SYMMETRIES, FIELDS, AND PARTICLES

IAN LIM
LAST UPDATED NOVEMBER 17, 2018

These notes were taken for the *Symmetries, Fields, and Particles* course taught by Nick Dorey at the University of Cambridge as part of the Mathematical Tripos Part III in Michaelmas Term 2018. I live-TEXed them using Overleaf, and as such there may be typos; please send questions, comments, complaints, and corrections to itel2@cam.ac.uk.
Many thanks to Arun Debray for the LATEX template for these lecture notes: as of the time of writing, you can find him at https://web.ma.utexas.edu/users/a.debray/.

Contents

Lecture 1.

Symmetries, Fields, and Start-icles: Thursday, October 4, 2018

Today we'll outline the content of this course and motivate it with a few examples. To begin with, symmetry as a principle has led physicists all the way to our current model of physics. This course's content will be almost exclusively mathematical, yet more pragmatic about introducing the necessary tools to apply symmetries to the physical systems we're interested in.

Resources

- Notes (online)
  - Nick Manton's notes (concise, more on geometry of Lie groups)
  - Hugh Osborn's notes (comprehensive, don't cover Cartan classification)

  - Jan Gutowski's notes (classification of Lie algebras). There is actually a second set of notes on an earlier version of the course which can be found here, but I believe the notes referred to in lecture are the first set.
- Books: "Symmetries, Lie Algebras and Representations", Fuchs & Schweigert Ch. 1-7.

Prof. Dorey has also provided his own handwritten notes, which I will be typing up and supplementing with lecture material here.

Introduction

Definition 1.1. We define a *symmetry* as a transformation of dynamical variables that leaves the form of physical laws invariant.

Example 1.2. A rotation is a transformation, e.g. on $\mathbf{x} \in \mathbb{R}^3$ such that $\mathbf{x}' = M \cdot \mathbf{x} \in \mathbb{R}^3$. There are *orthogonal* matrices which satisfy $MM^T = 1_3$ and also *special* matrices which satisfy $\det M = 1$.

It's also useful for us to define the notion of a group (likely familiar from an intro course on abstract algebra or mathematical methods).

Definition 1.3. A *group* G is a set equipped with a multiplication law (binary operation) obeying

- Closure ($\forall g_1, g_2 \in G, g_1 g_2 \in G$)
- Identity ($\exists e \in G \text{s.t.} \forall g \in G, eg = ge = g$)
- Existence of inverses ($\forall g \in G, \exists g^{-1} \in G$ s.t. $g^{-1}g = gg^{-1} = e$)
- Associativity ($\forall g_1, g_2, g_3 \in G, (g_1 g_2) g_3) = g_1(g_2 g_3)$).

Exercise 1.4. For rotations $G = SO(3)$, the group of 3-dimensional special orthogonal matrices, check that the group axioms apply (SO(3) forms a group).[1]

We also remark that the set may be finite or infinite[2].

Definition 1.5. A group G is called *abelian* if the multiplication law is commutative ($\forall g_1, g_2 \in G, g_1 g_2 = g_2 g_1$). Otherwise, it is called non-abelian.

We notice that a rotation in $\mathbb{R}^3$ depends continuously on 3 parameters: $\hat{n} \in S^2, \theta \in [0, \pi]$ (with $\hat{n}$ the axis of rotation, θ the angle of rotation). This leads us to introduce the idea of a Lie group.

Definition 1.6. A *Lie group* G is a group which is also a smooth manifold. It's key that the group and manifold structures must be compatible, and so G is (almost) completely determined by the behavior "near" e, i.e. by infinitesimal transformations in a small neighborhood of the identity element e. These correspond to the *tangent vectors* to G at e.

The tangent vectors are local objects which span the tangent space to the manifold at some given point. It turns out that $\forall v_1, v_2 \in T_e(G)$ the tangent space of G, we can define a binary operation $[,] : T_e(G) \times T_e(G) \to T_e(G)$ such that $[,]$ is bilinear, antisymmetric, and obeys the Jacobi identity.

Definition 1.7. The tangent space at the identity equipped with the Lie bracket defines a *Lie algebra* $\mathcal{L}(G)$.

It's a remarkable fact that *all* finite-dimensional semi-simple Lie algebras (over $\mathbb{C}$) can be classified into four infinite families A_n, B_n, C_n, D_n with $n \in \mathbb{N}$, plus five *exceptional cases* E_6, E_7, E_8, G_2, F_4.[3] We call this the *Cartan classification.*

Symmetries in physics In classical physics, (continuous) symmetries give rise to conserved quantities. This is the conclusion of Noether's theorem.

Example 1.8. Rotations in $\mathbb{R}^3$ correspond to conservation of angular momentum, $\mathbf{L} = (L_1, L_2, L_3)$.

In quantum mechanics, we have

- states: vectors in Hilbert space $|\psi\rangle \in \mathcal{H}$

[1] We'll prove this more generally for $SO(n)$ in a few lectures. The answer is in the footnote to Exercise 3.4.

[2] For example, cyclic groups $\mathbb{Z}_n$ (i.e. addition in modular arithmetic) vs. most matrix groups like GL_n.

[3] The exceptional groups have not yet come up in physical phenomena, but they seem to have a mysterious connection to the absence of anomalies in string theory.

