \in G$ acts on elements of F^n: let $e_1, e_2, \dots, e_n$ be a basis; then $g \cdot e_i = e_{g(i)}$. Then the vector space is a representation of G, as each element of the group is associated with a linear map on the vector space.

-

Example 2. (The Dual Representation): If V is an arbitrary vector space, take $\psi : V \to V$ with its dual $\psi^* : V^* \to V^*$ defined by $\psi^*(\rho) = \rho \circ \psi$ in the normal way for a vector space. In matrix form, ψ^* is ${}^t\psi$ and since ${}^t(\psi \cdot \phi) =^t \phi \cdot^t \psi$ for any two matrices ψ, ϕ, we have $(\psi \cdot \phi)^*(\rho) = \phi^*(\psi^*(\rho))$. If V is a representation of a group G, we have $\varphi_g$for each $g \in G$and we see $\varphi^*_{g \cdot h} = \varphi^*_h \circ \varphi^*_g$so this is an anti-homomorphism. By defining, in the case of group representations, $\varphi^*_g =^t \varphi_{g^{-1}}$, we get

$$\varphi^*_{g \cdot h} =^t \varphi_{(g \cdot h)^{-1}} =^t \left(\varphi_{h^{-1}} \cdot \varphi_{g^{-1}}\right) =^t \varphi_{g^{-1}} \cdot^t \varphi_{h^{-1}} = \varphi^*_g \circ \varphi^*_h$$

as desired. We have thus described V^*as a representation of G.

We now turn away from the group structure of a Lie group and explore the implications of its manifold structure.

2. Lie Algebras

2.1. **The Tangent Space at the Identity.** Since a Lie group is a finite-dimensional manifold, given a point g_o on its surface we may find the tangent space at g_o. That is, for G a Lie group we have a space of all analytic functions (not necessarily linear) $G \to G$. We may take the partial derivatives of any of these functions at our specified point g_o. We do this by fixing a chart (U, φ) and considering the space of operators $\left\{\xi^i \frac{\partial}{\partial x_i}\right\}$, applying them to analytic functions at g_o. The direction of the partial derivative is determined by ξ^i (or, specifically, the ratio of ξ^i to ξ^j for various coordinates i,j), and the space of such operators analyzed at $g_o$is the tangent space at g_o, $T_{g_o}G$. If we have a tangent vector at each point of the manifold, we have a vector field. We wish to invest the space of analytic vector fields A with the structure of a vector space, like its corresponding Lie group G, and so we need to have a multiplication. Merely taking two tangent vectors $\xi^i \frac{\partial}{\partial x_i}$ and $\eta^i \frac{\partial}{\partial x_i}$ at a point g_o and composing them as operators yields

$$\xi^i \frac{\partial}{\partial x_i}\left(\eta^j \frac{\partial}{\partial x_j}\right) = \xi^i \eta^j \frac{\partial^2}{\partial x_i \partial x_j} + \xi^i \frac{\partial \eta^j}{\partial x_i} \frac{\partial}{\partial x_j}$$

which is not a tangent vector because of the double partial derivative. However, if we take two such operators as before, we see that

$$\left(\xi^i \frac{\partial}{\partial x_i}\right)\left(\eta^j \frac{\partial}{\partial x_j}\right) - \left(\eta^j \frac{\partial}{\partial x_j}\right)\left(\xi^i \frac{\partial}{\partial x_i}\right) = \left(\xi^i \frac{\partial \eta^j}{\partial x_j} - \eta^j \frac{\partial \xi^i}{\partial x_j}\right) \frac{\partial}{\partial x_i}$$

which is itself another tangent vector. Hence, we can define multiplication of two vector fields $\mathsf{X}, \mathsf{Y} \in A$ to be $[\mathsf{X}, \mathsf{Y}] = \mathsf{XY} - \mathsf{YX}$, which is also known as the *commutator*.

