Lie theory and the eightfold way

John Huerta

CAMGSD
Departamento de Matemática
Instituto Superior Técnico
Universidade de Lisboa
john.huerta@tecnico.ulisboa.pt

20th January 2025

Abstract

This short expository note aims to give the minimal amount of Lie theory needed to appreciate the eightfold way in particle physics. We first give a quick but complete account of the finite-dimensional irreducible representations (irreps) of $\mathfrak{sl}(2,\mathbb{C})$. Then we sketch how the theory generalizes to the irreps of $\mathfrak{sl}(3,\mathbb{C})$, and close by gesturing at the role of these irreps in the eightfold way.

1 Introduction

Consider the space of traceless, 2×2 complex matrices, which we denote $\mathfrak{sl}(2,\mathbb{C})$:

$$\mathfrak{sl}(2,\mathbb{C}) = \left\{ \begin{pmatrix} a & b \\ c & -a \end{pmatrix} : a, b, c \in \mathbb{C} \right\}.$$

This is a vector space, but more importantly, it is a Lie algebra—it is closed under the Lie bracket, defined in this case to be the commutator:

$$[X, Y] = XY - YX, \text{ for } X, Y \in \mathfrak{sl}(2,\mathbb{C}).$$

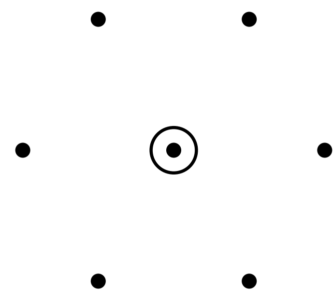
Lie algebras, like their cousins Lie groups, encode *symmetries*. In the case of $\mathfrak{sl}(2,\mathbb{C})$, for instance, each element acts on the vector space $\mathbb{C}^2$ as a linear transformation, and we think of such a transformation as a symmetry of $\mathbb{C}^2$. This situation, where a Lie algebra acts on a vector space, is the subject of Lie theory. The vector space equipped with such an action is called a 'representation'.

The representations of $\mathfrak{sl}(2,\mathbb{C})$ on finite-dimensional vector spaces are startling in their simplicity. Though there are infinitely many such representations, each one can be decomposed into simpler pieces, called the 'irreducible representations', or irreps, and each irrep is uniquely determined by the choice of a single natural number. Curiously, we will see that it is useful to describe the irrep associated to the natural number m with a diagram consisting of $2m + 1$ dots arranged in a line, labeled from $-m$ to m in increments of 2:

$$\begin{array}{ccccc} \bullet & \bullet & \cdots & \bullet & \bullet \\ -m & -m+2 & & m-2 & m \end{array}$$

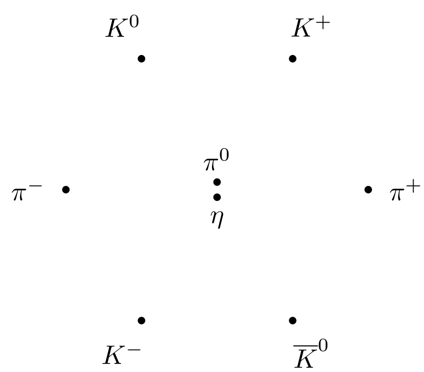
This picture is our first hint that irreps are surprisingly discrete, being determined by points in a lattice, here just the integers. They also exhibit some hidden symmetry, which we see here by the symmetry between m and $-m$.

Passing to the next case of $\mathfrak{sl}(3,\mathbb{C})$, the special linear Lie algebra on $\mathbb{C}^3$, reveals even more structure. Again, the irreducible representations, or irreps, are determined by points in a lattice with some symmetry:



Now, however, the lattice is two-dimensional, and the symmetry group is bigger, generated by all the reflection symmetries manifest in the picture. The circle around the center dot tells us to count it with multiplicity two.

This representation theory is pure linear algebra, but it magically appears in particle physics, where it is part of the 'eightfold way'. In the middle of the 20th century, experiments where physicists collided particles together at high energies produced scores of a new particles beyond the familiar electron, proton and neutron that constitute atoms. Searching for order in the chaos, physicists discovered that these new particles could be organized into representations of $\mathfrak{sl}(3,\mathbb{C})$:



1.0.1 Acknowledgments

I thank Gonçalo Oliveira for his encouragement and patience. As a member of CAMGSD, this work was funded by FCT/Portugal through project UIDB/04459/2020. This article was written while visiting the Mathematical Sciences Institute at the Australian National University, and I thank the MSI for their hospitality.