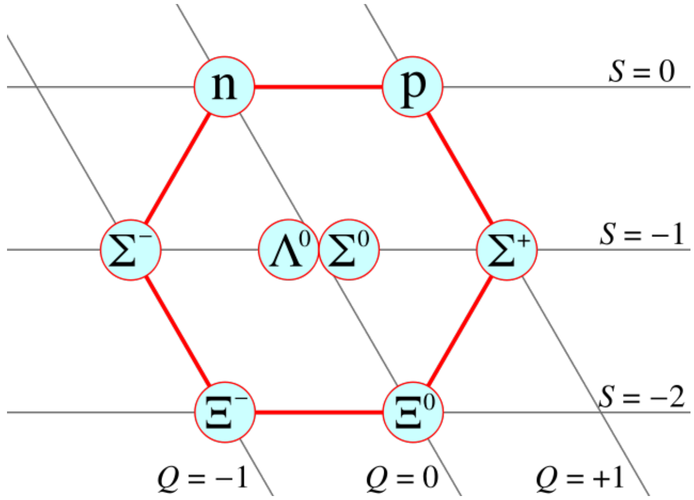

FIGURE 1. The baryon octet. Particles are arranged by their charge along the diagonals and by their strangeness on the horizontal lines.

- observables: linear operators $\hat{O} : \mathcal{H} \to \mathcal{H}$ with (generally) non-commutative multiplication.

We recall from previous courses in QM that operators which commute with the Hamiltonian (e.g. $[\hat{H}, \hat{L}_i] = 0, i = 1, 2, 3$) give rise to "quantum conserved quantities."

In fact, we recall that the angular momentum operators are associated to a Lie bracket: $[\hat{L}_i, \hat{L}_j] = i\epsilon_{ijk}\hat{L}_k$. But this is exactly the $\mathcal{L}(SO(3))$ Lie algebra.

Our angular momentum operators often act on finite-dimensional vector spaces, e.g. *electron spin*.

$$|\uparrow\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |\downarrow\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

This corresponds to a two-dimensional *representation* of $\mathcal{L}(SO(3)$, i.e. a set of 2×2 matrices $\Sigma_i, i = 1, 2, 3$ satisfying the same Lie algebra,

$$[\Sigma_i, \Sigma_J] = i\varepsilon_{ijk}\Sigma_k,$$

which is provided by setting $\Sigma_i = \frac{1}{2}\sigma_i$, our old friends the Pauli matrices.

More generally, we should think of a representation as a map e from a Lie group to some space of transformations on a vector space which preserves the Lie bracket, $e([v_1, v_2]) = [e(v_1), e(v_2)]$.

Now suppose we have a rotational symmetry in a quantum system,

$$[\hat{H}, \hat{L}_i] = 0, i = 1, 2, 3.$$

Then the spin states obey $\hat{H}|\uparrow\rangle = E|\uparrow\rangle, \hat{H}|\downarrow\rangle = E'|\downarrow\rangle$, with $E = E'$. More generally, degeneracies in the energy spectrum of quantum systems correspond to irreducible representations of symmetries.

Example 1.9. We have an approximate $SU(3)$ symmetry for the strong force, with

$$G = SU(3) \equiv \{3 \times 3 \text{ complex matrices } M \text{ with } MM^\dagger = I_3 \text{ and } \det M = 1.\}$$

The spectrum of mesons and baryons are thus defined by the representation of the Lie algebra $\mathcal{L}(SU(3))$. See also the "eightfold way," due to Murray Gell-Mann, who showed that plotting the various mesons and baryons with respect to certain quantum numbers (isospin and hypercharge) gives rise to a very nice picture corresponding to the 8-dimensional representation of the Lie algebra $\mathcal{L}(SU(3))$.

Lecture 2.

Symmetry Described Simp-Lie: Saturday, October 6, 2018

So far, we have discussed global symmetries.

- Spacetime symmetries:

    - Rotation, $SO(3)$.
    - Lorentz transformations, $SO(3,1)$. (Rotations in $\mathbb{R}^3$ plus boosts.)
    - The Poincaré group (not a simple Lie group, so does not fit Cartan classifications)
    - Supersymmetry? (i.e. a symmetry between fermions and bosons, described by "super" Lie algebra)
- Internal symmetries:
    - Electric charge
    - Flavor, $SU(3)$ in hadrons
    - Baryon number

But we also have gauge symmetry.

Definition 2.1. A *gauge symmetry* is a redundancy in our mathematical description of physics. For instance, the phase of the wavefunction in quantum mechanics has no physical meaning:

$$\psi \to e^{i\delta}\psi \tag{2.2}$$

leaves all the physics unchanged ($\delta \in \mathbb{R}$).[4]

Example 2.3. Another gauge symmetry familiar to us is the gauge transformation in electrodynamics,

$$\mathbf{A}(\mathbf{x}) \to \mathbf{A}(\mathbf{x}) + \boldsymbol{\nabla}\chi(\mathbf{x}).$$

By adding the gradient of some scalar function χ of $\mathbf{x}$, this leaves $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$ unchanged (since $\boldsymbol{\nabla} \times \boldsymbol{\nabla} F = 0$) and so the fields corresponding to the vector potential produce the same physics. Gauge invariance turns out to be key to our ability to quantize the spin-1 field corresponding to the photon.

Example 2.4. Another example (maybe less familiar in the exact details) is the Standard Model of particle physics.[5] The Standard Model is a non-abelian gauge theory based on the Lie group

$$G_{SM} = SU(3) \times SU(2) \times U(1).$$

We started to describe Lie groups last time. Let us repeat the definition here: a Lie group G is a group which is also a (smooth) manifold. Informally, a manifold is a space which locally looks like $\mathbb{R}^n$– for every point on the manifold, there is a smooth map from an open set of $\mathbb{R}^n$ to the manifold (that patch "looks flat"), and these maps are compatible. For cute wordplay reasons, the collection of such maps is known as an atlas.