This is a repetition of results in differential geometry, but there are some useful results for Lie groups. In particular, as with any group, we consider the action of a Lie group on itself by right (or left) multiplication, i.e. for $g \in G$ consider the map $\Phi_g(g') = g \cdot g'$. Any homomorphism of groups $\rho : G \to H$ must preserve this action, since

$$\rho(g \cdot g') = \rho(g) \cdot \rho(g') = \rho(g) \cdot \rho(g') = \Phi_{\rho(g)}(g').$$

That a map respects multiplication on the right or left is equivalent to saying that it is a homomorphism. Hence, finding group representations amounts in some sense to finding maps between groups that are compatible with right multiplication. Since right multiplication is everywhere continuous, we can take its differential at some point g_o. Let $X = \xi^i \frac{\partial}{\partial x_i}$ be a tangent vector at g_o and consider $X\Phi_g(v) = \xi^i \frac{\partial(\varphi \circ \Phi_g)(v)}{\partial v_i}\Big|_{v=g_o}$. Letting $\frac{\partial((\varphi \circ \Phi_g)(v))_k}{\partial v_i}\Big|_{v=g_o} = \Phi_g^{k,i}(g_o)$, we can define, for an arbitrary analytic function f, $X_g f(v) = X(f \circ \Phi_g)(v) = Xf(g \cdot v)$ which gives

$$\xi^i \frac{\partial f(g \cdot v)}{\partial v_i}\bigg|_{v=g_o} = \xi^i \left[\frac{\partial f(g \cdot v)}{\partial (g \cdot v)_k} \frac{\partial (g \cdot v)_k}{\partial v_i}\right]_{v=g_o} = \xi^i \Phi_g^{k,i}(g_o) \left.\frac{\partial f(g \cdot v)}{\partial (g \cdot v)_k}\right|_{v=g_o}$$

which as a tangent vector $\xi^i \Phi_g^{k,i}(g_o) \frac{\partial}{\partial(g \cdot v)}$ is an element an element of $T_{g \cdot g_o} G$. We get a function between tangent spaces defined by this "infinitesimal" right translation. We can map any tangent space $T_{g_o} G$to $T_e G$ by defining a map ω such that

$$\omega(X) = X_{g_o^{-1}} \in T_{g_o \cdot g_o^{-1}} G = T_e G$$

so that we have "pulled back" the tangent vector to lie in the tangent space to the identity. Suppose X were a right-invariant vector field; then ω must map every tangent vector to the same element, namely to $\mathsf{X}(e)$. But given a tangent vector at the identity, we can construct a right-invariant vector field by right-translating it, so these vector fields are in one-to-one correspondence with elements of $T_e G$. Further, the map ω preserves the commutator of two right-invariant vector fields:

$$\omega([\mathsf{X}, \mathsf{Y}]_i) = \omega(\mathsf{X}_i \frac{\partial \mathsf{Y}_j}{\partial x_i} - \mathsf{Y}_i \frac{\partial \mathsf{X}_j}{\partial x_i}) = \mathsf{X}(e) \frac{\partial \mathsf{Y}(e)_j}{\partial x_i} - \mathsf{Y}(e)_i \frac{\partial \mathsf{X}(e)_j}{\partial x_i} = [\omega(\mathsf{X}), \omega(\mathsf{Y})]_i$$

.

We have proved:

Proposition 1. *The tangent space at the identity $T_e G$ is isomorphic to the space of right-invariant vector fields.*

To see that T_eG encodes many desired properties of G, we will use this isomorphism to move from elements of the tangent space to the manifold itself.

2.2. **The Exponential Map.** Given $X \in T_eG$ let X be the associated right-invariant vector field, i.e. $\mathsf{X}(e) = X$. Then the results of differential equations allow us to integrate over X, getting a map $\phi_X : U \subset \mathbb{R} \to G$ such that $\phi_X(0) = g_o$ for some chosen point g_o and $\phi'_X(t) = \mathsf{X}(\phi(t))$ for $t \in U$, where U is some open ball around the origin. Essentially, we are solving the differential equation

$$\frac{d\phi_X}{dt} = \mathsf{X}(\phi_X(t))$$

given $\phi_X(0) = g_o$, so we are using the principle of least action and "following the tangent vectors." By the invariance of X and the definition of ϕ_X it follows that $\phi_X(s+t) = \phi_X(s)\phi_X(t)$ so ϕ_X is a homomorphism: Fix s and define $\alpha_X(t) = \phi_X(s+t)$, $\beta_X(t) = \phi_X(s)\phi_X(t)$; take the differential of both sides to see that, by invariance, $\mathsf{X}(\alpha(t)) = \alpha'(t)$ and $\mathsf{X}(\beta(t)) = \beta'(t)$. Since $\alpha(0) = \beta(0)$ their differentials must be the same at the origin, so by the uniqueness of the integral curve they are equal for all t. Then ϕ_X is a homomorphism of Lie groups and since it is defined around the origin, it extends uniquely to all of $\mathbb{R}^n$. As there is one such map for each tangent vector X , one might assume that they will somehow be useful. They do get their own special name:

Definition 3. The homomorphisms $\phi_X : \mathbb{R} \to G$ are the *one-parameter subgroups of G.*

Now let $g_o = e$ and define a map that will take elements of T_eG to the manifold:

Definition 4. The *exponential map* of a Lie group G is the map sending elements $X \in T_eG$ to the corresponding one-parameter subgroup of G.

$$\exp : X \mapsto \phi_X(1).$$

If the Lie group is a real or complex matrix group, the 'usual' properties of the exponential hold (see below). However, for any group, the exp map gives us the useful result

Proposition 2. *The map* exp *is natural, i.e. for any homomorphism of groups* Ψ*with differential* $\psi : T_eG \to T_eG$ *the following diagram commutes:*

$$\begin{array}{ccc} G & \xrightarrow{\Psi} & H \\ exp \uparrow & & exp \uparrow \\ T_{e_G}G & \xrightarrow{\psi} & T_{e_H}H \end{array} .$$

Proof. Let $X \in T_{e_G}G$ and take ϕ_X; then $\Psi \circ \phi_X : \mathbb{R} \to H$ is the one-parameter subgroup corresponding to $\psi(X) \in T_{e_H}H$, so by definition $\exp(\psi(X)) = (\Psi \circ \phi_X)(1) = \Psi \circ \exp(X)$. ☐

Note that this means exp lifts the tangent space to give a neighborhood of the identity in G; if G is connected then we generate the whole group from this neighborhood. The connectedness of Lie groups will be important for the construction of the Lie algebra (see Theorem 1 below). If our Lie group is indeed a real or complex matrix group, the exponential can be written

$$\exp(X) = I + X + \frac{X^2}{2} + \frac{X^3}{6} + \cdots$$

as with the 'usual' definition of exponential.[1] This will always converge; if we are dealing with an $n \times n$ matrix X whose largest value is m then the elements of X^i will be at most $\left((n^{i-1}m^i)\right)$ and since $m\sum_{i=1}^{\infty} \frac{(nm)^{i-1}}{i!}$ will converge for any number nm, the infinite sum $I + X + \frac{X^2}{2} + \cdots$ will always converge. One can additionally check that if $[X, Y] = 0$ then $\exp(X)\exp(Y) = \exp(X + Y)$ on a sufficiently small neighborhood of the identity. Because exp is injective on a neighborhood of the identity, we can introduce an inverse on this neighborhood. It will map from a subset of $GL_n(\mathbb{R})$ to the tangent space $T_e GL_n(\mathbb{R})$:

Definition 5. Define on a neighborhood of $e \in GL_n(\mathbb{R})$ the map

$$\log : g \mapsto (g - I) - \left(\frac{(g-I)^2}{2}\right) + \left(\frac{(g-I)^3}{6}\right) - \cdots$$

and on a neighborhood of 0 in $T_e GL_n(\mathbb{R})$ the binary operation $*$ such that

$$X * Y = \log\left(\exp(X) \cdot \exp(Y)\right).$$

This is the *Campbell-Hausdorff Formula.*

Careful computation gives the first few terms of the Campbell-Hausdorff formula:

$$X * Y = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] + \cdots$$

The Campbell-Hausdorff formula provides a way to expand the binary operation above without the use of exp or log, i.e. it relies only on the algebraic relations of the elements of the tangent space. We observe that this series always converges, so that $\exp(X) \cdot \exp(Y)$ must always be in the Lie group if X, Y are in the tangent space. The formula shows how $T_e G$ encodes the local group structure of the manifold while simplifying the topology via a "linearization."

As we said earlier, finding homomorphisms of groups is our main goal, and we employed right translation to aid us in our search. Before we can completely characterize homomorphisms of Lie groups, we prove a lemma.

[1] Since $\mathbb{R}$ is itself a Lie group, the 'usual' exponential might be seen as a special case of the map described here.

Lemma 1. *Let G be a real matrix Lie group with $\mathfrak{h} \leq T_e G$ a subspace of the tangent space at the identity of G. Then the subgroup of G obtained via $\exp(\mathfrak{h})$ is an immersed[2] subgroup H with tangent space $T_e H = \mathfrak{h}$.*

Proof. Let $h \in H$ and consider the map $\rho_g : H \to T_e H$ taking $y \mapsto \log(g^{-1} \cdot y)$. This gives an atlas of local coordinatization of elements of H because log is defined only on a neighborhood of the identity. Hence only those elements for which g^{-1} is sufficiently near y will map under ρ_g to the tangent space, and so we have identified a neighborhood of y in G. □

Now we can prove the final result, showing a correspondence between homomorphisms of Lie groups and homomorphisms of their tangent spaces at the identity.