2 Lie algebras

To get started, let us give some precise definitions of the objects we want to study, namely Lie algebras and their representations. Although we could work over any field, we choose to work over the complex numbers, $\mathbb{C}$. Thanks to $\mathbb{C}$ being algebraically closed, every matrix has an eigenvalue. Since computing eigenvalues and diagonalizing matrices will play an essential role in our analysis, it pays to work over $\mathbb{C}$.

Definition 1. A **Lie algebra** $\mathfrak{g}$ is a complex vector space equipped with a bilinear operation called the Lie bracket

$$[-,-]\colon \mathfrak{g}\otimes\mathfrak{g}\to\mathfrak{g},$$

satisfying the following axioms:

- skew-symmetry: $[X,Y]=-[Y,X]$ for all $X,Y\in\mathfrak{g}$;
- the Jacobi identity: $[X,[Y,Z]]=[[X,Y],Z]+[Y,[X,Z]]$, for all $X,Y,Z\in\mathfrak{g}$.

Example 2. We have already met two examples of our favorite Lie algebra in the Introduction, namely the **special linear Lie algebra** of traceless $n\times n$ complex matrices:

$$\mathfrak{sl}(n,\mathbb{C})=\left\{X\in \mathrm{Mat}_{n\times n}(\mathbb{C})\ :\ \mathrm{tr}(X)=0\right\},$$

where the Lie bracket is given by the commutator:

$$[X,Y]=XY-YX,\quad X,Y\in\mathfrak{sl}(n,\mathbb{C}).$$

It is a worthwhile if somewhat tedious exercise to check that the Jacobi identity holds.

Example 3. Dropping the condition on the trace, we have the **general linear Lie algebra** of all $n\times n$ complex matrices:

$$\mathfrak{gl}(n,\mathbb{C})=\mathrm{Mat}_{n\times n}(\mathbb{C}).$$

Again, the Lie bracket is given by the commutator.

Example 4. More abstractly, for any complex vector space V, we have the **general linear Lie algebra on V**, consisting of all linear maps:

$$\mathfrak{gl}(V)=\left\{T\colon V\to V\ :\ T \text{ linear}\right\}.$$

Once again, the Lie bracket is given by the commutator. Of course, fixing a basis of V gives us an isomorphism of Lie algebras, $\mathfrak{gl}(V)\cong\mathfrak{gl}(n,\mathbb{C})$, where $n=\dim(V)$.

Definition 5. Let $\mathfrak{g}$ be a Lie algebra. A **representation** of $\mathfrak{g}$ is a pair (V,ρ) where V is a finite-dimensional complex vector space, and $\rho\colon\mathfrak{g}\to\mathfrak{gl}(V)$ is a homomorphism of Lie algebras. Explicitly, this means that ρ is a linear map such that

$$\rho\left([X,Y]\right)=\rho(X)\rho(Y)-\rho(Y)\rho(X),\ \text{for all } X,Y\in\mathfrak{g},$$

since the bracket on $\mathfrak{gl}(V)$ is the commutator.

Example 6. Every Lie algebra $\mathfrak{g}$ has a god-given representation on itself, called the **adjoint representation**, $\mathrm{ad}\colon\mathfrak{g}\to\mathfrak{gl}(\mathfrak{g})$. An element $X\in\mathfrak{g}$ acts on $Y\in\mathfrak{g}$ by bracketing: $\mathrm{ad}(X)Y=[X,Y]$.

It is a useful exercise to check that this is a representation. The Jacobi identity will play a key role.

3 The representations of $\mathfrak{sl}(2,\mathbb{C})$

We will study the representations of the complex special linear Lie algebra, $\mathfrak{sl}(n,\mathbb{C})$. In fact, we are mainly interested in $\mathfrak{sl}(3,\mathbb{C})$, because of the role it plays in particle physics. But to get started, we need to study $\mathfrak{sl}(2,\mathbb{C})$.

Definition 7. A **Cartan subalgebra** $\mathfrak{h}\subseteq\mathfrak{sl}(n,\mathbb{C})$ is a maximal abelian subalgebra such that the adjoint action $\mathrm{ad}(H)$ on $\mathfrak{sl}(n,\mathbb{C})$ can be diagonalized for all $H\in\mathfrak{h}$.