Sometimes it is useful to consider a manifold as embedded in an ambient space, e.g. S^2 embedded in $\mathbb{R}^3$: $\mathbf{x}(x,y,z) \in \mathbb{R}^3$ such that $x^2 + y^2 + z^2 = r^2, r > 0$.

More generally, we can take the set of all $\mathbf{x} = (x_1, x_2, \ldots, x_{n+m}) \in \mathbb{R}^{n+m}$ such that for a continuous, differentiable set of functions $F^\alpha(\mathbf{x}) : \mathbb{R}^{n+m} \to \mathbb{R}, \alpha = 1, \ldots, m$, a space M is defined by all such $\mathbf{x}$ satisfying $F^\alpha(\mathbf{x}) = 0, \alpha =, 1 \ldots, m$. That is,

$$M = \{\mathbf{x} \in \mathbb{R}^{n+m} : F^\alpha(\mathbf{x}) = 0, \alpha =, 1 \ldots, m\} \tag{2.5}$$

Then the following theorem holds.

Theorem 2.6. *M is a smooth manifold of dimension n if the Jacobian matrix J has rank m, with the Jacobian defined*

$$J_i^\alpha = \frac{\partial F^\alpha}{\partial x_i}.$$

In words, all this says is that M is a manifold if F^α imposes a nice independent set of m constraints on our $n+m$ variables, leaving us with a manifold of dimension n.

Example 2.7. For the sphere S^2, we have $m = 1, n = 2$ and we have the constraint $F^1(\mathbf{x}) = x^2 + y^2 + z^2 - r^2$ for some r. Then the Jacobian is simply

$$J = \left(\frac{\partial F^1}{\partial x}, \frac{\partial F^1}{\partial y}, \frac{\partial F^1}{\partial z}\right) = 2(x,y,z),$$

and this matrix indeed has rank 1 unless $x = y = z = 0$. Therefore we can represent S^2 as a manifold of dimension 2 embedded in $\mathbb{R}^3$.

[4] However, differences in phase can have significant effects– see for instance the Aharanov-Bohm effect.

[5] We'll unpack the Standard Model more in next term's Standard Model class.

Group operations (multiplication, inverses) define smooth maps on the manifold. The *dimension* of G, denoted $\dim(G)$, is the dimension of the group manifold $M(G)$. We may introduce coordinates $\{\theta^i\}, i = 1, \ldots, D = \dim(G)$ in some local coordinate patch P containing the identity $e \in G$. Then the group elements depend continuously on $\{\theta^i\}$, such that $g = g(\theta) \in G$ (the manifold structure is compatible with group elements). Set $g(0) = e$.

Thus if we choose two points θ, θ' on the manifold M, group multiplication,

$$g(\theta)g(\theta') = g(\phi) \in G,$$

corresponds to (induces) a smooth map $\phi : G \times G \to G$ which can be expressed in coordinates

$$\phi^i = \phi^i(\theta, \theta'), i = 1, \ldots, D$$

such that $g(0) = e \implies$

$$\phi^i(\theta, 0) = \theta^i, \phi^i(0, \theta') = \theta'^i.$$

We ought to be a little careful that our group multiplication doesn't take us out of the coordinate patch we've defined our coordinates on, but in practice this shouldn't cause us too many problems.

Similarly, group inversion defines a smooth map, $G \to G$. This map can be written as follows:

$$\forall g(\theta) \in G, \exists g^{-1}(\theta) = g(\tilde{\theta}) \in G$$

such that

$$g(\theta)g(\tilde{\theta}) = g(\tilde{\theta})g(\theta) = e.$$

In coordinates, the map

$$\tilde{\theta}^i = \tilde{\theta}^i(\theta), i = 1, \ldots, D$$

is continuous and differentiable.

Example 2.8. Take the Lie group $G = (\mathbb{R}^D, +)$ (Euclidean D-dimensional space with addition as the group operation). Then the map defined by group multiplication is simply

$$\mathbf{x}'' = \mathbf{x} + \mathbf{x}' \forall \mathbf{x}, \mathbf{x}' \in \mathbb{R}^D$$

and similarly the map defined by group inversion is

$$\mathbf{x}^{-1} = -\mathbf{x} \forall \mathbf{x} \in \mathbb{R}^D.$$

This is a bit boring since the group multiplication law is commutative, so we'll next look at some important non-abelian groups– namely, the matrix groups.

Matrix groups Let $\mathrm{Mat}_n(F)$ denote the set of $n \times n$ matrices with entries in a field $F = \mathbb{R}$ or $\mathbb{C}$. These satisfy some of the group axioms– matrix multiplication is closed and associative, and there is an obvious unit element, $e = I_n \in \mathrm{Mat}_n(F)$ (with I_n the $n \times n$ unit matrix). However, $\mathrm{Mat}_n(F)$ is not a (multiplicative) group because not all matrices are invertible (e.g. with $\det M = 0$). (Since it is not a group, it is also not a Lie group, though it does have a manifold structure, that of $\mathbb{R}^{n^2}$.) Thus, we define the *general linear groups.*

Definition 2.9. The general linear group $GL(n, F)$ is the set of matrices defined by

$$GL(n, F) \equiv \{M \in \mathrm{Mat}_n(F) : \det M \neq 0\}. \tag{2.10}$$

Definition 2.11. We also define the *special linear groups* $SL(n, F)$ as follows:

$$SL(n, F) \equiv \{M \in GL(n, F) : \det M = 1.\} \tag{2.12}$$

Here, closure follows from the fact that determinants multiply nicely, $\forall M_1, M_2 \in GL(n, F), \det(M_1 M_2) = \det(M_1)\det(M_2) = 1$ for $SL(n, F)$ (is nonzero for $GL(n, F)$), and existence of inverses follows from the defining condition that $\det M \neq 0$.