Theorem 1. *Let G, H be Lie groups with G simply connected. Let $T_{e_G}G, T_{e_H}H$ be the corresponding tangent spaces. Then $\Psi : G \to H$ is a (differentiable) group homomorphism iff $d\Psi_{e_G} = \psi : T_{e_G}G \to T_{e_H}H$ is a differentiable homomorphism of the tangent spaces (that is, it preserves the commutator).*

Proof. Consider the Lie group $G \times H$, which has tangent space $T_{e_G}G \times T_{e_H}H$. Then if ψ is a homomorphism as hypothesized, the graph j of ψ is a subspace of $T_{e_G}G \times T_{e_H}H$. By Lemma 1, this would imply that we have an immersed subgroup $J \subset G \times H$ with tangent space $T_{e_{G\times H}}J$. Consider the projection map on the first factor $\pi_1 : J \to G$, which has differential $d\pi_1 : j \to T_e G$ that is an isomorphism by hypothesis. But since G is simply connected, π_1 is also an isomorphism. Hence the projection on the second factor $\pi_2 : J \to H$ must be a homomorphism with differential at the identity equal to ψ. The other direction follows from the diagram above; that is, if Ψ is a homomorphism then its differential is also. □

We used the fact that G is simply connected in order to show that the map Ψ is defined on all of G while preserving its differentiability. Hopefully it is clear now that the tangent space at the identity is quite important, because as Theorem 1 shows, any representation of a Lie group will induce a representation of the tangent space. Although we have seen before now that the tangent space has the structure of an algebra (bilinearity in its commutator), its importance was not illuminated until now, and so we finally give the tangent space a special name.

Definition 6. The tangent space at the identity of a Lie group, $T_e G$, is called the *Lie algebra* of the group and is labeled $\mathfrak{g}$.

Hence we can now name the Lie algebras of the classical Lie groups: for the special linear group $SL(n)$ we have $\mathfrak{sl}_n$, for the symplectic group $Sp(2n)$ we have $\mathfrak{sp}_{2n}$, etc.

Note that we may axiomatize Lie algebras:

[2] The term *immersed* borrows from [6] and corresponds to what is normally termed a Lie subgroup, not necessarily closed.

Definition 7. A *Lie algebra* is a vector space $\mathfrak{g}$ over a field F with a map $[\cdot,\cdot] : \mathfrak{g} \to \mathfrak{g}$ satisfying three axioms:

1. Bilinearity: $[aX + bY, Z] = a\,[X, Z] + b\,[Y, Z]$
2. Anticommutativity: $[X, Y] = -\,[Y, X]$
3. Jacobi identity: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$

Property 3 was not verified for the tangent space of a Lie group, but it holds. Definition 7 would imply that perhaps there exist Lie algebras without an associated Lie group - indeed there do, but they must be infinite-dimensional. On the other hand, we could construct an 'abstract' Lie algebra by taking generators of any vector space such as a Lie group and applying commutation relations to them to determine what the commutators are. This is the method we use in Section 3 to develop the Eight-Fold Way. Other instances of a Lie algebra being developed 'from scratch' rather than from a Lie group would be the Heisenberg group:

$$\mathfrak{h}_3 = \langle x, y, z : [x, y] = z, [x, z] = 0 = [y, z]\rangle$$

which has for elements the upper-triangular matrices with diagonal elements equal to 1. This has additional applications to quantum mechanics.

Now that Lie algebras have been developed and their relationships to Lie groups and Lie subgroups have been discussed,[3] we describe an example in which representations of a Lie algebra prove useful in examining a Lie group. Lie algebras will be important in the following exposition because they are endowed with a commutator. Lacking any other information about how to find a representation, we know that any representation of a Lie group must preserve the commutation relations of the Lie algebra, so we can construct an 'abstract' Lie algebra using the generators of the Lie group and find the commutation relations between them, thereby gaining some structure for the action of the representation.

3. The Eight-Fold Way

3.1. **Historical Background.** One common method for predicting the potential for certain interactions to occur in physics is to determine what observable quantities must be conserved during the reaction; if a given interaction conserves all of these quantities, then it is considered "possible" and can

[3]Note: One may equivalently develop Lie groups using the adjoint representation of a Lie group. This is perfectly equivalent to the method outlined above but seems rather artificial because one doesn't know *a priori* that the adjoint operator will be useful. Hence, examining the tangent space, which is a very normal endeavor when dealing with manifolds, seems to be the most natural method of describing the Lie group, and is also closer to what Sophus Lie originally did.