Example 8. Let $\mathfrak{h}$ be the diagonal matrices in $\mathfrak{sl}(n, \mathbb{C})$. This subalgebra is:

- abelian, because diagonal matrices commute;
- maximal, because additional elements would have off-diagonal entries and no longer commute;
- $\operatorname{ad}(H)$ is diagonalizable for all $H \in \mathfrak{h}$.

Let us check this last claim: let H be the diagonal matrix with entries $a_1, \ldots, a_n$ on the diagonal, let E_{ij} be the elementary matrix with 1 in the ijth entry and zeroes elsewhere. A quick computation shows that $[H, E_{ij}] = (a_i - a_j)E_{ij}$. For $i \neq j$ (why?), this shows that E_{ij} is an eigenvector of $\operatorname{ad}(H)$ with eigenvalue $a_i - a_j$, and we can thus write down a basis of eigenvectors for $\mathfrak{sl}(n, \mathbb{C})$. This diagonalizes $\operatorname{ad}(H)$.

For $\mathfrak{sl}(2, \mathbb{C})$, there's a one-dimensional Cartan subalgebra $\mathfrak{h} = \operatorname{span}(H)$, spanned by the element $H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. This matrix is part of the standard basis for $\mathfrak{sl}(2, \mathbb{C})$:

$$E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

These matrices satisfy the important relations:

$$[H, E] = 2E, \quad [E, F] = H, \quad [H, F] = -2F.$$

Definition 9. Let (V, ρ) be a representation of $\mathfrak{g}$. A subspace $W \subseteq V$ is called **invariant** under $\mathfrak{g}$ if $\rho(X)w \in W$ for all $X \in \mathfrak{g}$ and $w \in W$. A representation V of $\mathfrak{g}$ is **irreducible** if the only invariant subspaces of V are 0 and V. A representation V is called **completely reducible** if it is the direct sum of irreducible representations, i.e., $V = \bigoplus_\lambda V_\lambda$, where each V_λ is irreducible.

To help us analyze the representations of $\mathfrak{sl}(2, \mathbb{C})$, we note the following without proof:

Theorem 10. *All complex, finite-dimensional representations of $\mathfrak{sl}(n, \mathbb{C})$ are completely reducible.*

This theorem says we can focus on the irreducible representations, or **irreps**, since all others arise by taking direct sums. From now on, we assume V is an irreducible, finite-dimensional representation over $\mathbb{C}$. The key result for understanding V is a bit deep, and we give it without proof:

Theorem 11. *Given any complex finite-dimensional representation (V, ρ) of $\mathfrak{sl}(2, \mathbb{C})$, $\rho(H)$ is diagonalizable.*

This should be plausible, since H itself is a diagonal matrix, and we already know that $\operatorname{ad}(H)$ is diagonalizable. The remarkable thing is that $\rho(H)$ is diagonalizable for any ρ. We use this diagonalizability as follows: decompose the irrep V into a direct sum of eigenspaces

$$V = \bigoplus_\lambda V_\lambda,$$

where the direct sum is over all complex numbers λ which are eigenvalues of $\rho(H)$, and each summand V_λ is an eigenspace for λ. In other words, for all $v \in V_\lambda$, we have $Hv = \lambda v$. (Really, $\rho(H)v = \lambda v$, but it is standard to suppress ρ.)

In Lie theory, the eigenvalues occurring here have a special name: they are called the **weights** of V. The eigenspaces V_λ are called **weight spaces**, and an eigenvector $v \in V_\lambda$ is called a **weight vector**. Given the decomposition of V into weight spaces, we thus know how H acts—it acts diagonally, by multiplication by the corresponding weight. Next, we would like to determine how the other basis elements E and F act:

Proposition 12. *If $v \in V_\lambda$, then $Ev \in V_{\lambda+2}$ and $Fv \in V_{\lambda-2}$. (Really, $\rho(E)v$ and $\rho(F)v$, but we're continuing to suppress ρ.)*

The proof of this proposition is so important, that we call it the **fundamental calculation**: fix $v \in V_\lambda$, and compute

$$HEv = EHv + [H, E]v = \lambda Ev + 2Ev = (\lambda + 2)Ev.$$

Similarly, $HFv = (\lambda - 2)Fv$. This is what we wanted to check.