It's less obvious that $GL(n, F)$ and $SL(n, F)$ are also Lie groups. In fact, our theorem (Thm. 2.6) applies here: the condition that $\det M = \pm 1$ corresponds to a nice $F(\mathbf{x}) = \det M - 1, \mathbf{x} \in \mathbb{R}^{n^2}$, which is sufficently nice as to define a manifold. The same is true for $SL(n, F)$, so these are indeed Lie groups. Note the dimensions of these sets are as follows.

$$\begin{array}{cc} \dim(GL(n, \mathbb{R})) = n^2 & \dim(GL(n, \mathbb{C})) = 2n^2 \\ \dim(GL(n, \mathbb{R})) = n^2 - 1 & \dim(SL(n, \mathbb{C})) = 2n^2 - 2 \end{array}$$

And now, a bit of extra detail on the dimensions and manifold properties of these Lie groups. In $\mathrm{Mat}_n(F)$, we have our free choice of any numbers we like in F for the n^2 elements of our matrix. It turns out that imposing $\det M \neq 0$ is not too strong a constraint– it eliminates a set of zero measure from the space of possible $n \times n$ matrices, so we have our choice of n^2 real numbers in $GL(n, \mathbb{R})$ and n^2 complex numbers (so $2n^2$ real numbers) in $GL(n, \mathbb{C})$. Requiring that $\det M \neq 0$ means we can equivalently view $GL(n, \mathbb{R})$ as the preimage of an open set in $\mathbb{R}$ (since $\det M : \mathbb{R}^{n^2} \to \mathbb{R}$) under a continuous (and smooth!) map, which is therefore an open set in $\mathbb{R}^{n^2}$. It turns out that any open set in $\mathbb{R}^{n^2}$ is itself a manifold (really, any open subset of a manifold), so $GL(n, \mathbb{R})$ is indeed a manifold.

Note that the situation is easier in $SL(n, F)$, since our theorem then applies with $F = \det M - 1$. The corresponding Jacobian has rank 1 unless all the matrix elements vanish identically, so $SL(n, F)$ is a manifold Imposing the restriction that $\det M = 1$ is now a stronger algebraic condition– it reduces our choice of values by 1, since if we have picked $n^2 - 1$ values of the matrix, the last value is completely determined by $\det M = 1$. Thus the dimension of $SL(n, \mathbb{R})$ is $n^2 - 1$. Since we get to pick $n^2 - 1$ complex numbers in $SL(n, \mathbb{C})$ (equivalently there are two real constraints, one on the real components and one on the imaginary ones), that amounts to $2(n^2 - 1) = 2n^2 - 2$ real numbers. Hence, dimension $2n^2 - 2$.

Definition 2.13. A *subgroup* H of a group G is a subset ($H \subseteq G$) which is also a group. We write it as $H \leq G$. If H is also a smooth submanifold of G, we call H a *Lie subgroup* of G.

Lecture 3.

Here Comes the SO(n): Tuesday, October 9, 2018

Having introduced the matrix groups, we'll next discuss some important subgroups of $GL(n, \mathbb{R})$. First, the *orthogonal groups.*

Definition 3.1. Orthogonal groups $O(n)$ are the matrix groups which preserve the Euclidean inner product,

$$O(n) = \{M \in GL(n, \mathbb{R}) : M^T M = I_N\}. \tag{3.2}$$

Their elements correspond to orthogonal transformations, so that for $\mathbf{v} \in \mathbb{R}^n$, an orthogonal matrix M acts on $\mathbf{v}$ by matrix multiplication,

$$\mathbf{v}' = M \cdot \mathbf{v}$$

and so in particular

$$|\mathbf{v}'|^2 = \mathbf{v}'^T \cdot \mathbf{v}' = \mathbf{v}^T \cdot M^T M \cdot \mathbf{v} = \mathbf{v}^T \cdot \mathbf{v} = |\mathbf{v}|^2.$$

It also follows that $\forall M \in O(n), \det(M^T M) = \det(M)^2 = \det(I_n) = 1 \implies \det M = \pm 1$.

$\det M$ is a smooth function of the coordinates, but our constraint equation means that $\det M$ can only take on one of two discrete values. The orthogonal group $O(n)$ has therefore two connected components corresponding to $\det M = +1$ and $\det M = -1$. The connected component containing the origin ($\det M = +1$) is the special orthogonal group $SO(n)$.

Definition 3.3. The *special orthogonal groups* $SO(n)$ are the subset of orthogonal groups which also preserve orientation (i.e. no reflections):

$$SO(n) \equiv \{M \in O(n) : \det M = +1\}.$$

That is, elements of $SO(n)$ preserve the sign of the volume element in $\mathbb{R}^n$,

$$\Omega = \epsilon^{i_1 i_2 \ldots i_n} v_1^{i_1} v_2^{i_2} \ldots v_n^{i_n}.$$

In contrast, $O(n)$ matrices may include reflections as well as rotations when $\det M = -1$.

Exercise 3.4. Check the group axioms for $SO(n)$.[6] Show that $\dim(O(n)) = \dim(SO(n)) = \frac{1}{2}n(n-1)$.[7]

Orthogonal matrices have some nice properties. Let $M \in O(n)$ be an orthogonal matrix and suppose that $\mathbf{v}_\lambda$ is an eigenvector of M with eigenvalue λ. Then the following is true:

(a) If λ is an eigenvalue, then λ^* is also an eigenvalue (eigenvalues of M come in complex conjugate pairs).

(b) $|\lambda|^2 = 1$.