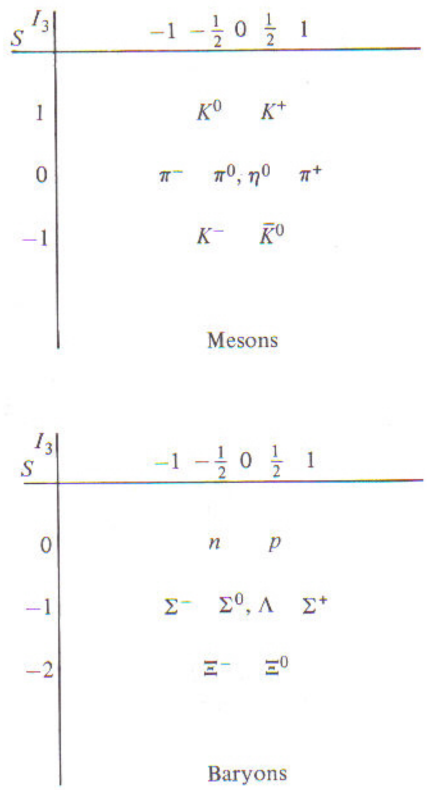
be looked for in an experiment in order to confirm the theory. Most people are familiar with the concepts of conservation of energy and momentum, but there are other quantities that must be conserved, depending on the reaction. Two of these observable are known as isospin and strangeness. As science began to uncover the inner structure of the atom, it became clear that some force must be binding the protons and neutrons together in the nucleus. This force must be stronger over short distances than the electromagnetic force but ultimately weaker at far distances (where "far" can mean billionths of a meter). This so-called "strong" force is charge-independent, since neutron and proton do not have the same charge but act the same under the force. Heisenberg postulated in 1932 that these two particles are just different states of a single particle, the nucleon. In fact, this would imply that they could be transformed from one to the other as matrices might be transformed under the action of a symmetry group. Heisenberg assumed that the associated group was $SU(2)$ and that the two eigenstates of the nucleon were $|n\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $|p\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (the similarity to the spin-up/spin-down states of an electron caused this nucleon matrix group to be known as isospin, with each particle having an observable "isospin," and the total isospin must be conserved in any interaction).[4] Note that the choice of $SU(2)$ was probably due to the fact that there are generators of $SU(2)$ which take $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and vice-versa. In addition, a few other particles involved in the strong interaction (baryons, including the proton and neutron, and mesons, such as pions and kaons) must be related by matrix relations of the same group, as they are involved in interactions with each other. Then the transformation between the states of the nucleon would be equivalent to the action of the two-dimensional representation of $SU(2)$. We won't go into the details of this representation - finding generators and so forth - because some of the reactions that were hypothesized to occur were not found in high-energy accelerator experiments. Such interactions conserved all the supposedly necessary observables required by the strong interaction. Because these reactions did not occur, Murray Gell-Mann and Nishijima hypothesized that there must be an additional observable quantity which these "unseen" reactions did not conserve. Because this quantity was theretofore unknown, they called it "strangeness." By normalizing the strangeness of the proton and neutron to zero, the strangeness of other particles could be deduced from their production or decay reactions. A new quantum number, the hypercharge, was introduced and related to the strangeness S and the baryon number B:

$$Y = S + B$$

.

[4]The $|\cdot\rangle$"bra-ket" notation is a product of Paul Dirac and indicates that the corresponding matrix represents a particle's state, in this case the state of the neutron and proton, respectively.

FIGURE 1. Quantum Numbers of Mesons and Baryons [7]



After trying alternatives, a model for the group of strong interactions was developed which described the reactions as seen in experiment. It is called the "flavor" model and is the foundation for the quark model of subatomic particles. By introducing these unobservable states called "quarks," observed mesons and baryons could be built. This all corresponds to a representation of the group $SU(3)$, which was chosen over the smaller group $SU(2)$ to account for the new quantum number Y. Diagrams displaying the quantum numbers I_3 (isospin) and Y are shown in Figure 1: the vertical axis shows the isospin and the horizontal axis shows the strangeness, with the particles' names written at the intersection of their quantum numbers.