So, we have arrived at the following picture of V:

$$V = \cdots \oplus V_{\lambda-2} \oplus V_\lambda \oplus V_{\lambda+2} \oplus \cdots$$

where H multiplies by the weight, E raises the weight, and F lowers the weight. We do not know that all the weights of V lie in this sequence—we will see that soon!—but because V is finite-dimensional, we know this sequence cannot go on forever, so there must be a largest weight:

$$V = \cdots \oplus V_{\lambda-2} \oplus V_\lambda \oplus V_{\lambda+2} \oplus \cdots \oplus V_{\lambda_{\max}}.$$

The weight $\lambda_{\max}$ is called the **highest weight**, and a nonzero vector $v \in V_{\lambda_{\max}}$ is called a **highest weight vector**. A highest weight vector has the property that $Ev = 0$.

Similarly, there must be a lowest weight:

$$V = V_{\lambda_{\min}} \oplus \cdots \oplus V_{\lambda-2} \oplus V_\lambda \oplus V_{\lambda+2} \oplus \cdots \oplus V_{\lambda_{\max}},$$

For any $v \in V_{\lambda_{\min}}$, we must have $Fv = 0$.

Now, if we pick a highest weight vector $v \in V_{\lambda_{\max}}$ and keep lowering it with F, we will eventually get a vector in $V_{\lambda_{\min}}$. Let us suppose this happens in m steps. That is, m is the natural number such that $F^m v \neq 0$, but $F^{m+1}v = 0$.

Proposition 13. *The vectors $\{v, Fv, \ldots, F^m v\}$ form a basis of V.*

Proof. These vectors are linearly independent because they are eigenvectors (weight vectors) with distinct eigenvalues (weights). To show they span V, let $W = \text{span}(v, Fv, \ldots, F^m v)$. The nonzero subspace W is preserved by the action of E, F , and H. Hence, W is invariant and we conclude $W = V$, because V is irreducible. $\square$

In this basis, we know exactly how H and F act:

$$HF^k v = (\lambda_{\max} - 2k)F^k v, \quad FF^k v = F^{k+1}v,$$

but it is less clear how E acts. Let us derive a formula for the action of E, inductively. First of all, we know $Ev = 0$. For EFv, we compute:

$$EFv = FEv + [E, F]v = 0 + Hv = \lambda_{\max} v.$$

And for $EF^2 v$, we have:

$$EF^2 v = FEFv + [E, F]Fv = \lambda_{\max} Fv + HFv = (2\lambda_{\max} - 2)Fv.$$

Continuing in this way, we can discover the pattern:

$$EF^{k+1}v = (\lambda_{\max} + (\lambda_{\max} - 2) + \cdots + (\lambda_{\max} - 2k))F^k v,$$

which simplifies to $EF^{k+1}v = (k + 1)(\lambda_{\max} - k)F^k v$.

We learn something magical from this formula when we set $k = m$:

$$EF^{m+1}v = 0 = (m+1)(\lambda_{\max} - m)F^m v.$$

It vanishes because $F^m v$ is in the lowest weight space, so $F^{m+1}v = 0$. But on the right hand side, the $m + 1$ is nonzero, and the vector $F^m v$ is nonzero. So the only way this can vanish is if

$$\lambda_{\max} = m.$$

Look! The highest weight $\lambda_{\max}$ is a natural number! Specifically, it is the number of times we need to apply F to go from the highest weight vector v to the lowest. We have nearly proved:

Theorem 14. *For each natural number m (including zero), there is a unique finite-dimensional irreducible representation $(V^{(m)}, \rho_{(m)})$ of $\mathfrak{sl}(2, \mathbb{C})$ with highest weight m. All finite-dimensional irreps of $\mathfrak{sl}(2, \mathbb{C})$ have this form.*

To recap, if V is an irrep with highest weight m, V decomposes into the weight spaces $V = V_{-m} \oplus V_{-m+2} \oplus \cdots \oplus V_{m-2} \oplus V_m$. Each weight space is one-dimensional, spanned by one of the basis vectors in $\{v, Fv, \ldots, F^m v\}$. We summarize all of these facts in the **weight diagram** of V:

$$\begin{array}{ccccc} \bullet & \bullet & \cdots & \bullet & \bullet \\ -m & -m+2 & & m-2 & m \end{array}$$

Each dot represents a weight space. In more general weight diagrams such as those in the next section, the dots can have multiplicities. Here, they all have multiplicity one, telling us that each weight space is one-dimensional.