The proof is as follows:

(a) $M \cdot \mathbf{v}_\lambda = \lambda \mathbf{v}_\lambda \implies M \cdot \mathbf{v}_\lambda^* = \lambda^* \mathbf{v}_\lambda^*$ (since M is a real matrix).[8]

(b) For any complex vector $\mathbf{v}$, we have

$$(M \cdot \mathbf{v}^*)^T \cdot M \cdot \mathbf{v} = \mathbf{v}^\dagger \cdot M^T M \cdot \mathbf{v} = \mathbf{v}^\dagger \cdot \mathbf{v}.$$

Now if $\mathbf{v} = \mathbf{v}_\lambda$, then

$$(M \cdot \mathbf{v}_\lambda^*)^T \cdot M \cdot \mathbf{v}_\lambda = (\lambda^* \mathbf{v}_\lambda^*)^T \cdot (\lambda \mathbf{v}_\lambda) = |\lambda|^2 \mathbf{v}_\lambda^\dagger \cdot \mathbf{v}_\lambda.$$

By comparison to the first expression, we see that $|\lambda|^2 = 1$. ☒

Example 3.5. For the group $G = SO(2)$, $M \in SO(2) \implies M$ has eigenvalues

$$\lambda = e^{i\theta}, e^{-i\theta}$$

for some $\theta \in \mathbb{R}, \theta \sim \theta + 2\pi$ (identified up to a phase of 2π). A group element may be written explicitly as

$$M = M(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

which is uniquely specified by a rotation angle θ. Therefore the group manifold of $SO(2)$ is $M(SO(2)) \cong S^1$, the circle, and we see that $SO(2)$ is an abelian group..

It's not too hard to check using the trig addition formulas that the matrices M written this way really do form a representation of $SO(2)$, since $M(\theta_1)M(\theta_2) = M(\theta_1 + \theta_1)$.

Example 3.6. For the group $G = SO(3)$, we have instead $M \in SO(3) \implies M$ has eigenvalues

$$\lambda = e^{i\theta}, e^{-i\theta}, 1$$

for $\theta \in \mathbb{R}, \theta \sim \theta + 2\pi$, using our two properties again of paired eigenvalues and modulus 1. The normalized eigenvector for $\lambda = 1$, $\hat{\mathbf{n}} \in \mathbb{R}^3$, specifies the axis of rotation ($M \cdot \hat{\mathbf{n}} = \hat{\mathbf{n}}$ and $\hat{\mathbf{n}} \cdot \hat{\mathbf{n}} = 0$).

A general group element of $SO(3)$ can be written explicitly as

$$M(\hat{\mathbf{n}}, \theta)_{ij} = \cos\theta \delta_{ij} + (1 - \cos\theta) n_i n_j - \sin\theta \epsilon_{ijk} n_k. \tag{3.7}$$

Let us remark that our group is invariant under the identification $\theta \to 2\pi - \theta, \hat{\mathbf{n}} \to -\hat{\mathbf{n}}$. It's also true that we should identify all M with $\theta = 0$ since $M(\hat{\mathbf{n}}, 0) = I_3 \forall \hat{\mathbf{n}}$.

We also observe that we can consider the vector

$$\mathbf{w} \equiv \theta \hat{\mathbf{n}}$$

[6] As usual, we need to check closure and inverses. The identity matrix I satisfies $I^T I = I$ and $\det I = 1$, and associativity follows from standard matrix multiplication. Inverses: if $M \in SO(n)$, then M^{-1} is defined by $MM^{-1} = I$. But $\det(MM^{-1}) = \det(M)\det(M^{-1}) = (1)\det(M^{-1}) = \det I = 1$, so $\det(M^{-1}) = 1$. We also check that the inverse of an orthogonal matrix is also orthogonal: $MM^{-1} = I$, so $(M^{-1})^T(M^T) = (M^{-1})^T M^{-1} = I^T = I$. Closure: $\forall M, N \in SO(n), \det(MN) = \det(M)\det(N) = (1)(1) = 1$ and $(MN)^T(MN) = N^T M^T M N = I$, so $MN \in SO(n)$. ☒

[7] This can be seen by writing a matrix $M \in SO(n)$ as a row of n column vectors $(\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n})$. Then the condition that $M^T M = 1$ is equivalent to $\begin{pmatrix} \mathbf{x_1}\cdot\mathbf{x_1} & \mathbf{x_1}\cdot\mathbf{x_2} & \ldots & \mathbf{x_1}\cdot\mathbf{x_n} \\ \mathbf{x_2}\cdot\mathbf{x_1} & \mathbf{x_2}\cdot\mathbf{x_2} & \ldots & \mathbf{x_2}\cdot\mathbf{x_n} \\ \vdots & & & \\ \mathbf{x_n}\cdot\mathbf{x_1} & \ldots & \ldots & \mathbf{x_n}\cdot\mathbf{x_n} \end{pmatrix} = I_n$. We see that by the symmetry of the explicit form of $M^T M$, we get $1 + 2 + 3 + \ldots + n = n(n+1)/2$ independent constraints on the n^2 entries of M. Applying our theorem, we find that the resulting manifold has dimension $n^2 - n(n+1)/2 = n(n-1)/2$.

[8] This is generally true of real matrices with complex eigenvalues– it's not specific to orthogonal matrices.