3.2. **The Mathematics.** $SU(3)$ is the space of unitary 3×3 matrices with determinant 1. These are all traceless, as can be seen thus:

This space has dimension eight, and the standard example of a basis is the set of Gell-Mann λ-matrices:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(the Pauli matrices, which generate $SU(2)$, with extra row of zeroes)

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Without actually finding the corresponding Lie algebra, we can construct an "abstract" version by implementing the fact that elements of the Lie algebra obey the commutation relation (in fact, the Lie algebra of $SU(3)$ is isomorphic to that of $SL_3(\mathbb{C})$, so we can call it $\mathfrak{sl}_3(\mathbb{C})$). Out of eventual convenience (and partly to reflect the structure of the basis of $SU(2)$ that we didn't consider) we define some new matrices (the names are historical artifices):

$$T_x = \frac{1}{2}\lambda_1, T_y = \frac{1}{2}\lambda_2, T_z = \frac{1}{2}\lambda_3, V_x = \frac{1}{2}\lambda_4, V_y = \frac{1}{2}\lambda_5, U_x = \frac{1}{2}\lambda_6, U_y = \frac{1}{2}\lambda_7, Y = \frac{1}{\sqrt{3}}\lambda_8$$

and

$$T_\pm = T_x \pm iT_y, V_\pm = V_x \pm iV_y, U_\pm = U_x \pm iU_y \text{ (the so} - \text{called "shift'' operators).}$$

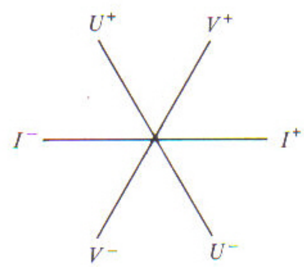
The names should give some hint of what the matrices mean; Y, for instance, corresponds to the hypercharge and accordingly, T_z, the only other matrix not encoded in a "$\pm$" relation, corresponds to isospin (note that these are the only two diagonalizable matrices, and so the only ones with eigenvectors). The "convenience" comes when we consider the commutation relations (in Table 1), so that all commutators are just scalar multiples of other matrices in the basis. We construct an abstract Lie algebra by assuming elements t_x, t_y, t_z, etc. whose commutators are the same as the commutators of the matrices above.

We see that y and t_z commute, so they are simultaneously diagonalizable; hence we might write simultaneous eigenvectors as particle states in the Dirac notation $|\lambda, \mu\rangle$ where λ, μ are eigenvalues of the matrices on that vector. Then the notation $|\lambda, \mu\rangle$ would correspond to the state of a particle with isospin λ and strangeness μ. By examination we see that T_+ shifts the eigenstate to $|\lambda + 1, \mu\rangle$, U_+ shifts it to $\left|\lambda - \frac{1}{2}, \mu + 1\right\rangle$, etc. so that the shift operators (hence the reason for the subscripts) form a diagram with vectors $T_+ = (1, 0)$, $U_+ = (-\frac{1}{2}, 1)$, etc. The shift operators, then, actually transform *between particles*, which are themselves *eigenstates of the isospin and strangeness operators.* Figure 2 shows the six vectors (note that in the text from which this came, the operators T are called I; someone

TABLE 1. Commutation Relations for Elements of Abstract Lie Algebra

	$\mathbf{t}_+$	$\mathbf{t}_-$	$\mathbf{t_z}$	$\mathbf{v}_+$	$\mathbf{v}_-$	$\mathbf{u}_+$	$\mathbf{u}_-$	$\mathbf{y}$
$\mathbf{t}_+$	0	$2t_z$	$-t_+$	0	$-u_-$	v_+	0	0
$\mathbf{t}_-$	$-2t_z$	0	t_-	u_+	0	0	$-v_-$	0
$\mathbf{t}_z$	t_+	$-t_-$	0	$\frac{1}{2}v_+$	$-\frac{1}{2}v_-$	$-\frac{1}{2}u_+$	$\frac{1}{2}u_z$	0
$\mathbf{v}_+$	0	$-u_+$	$-\frac{1}{2}v_+$	0	$\frac{3}{2}y+t_z$	0	t_-	$-v_+$
$\mathbf{v}_-$	$-u_-$	0	$\frac{1}{2}v_-$	$-\frac{3}{2}y-t_z$	0	$-t_-$	0	v_-
$\mathbf{u}_+$	$-v_+$	0	$\frac{1}{2}u_+$	0	t_-	0	$\frac{3}{2}y-t_z$	$-u_+$
$\mathbf{u}_-$	0	v_-	$-\frac{1}{2}u_z$	$-t_-$	0	$-\frac{3}{2}y+t_z$	0	u_-
$\mathbf{y}$	0	0	0	v_+	$-v_-$	u_+	$-u_-$	0

FIGURE 2. Root Diagram for Shift Operators(A_2) [7]



with previous background might recognize this chart as the root diagram A_2, which corresponds to the Lie algebra we are considerinig). A representation of this Lie algebra would be any map from the generators t_+, t_-, etc. to a vector space while preserving the commutator relations.