4 The representations of $\mathfrak{sl}(3, \mathbb{C})$

The representation theory of $\mathfrak{sl}(3, \mathbb{C})$ begins the same way: we choose a Cartan subalgebra $\mathfrak{h} \subseteq \mathfrak{sl}(3, \mathbb{C})$. As before, we take $\mathfrak{h}$ to consist of traceless diagonal matrices. Thus $\mathfrak{h} = \operatorname{span}(H_1, H_2)$ is two-dimensional, and we pick:

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad H_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

As before, the Cartan subalgebra $\mathfrak{h}$ is maximal abelian, and $\operatorname{ad}(H)$ is diagonalizable for any $H \in \mathfrak{h}$, thanks to the formula $[H, E_{ij}] = (a_i - a_j)E_{ij}$, where

$$H = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix},$$

and E_{ij} is the matrix with 1 in the ijth entry, and zeroes elsewhere.

To analyze representations, we need a version of Theorem 11 for $\mathfrak{sl}(n, \mathbb{C})$:

Theorem 15. *For any complex finite-dimensional representation (V, ρ) of $\mathfrak{sl}(n, \mathbb{C})$, and any choice of Cartan subalgebra $\mathfrak{h} \subseteq \mathfrak{sl}(n, \mathbb{C})$, $\rho(H)$ is diagonalizable for all $H \in \mathfrak{h}$.*

In analogy with $\mathfrak{sl}(2, \mathbb{C})$, we use this to write any representation as a direct sum over weights:

$$V = \bigoplus_{\lambda} V_{\lambda}.$$

For $\mathfrak{sl}(2, \mathbb{C})$, there was only one H, and a weight was simply an eigenvalue of H. But now $\mathfrak{h}$ is two-dimensional, and there are many $H \in \mathfrak{h}$. In this instance, what is a weight?

Definition 16. Given a representation (V, ρ), a nonzero vector $v \in V$ is a **weight vector** if $\rho(H)v = \lambda(H)v$ for all $H \in \mathfrak{h}$ and some linear map $\lambda\colon \mathfrak{h} \to \mathbb{C}$. Here, λ is called a **weight** of the representation V.

Weights are a generalization of eigenvalues, and weight vectors are a generalization of eigenvectors that allow us to diagonalize all the $H \in \mathfrak{h}$ at once. And we really can diagonalize all the $H \in \mathfrak{h}$ simultaneously, because they commute!

To get a feel for the weights of a representation of $\mathfrak{sl}(3, \mathbb{C})$, let us consider an example:

Example 17 (The weights of the adjoint representation.)**.** We already know the weight vectors—they are the elementary matrices E_{ij}, at least when $i \neq j$. This is because of the formula:

$$\mathrm{ad}(H)E_{ij} = (a_i - a_j)E_{ij}.$$

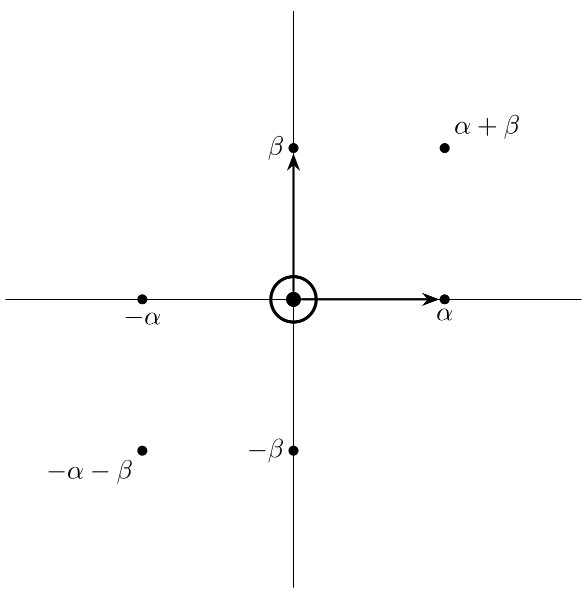
(For $i = j$, E_{ii} has trace 1 and is not in $\mathfrak{sl}(3, \mathbb{C})$.)

So let us define the weight $\alpha_{ij}\colon \mathfrak{h} \to \mathbb{C}$ by the formula $\alpha_{ij}(H) = a_i - a_j$, where a_i and a_j are the ith and jth entries of the diagonal matrix H. Then we have $\mathrm{ad}(H)E_{ij} = \alpha_{ij}(H)E_{ij}$.