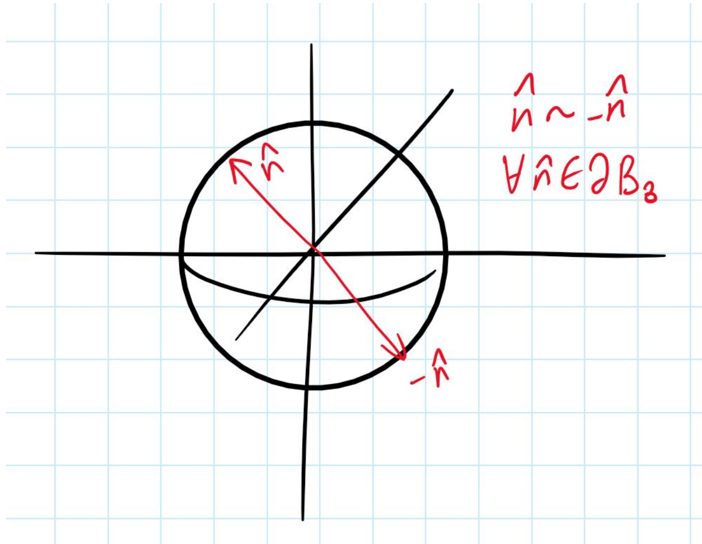

FIGURE 2. The group manifold $M(SO(3))$ is isomorphic to the 3-ball B^3 with antipodal points on the boundary identified, $\mathbf{w} \sim -\mathbf{w} \forall \mathbf{w} \in \partial B_3$.

which lives in the region

$$B_3 = \{\mathbf{w} \in \mathbb{R}^3 : |\mathbf{w}| \leq \pi\} \subset \mathbb{R}^3$$

with boundary

$$\partial B_3 = \{\mathbf{w} \in \mathbb{R}^3 : |\mathbf{w}| = \pi\} \cong S^2.$$

We say that the group manifold $M(SO(3))$ then comes from identifying antipodal points on ∂B_3 ($\mathbf{w} \sim -\mathbf{w} \forall \mathbf{w} \in \partial B_3$). See Fig. 2 for an illustration.

Definition 3.8. A *compact* set is any bounded, closed set in $\mathbb{R}^n$ with $n \geq 0$. For instance, the 2-sphere S^2 is clearly bounded in $\mathbb{R}^3$. But the hyperboloid H^2 (embedded in $\mathbb{R}^3$ as $x^2 + y^2 - z^2 = r^2$) is not bounded, since for any distance r_0 one can construct a point $\mathbf{x}$ on H^2 which has $|\mathbf{x}| > r_0$.

Let us note some properties of the group manifold $M(SO(3))$. It is compact and connected, but it is not simply connected.

Definition 3.9. A space is *simply connected* if all loops on the space are contractible (in the language of algebraic topology, its fundamental group π_1 is trivial).

A bit of intuition for why $M(SO(3))$ is topologically non-trivial: draw a path to the boundary, come out on the antipodal side, and go back to the origin. As it turns out, this is different from S^1 or the torus T^2: whereas these have the full $\mathbb{Z}$ as (part of) their fundamental groups (T^2 is simply $S^1 \times S^1$), if we go around twice in $SO(3)$ we find that this new loop is actually a trivial loop (see Fig. 3). Therefore the fundamental group of $SO(3)$ is not infinite but the cyclic group $\mathbb{Z}_2$ (i.e. the set $\{0, 1\}$ under the group operation $+$ mod 2).

Lecture 4.

Here Comes the SU(n): Thursday, October 11, 2018

Last time, we discussed $SO(3)$ which was a compact submanifold of $GL(n, \mathbb{R})$. But there are also non-compact subgroups we should consider. We introduced the orthogonal group of matrices $M \in O(n)$ which preserve the Euclidean metric on $\mathbb{R}^n$, i.e.

$$g = \text{diag}\{+1, +1, \ldots +1\}, M^T g M = g.$$

But we may also generalize almost immediately to a metric with a different signature.

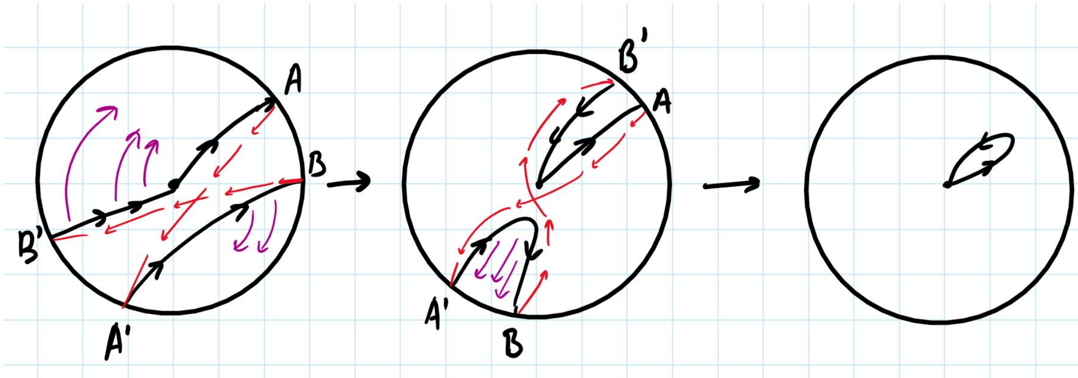

FIGURE 3. A sketch of why the loop which goes through the boundary ∂B_3 twice is homotopic to (can be continuously deformed into) the trivial loop. For simplicity, consider a circular cross-section of B_3 and suppose the loop passes through the boundary at points A ($\sim A'$) and B ($\sim B'$). As we continuously move the point B to A', B' must also move towards A, as we see in the second image. We then pull the bit of loop from A' to B through the boundary and find that the resulting loop is trivial (sketch 3). Solid black lines indicate the actual loop path, red dashed arrows indicate the effect of identifying antipodal points, and purple arrows suggest the direction of loop deformation between each drawing.