Beginning with the standard representation from Example 1 above, with the vector space $V = \mathbb{C}^3$, we have basis vectors $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. The Lie group acts on these by multiplying on the left. We (suggestively) write these $e_1 = u\,(\text{up quark})$, $e_2 = d\,(\text{down quark})$, $e_3 = s\,(\text{strange quark})$. Using the definition of the dual representation as stated in Example 2 above, the basis vectors are $u^* = \begin{pmatrix} -1 & 0 & 0 \end{pmatrix}$, $d^* = \begin{pmatrix} 0 & -1 & 0 \end{pmatrix}$, $s^* = \begin{pmatrix} 0 & 0 & -1 \end{pmatrix}$ which are anti-up, anti-down, and anti-strange. The Lie group acts on these by multiplying on the right by a matrix. For example,

$$u^* V_+ = \begin{pmatrix} -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 \end{pmatrix} = s^*$$

.

FIGURE 3. Weight Diagram for Standard Representation (V) [7]

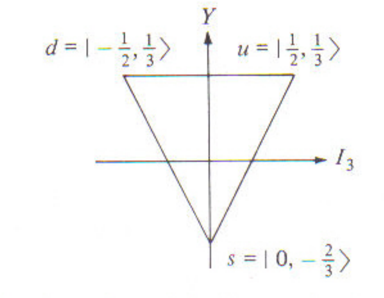
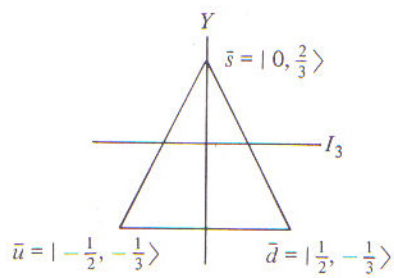


FIGURE 4. Weight Diagram for Dual Representation (V^*) [7]



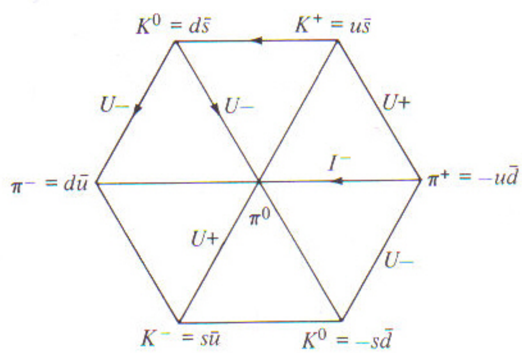
Then taking the tensor product $V \otimes V^*$ yields a representation with 9 generators (hence 9-dimensional). Given the Lie algebra relations in Table 1 and the shift operator relations in Figure 2, we can determine the eigenstate corresponding to each vector u, d, s, etc. by finding their eigenvalues under the action by T_z and Y. For instance, $T_z u = \frac{1}{2}u$ and $Yu = \frac{1}{3}u$ so $u = \left|\frac{1}{2}, \frac{1}{3}\right\rangle$ as an eigenstate. The diagram formed by charting these is called a *weight diagram*; the weight diagrams for the standard and dual representations are shown in Figures 3 and 4.

We can see by comparison to Figure 1 that there are no particles corresponding to these eigenvectors; hence, particles and quarks are distinct and the particles are actually *tensor products of quarks.* To actually find the elements of the tensor product, we compute them in the normal way: if L_1, L_2 are generators for two representations V_1, V_2 then $L(v_1 \otimes v_2) = L_1(v_1) \otimes v_2 + v_1 \otimes L_2(v_2)$. For example,

$$T_z(ud^*) = T_z(u)d^* + uT(d^*) = \frac{1}{2}ud^* + \frac{1}{2}ud^* = 1 \cdot ud^*$$

$$Y(ud^*) = Y(u)d^* + uY(d^*) = 0$$

FIGURE 5. The Meson Octet [7]



so ud^* corresponds to the state $|1,0\rangle$ which, as one can see from Diagram 1, is the same as that of the pion π^+. Further, we can get relationships using the shift operators. For example,

$U_+ud^* = us^* = \left|\frac{1}{2},1\right\rangle$ which is the kaon K^+.