To get a feel for the weights of the adjoint representation, note that we have the relations:

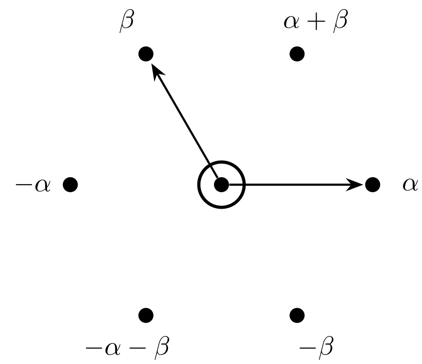
$$\alpha_{ij} = -\alpha_{ji}, \quad \alpha_{ii} = 0, \quad \alpha_{ij} + \alpha_{jk} = \alpha_{ik}.$$

These imply that all the weights of the adjoint can be expressed as a linear combination of two such weights. Let us pick $\alpha = \alpha_{12}$ and $\beta = \alpha_{23}$ as a basis. Then the other nonzero weights are $\alpha_{13} = \alpha + \beta$, $\alpha_{21} = -\alpha$, $\alpha_{32} = -\beta$, and $\alpha_{31} = -\alpha - \beta$. To really get a picture, we plot these weights:



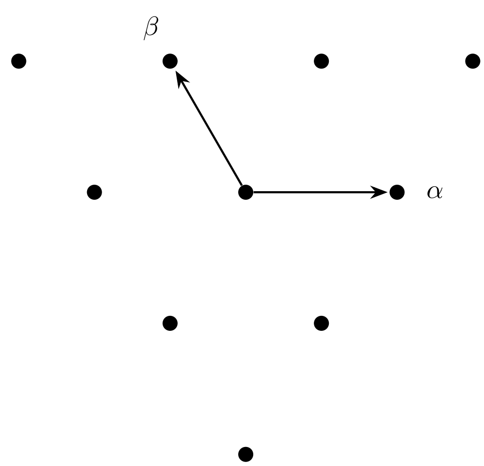
In the plot, we draw one dot for each weight space in the adjoint representation, except in the middle: the weight space of weight zero is two-dimensional—it is $\mathfrak{h}$!—and we depict this by adding the extra circle around the zero weight. We have six dots around the outside and two in the middle. The total is eight, as it must be: $\dim(\mathfrak{sl}(3, \mathbb{C})) = 8$.

This picture is usually drawn with more symmetry, as a regular hexagon with two dots in the middle:



This is the weight diagram of the adjoint representation.

Remarkably, all the irreps of $\mathfrak{sl}(3, \mathbb{C})$ have similar weight diagrams. For instance, here is the weight diagram of a 10-dimensional irrep:



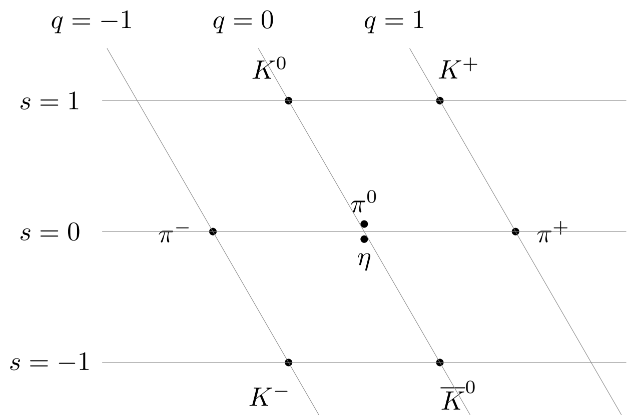
We have indicated α and β on this diagram to show how it compares to the adjoint.

5 The eightfold way

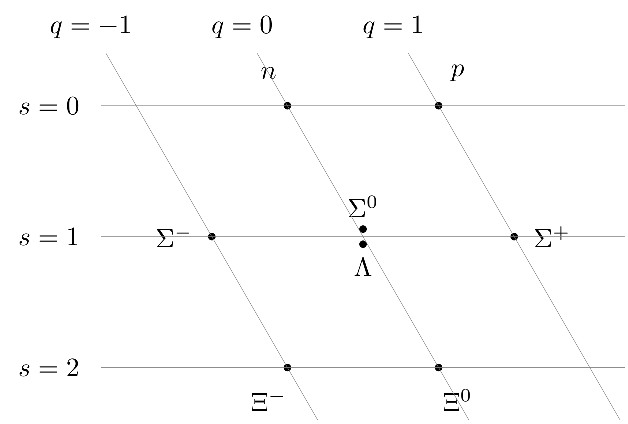
In the 1940s and 1950s, physicists built particle accelerators and began to collide protons together at high energies. Protons are strongly interacting particles, in the sense that they feel the strong nuclear force that binds them together in the atomic nucleus. Colliding protons produced *many* new, hitherto unknown particles, likewise strongly interacting. In physics, strongly interacting particles are called **hadrons**—the Greek root *hadros* means strong.