Definition 4.1. $O(p,q)$ transformations preserve the metric of signature (p,q) on $\mathbb{R}^{p,q}$, where

$$\eta = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}.$$

Then $O(p,q)$ is defined by

$$O(p,q) = \{M \in GL(p+q,\mathbb{R}) : M^T \eta M = \eta\}.$$

$SO(p,q)$ is defined equivalently as

$$SO(p,q) = \{M \in O(p,q) : \det M = 1\}.$$

Example 4.2. The (full) Lorentz group $O(3,1)$ preserves the Minkowski metric. We could consider $SO(1,1)$, which takes the form

$$M = \begin{pmatrix} \cosh\phi & \sinh\phi \\ \sinh\phi & \cosh\phi \end{pmatrix}$$

with $\phi \in \mathbb{R}$ the rapidity. This is just a Lorentz boost in one direction, parametrized by the rapidity.

It's also useful to discuss subgroups of $GL(n,\mathbb{C})$ (matrices with complex entries).

Definition 4.3. We introduce the *unitary transformations*, defined by

$$U(n) = \{U \in GL(N,\mathbb{C}) : UU^\dagger = I_n\}.$$

Such transformations therefore preserve the inner product of complex vectors $\mathbf{v} \in \mathbb{C}^n$, with $|\mathbf{v}|^2 = \mathbf{v}^\dagger \cdot \mathbf{v}$. These also form a Lie group (we need to look at the constraints imposed by the $UU^\dagger$ condition and apply our implicit function theorem to confirm that this is really a manifold).

The unitary transformations have the condition that since $U \in U(n) \implies U^\dagger U = I_n \implies |\det U|^2 = 1$. Thus $\det U = e^{i\delta}, \delta \in \mathbb{R}$. Whereas in $O(n)$ we had two discrete possibilities for $\det M$ leading to two connected components, we see that in $U(n)$ we can parametrize our matrices by a continuous δ and so we expect $O(n)$ as a manifold to be connected.

Definition 4.4. We may also define the special unitary group, $SU(N)$.

$$SU(n) = \{U \in U(n) : \det U = 1\}.$$

How big is $U(n)$? A priori we get $2n^2$ choices of real numbers. But the matrix equation $UU^\dagger = I$ is constrained since $UU^\dagger$ is Hermitian, and so we get $2 \times \frac{1}{2}n(n-1)$ constraints from the entries above the diagonal $+n$ constraints since the elements on the diagonal are real. Therefore we get $N^2 - n + n = n^2$ constraints, and

$$\dim(U(n)) = 2n^2 - n^2 = n^2.$$

What about for $SU(n)$? Normally $\det U = 1$ would give two constraints for a general complex number, but we know that $\det U = e^{i\delta}$ for $U \in U(n)$, so we only get one constraint out of this condition (effectively setting our parameter δ to 1). Thus

$$\dim(SU(n)) = n^2 - 1.$$

Example 4.5. $SU(1)$ would have dimension $1 - 1 = 0$, which is not interesting, so the first interesting subgroup of $GL(n, \mathbb{C})$ is then $U(1)$, with dimension 1:

$$U(1) = \{z \in \mathbb{C} : |z| = 1\}.$$

This has the group manifold structure of a circle, but we've seen another group with the same manifold structure: $SO(2)$! In light of this, we would like to have some notion that two groups are really "the same," motivating the following definition.

Definition 4.6. A *group homomorphism* is a function $J : G \to G'$ such that

$$\forall g_1, g_2 \in G, J(g_1 g_2) = J(g_1)J(g_2).$$

In other words, the group structure is preserved and group multiplication commutes with applying the homomorphism.

Definition 4.7. An *isomorphism* is a group homomorphism which is a one-to-one smooth map $G \leftrightarrow G'$. We say that two Lie groups G, G' are isomorphic if there exists an isomorphism between them.

Example 4.8. Take a general element $z = e^{i\theta} \in G = U(1, \theta \in \mathbb{R}$. Thus define

$$M(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in G' = SO(2).$$

Then our group homomorphism is

$$J : z(\theta) = e^{i\theta} \to M(\theta) \in SO(2).$$

It's straightforward to check that

$$\begin{aligned} J(z(\theta_1)z(\theta_2)) &= M(\theta_1 + \theta_2) \\ &= M(\theta_1)M(\theta_2) \\ &= J(z(\theta_1))J(z(\theta_2)) \\ &\implies U(1) \simeq SO(2). \end{aligned}$$

Example 4.9. Now consider $G = SU(2)$. $\dim(SU(2) = 2^1 - 1 = 3$, and we can write elements of $SU(2)$ as

$$U = a_0 I_2 + i\mathbf{a} \cdot \boldsymbol{\sigma},$$

where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the Pauli matrices, $a_0 \in \mathbb{R}, \mathbf{a} \in \mathbb{R}^3$, and

$$a_0^2 + |\mathbf{a}|^2 = 1.$$

We've seen another group of the same dimension, $SO(3)$, but we remark that these are *not* isomorphic to each other. From our parametrization of $SU(2)$, we see that $M(SU(2) = S^3$ the three-sphere, but

$$\pi_1(S_3) = \emptyset, \pi_1(M(SO(3)) = \mathbb{Z}_2,$$

so they cannot be isomorphic.