From these relationships, we can generate every meson; however, the particle $\eta = \frac{uu^*+dd^*+ss^*}{\sqrt{6}} = |0,0\rangle$ is invariant under the action (one can check that $uu*$, dd^*, ss^* go to zero under the shift operators) so it cannot be transformed into any other particle. Hence the representation $V \otimes V^*$ reduces to the direct sum of two irreducible representations of degrees 8 and 1.[5] The tensor product of the standard representation and its dual, then, accurately explain the observed interactions of mesons, and the group of eight mesons generated by the shift operators gives the Eight-Fold Way its name. The meson octet is shown in Figure 5 (by historical convention, $\pi^+ = -ud^*$ instead of ud^* as we determined).

3.3. **Physical Re-Interpretation.** As stated above, quarks are the basis vectors for the standard representation of $SU(3)$ and anti-quarks are the basis vectors for the dual representation (although today there are six quarks in the model of the universe). The tensor product has generators consisting of quark-anti-quark pairs, and these correspond to the mesons. Therefore, the elementary particles are the generators for a vector space into which there is a homomorphism from a Lie group. That Lie group has eight elements for generators; two of them have the particles as eigenvectors, and we can uniquely identify the particles from their corresponding two eigenvalues. The other six generators of the Lie group act on the mesons by transforming them, turning them into other mesons. Therefore, one might say that *all mesons are basis vectors for* $\mathbb{C}^3 \otimes \mathbb{C}^{3*}$ *and meson interactions are characterized by the action of linear combinations of generators of the Lie group* $SU(3)$. Why this should be - why

[5]Actually, the representation of degree 8 is the adjoint representation mentioned in footnote 2; one may actually begin with this representation as teh basis for the eight-fold way, but it makes more sense to me to start with the standard and dual representations, as their actions will end up corresponding to the actions of quarks and antiquarks, respectively.

Lie groups should be woven into the background of the universe's fabric - is a question I simply can't answer.

We have shown that all mesons are composed of one or more quark-antiquark pairs. The baryons (protons, neutrons, etc.), it turns out, correspond to the triple tensor product representation $V \otimes V \otimes V$ so each of them have 3 quarks (the proton, for instance, is up-up-down). This splits into a 10-dimensional irreducible representation, two 8-dimensional ones (the same as the 8-dimensional one above) and a 1-dimensional representation. The benefit of this mathematical model is that given enough information from the mathematics, one can tell when the physics is incomplete. In 1964, using a theoretical prediction from the representation theory used here, the baryon Ω^- was discovered. While the number of known elementary particles has grown to several hundred, some are still only theoretical predictions. That the mathematics of homomorphisms between groups should help in the endeavor to ascertain the structure of the universe is exciting and gives a unique perspective on just what the interactions between matter really are.

Lie algebras were indispensable here because, although we did all our work with the matrices of the Lie group itself, we used the fact that the Lie algebra is endowed with a commutator, and since homomorphisms between groups must correspond to homomorphisms between algebras, we use the information of the commutators to guarantee that we are finding workable representations of the groups at hand. Indeed, in this case one might not even know what vector space V to use as a representation; what ended up happening was that the Lie algebra itself acted as that vector space, as we found an 8-dimensional representation corresponding to a map into the Lie algebra.[6] Hence the algebraic structure of Lie algebras are a vital tool in the study of representations of Lie groups.

References

[1] Adams, J. Frank. *Lectures on Lie Groups*. Chicago: University of Chicago Press, 1969.

[2] Cahn, Robert N. *Semisimple Lie Algebras and Their Representations*. Accessed online: http://www-physics.lbl.gov/~rncahn/book.html.

[3] Cohn, P.M. *Lie Groups*. Cambridge: Cambridge University Press, 1957.

[4] Duistermat, J.J. and J.A.C. Kolk. *Lie Groups*. New York: Springer-Verlag, 2000.

[5] Dummit, David S. and Richard M. Foote. *Abstract Algebra*. Hoboken, NJ: John Wiley and Sons, Inc., 2004.

[6] Fulton, William and Joe Harris. *Representation Theory: A First Course*. New York: Springer Verlag, 1991.

[7] Sattinger, D. H. and O. L. Weaver. *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*. New York: Springer-Verlag, 1986.

[6] A greater elaboration on the adjoint representation would demonstrate immediately that this is true. In short, if Θ_g is conjugation by $g \in G$ then $\mathrm{Ad}(g) = (d\Theta_g)_e$ is a map from the Lie group to the space of automorphisms of its Lie algebra, $\mathrm{Aut}(\mathfrak{g})$.