No one expected such a zoo of new particles, and so a search was on for some kind of order, some system to classify the hadrons. Many properties of the particles were measured. Each particle X had an electric charge $q(X)$, which is an integer in suitable units. But several other kinds of "charge" were discovered. It turned out each particle had a property, called **strangeness** $s(X)$, which was also an integer.

When you plot the charge and the strangeness of hadrons on a plane, certain patterns emerge. For instance, here is the **spin-0 meson octet** (mesons are a type of hadron):



And here is the **spin-1/2 baryon octet**. (Baryons are another type of hadron, which includes the proton and neutron. In fact, the proton and neutron are the particles n and p at the top):

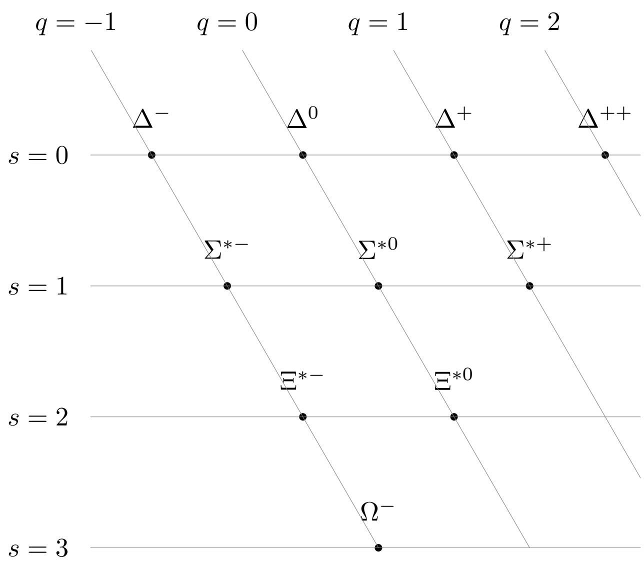


As you can clearly see, both of these are pictures of the adjoint representation of $\mathfrak{sl}(3, \mathbb{C})$! This led the American physicist Murray Gell-Mann and the Israeli physicist Yuval Ne'emann to propose independently **the eightfold way** hypothesis. The name comes from appearance of eight particles in the octets, and was Gell-Mann's allusion to the eightfold path to enlightenment in Buddhism. Here is the hypothesis:

Hypothesis (The eightfold way)**.** Hadrons are classified by representations of $\mathfrak{sl}(3, \mathbb{C})$.

In its original form, the eightfold way used the Lie group SU(3) in place of the Lie algebra $\mathfrak{sl}(3, \mathbb{C})$. It is a marvelous result of Lie theory that these objects have equivalent representation theory, so we have substituted $\mathfrak{sl}(3, \mathbb{C})$ to ease exposition.

The vindication of the eightfold way came with the prediction of new particles. This followed not from the octets above, but from the **spin-3/2 baryon decuplet**:



The particle at the bottom, the Ω^-, was previously unknown. Gell-Mann predicted it in 1962 on the basis of the eightfold way, and it was discovered in 1964.

6 Further reading

The best reference for the Lie theory we have discussed is the book by Fulton and Harris [1], to which our treatment of $\mathfrak{sl}(2,\mathbb{C})$ owes everything. Of course, Lie algebras are closely related to Lie groups, and a good first introduction can be found in the book of Hall [2]. For the eightfold way, a wonderful treatment can be found in Sternberg [3], who frames the question in terms of the representations of the Lie group SU(3), rather than the Lie algebra $\mathfrak{sl}(3,\mathbb{C})$ we used here.

References

[1] William Fulton and Joe Harris, *Representation theory: a first course*, Springer–Verlag, New York, 1991.

[2] Brian Hall, *Lie groups, Lie algebras, and representations: an elementary introduction*, 2nd ed., Springer, Cham, 2015.

[3] Shlomo Sternberg, *Group theory and physics*, CUP, Cambridge, 1994.