Lie algebras

Definition 4.10. A *Lie algebra* $\mathfrak{g}$ is a vector space (over a field $F = \mathbb{R}$ or $\mathbb{C}$) equipped with a *bracket*. A *bracket* is an operation

$$[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

which has the following properties:

(a) antisymmetry, $\forall X, Y \in \mathfrak{g}, [X, Y] = -[Y, X]$
(b) linearity, $[\alpha X + \beta Y, Z] = \alpha[X, Z] + \beta[Y, Z] \forall \alpha, \beta \in F, \forall X, Y, Z \in \mathfrak{g}$
(c) the Jacobi identity, $\forall X, Y, Z \in \mathfrak{g}, [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.

Note that if a vector space V has an associative multiplication law $* : V \times V \to V$ (that is, $(X * Y) * Z = X * (Y * Z)$), we can make a Lie algebra by simply defining the bracket as

$$[,] = X * Y - Y * X \forall X, Y \in V.$$

This is pretty easy to prove and we will do so on an example sheet. The most obvious choice is V a vector space of matrices and $*$ ordinary matrix multiplication.

The dimension of $\mathfrak{g}$ is the same as the dimension of the underlying vector space V (since we have just equipped V with some extra structure).

Note that we could choose a basis

$$B = \{T^a, a = 1, \dots, n = \dim(\mathfrak{g})\}$$

such that

$$\forall X \in \mathfrak{g}, X = X_a T^a \equiv \sum_{a=1}^{n} X_a T^a, X_a \in F.$$

That is, we can decompose a general element of $\mathfrak{g}$ into its components X_a. Then we observe that for $X, Y \in \mathfrak{g}$, we can always compute

$$[X, Y] = X_a Y_b [T^a, T^b]$$

in this basis T^a.

Definition 4.11. We therefore see that a general Lie bracket is defined by the *structure constants* f_c^{ab}, given by

$$[T^a, T^b] = f_c^{ab} T^c.$$

Once we compute these with respect to a basis, we know how to compute any Lie bracket of two general elements. Since the structure constants come from a Lie bracket, they obey antisymmetry in the upper indices,

$$f_c^{ab} = -f_c^{ab},$$

and also (exercise) a variation of the Jacobi identity,

$$f_c^{ab} f_e^{cd} + f_c^{da} f_e^{cb} + f_c^{bd} f_e^{ca} = 0.$$

Lecture 5.

Lie Algebras from Lie Groups: Saturday, October 13, 2018

Last time, we defined a Lie algebra as a vector space with some extra structure, the Lie bracket $[,]$.

Definition 5.1. Two Lie algebras $\mathfrak{g}, \mathfrak{g}'$ are isomorphic if $\exists$ a one-to-one linear map $f : \mathfrak{g} \to \mathfrak{g}'$ such that

$$[f(X), f(Y)] = f([X, Y] \forall X, Y \in \mathfrak{g}.$$

Therefore the isomorphism respects the Lie bracket structure (with the bracket being taken in $\mathfrak{g}$ or $\mathfrak{g}'$ as appropriate).

Definition 5.2. A subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ is a subset which is also a Lie algebra. This is equivalent to a subgroup in group theory.

Definition 5.3. An ideal of $\mathfrak{g}$ is a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ such that

$$[X, Y] \in \mathfrak{h} \forall X \in \mathfrak{g}, Y \in \mathfrak{h}.$$

This is the equivalent to a normal subgroup in group theory. Note that every $\mathfrak{g}$ has two trivial ideals:

$$\mathfrak{h} = \{0\}, \mathfrak{h} = \mathfrak{g}.$$

Every $\mathfrak{g}$ also has the following two ideals:

Example 5.4. The derived algebra, all elements i such that

$$i = \{[X, Y] : X, Y \in \mathfrak{g}\}.$$

Example 5.5. The centre (center) of $\mathfrak{g}$, $\xi(\mathfrak{g})$:

$$\xi(\mathfrak{g}) = \{X \in g : [X, Y] = 0 \forall Y \in \mathfrak{g}.\}$$

Definition 5.6. An abelian Lie algebra $\mathfrak{g}$ is then one for which $[X, Y] = 0 \forall X, Y \in \mathfrak{g}$ (i.e. $\xi(\mathfrak{g}) = \mathfrak{g}$, the center of the group is the whole group).

Definition 5.7. $\mathfrak{g}$ is simple if it is non-abelian and has no non-trivial ideals. This is equivalent to saying that

$$i(\mathfrak{g}) = \mathfrak{g}.$$

Simple Lie algebras are important in physics because they admit a non-degenerate inner product (related to Killing forms). These ideas will also lead us to classify all complex simple Lie algebras of finite dimension.

Lie algebras from Lie groups The names of these structures makes it seem that they ought to be related in some way. Let's see now what the connection is. Let M be a smooth manifold of dimension D and take $p \in M$ a point on the manifold. Since M is a manifold, we may introduce coordinates in some open set containing p.

Let us call the coordinates

$$\{x_i\}, i = 1, \dots, D$$

and set p to lie at the origin, $x^i = 0$. Now we will denote the tangent space to M at p by $\mathcal{T}_p(M)$, and define the tangent space as the vector space of dimension D spanned by

$$\{\frac{\partial}{\partial x_i}\}, i = 1, \dots, D.$$

A general tangent vector V is then a linear combination of the basis elements, given by components V^i:

$$V = V^i \frac{\partial}{\partial x^i} \in \mathcal{T}_p(M), V^i \in \mathbb{R}.$$

Tangent vectors then act on functions of the coordinates $f(x)$ by

$$Vf = v^i \frac{\partial f(x)}{\partial x^i}|_{x=0}$$

(they are local objects, so they only live at the point $x = 0$). Consider now a smooth curve

$$C : I \subset \mathbb{R} \to M$$

(if we like, one can normalize I to a unit interval) passing through the point p. In coordinates,

$$C : t \in I \mapsto x^i(t) \in \mathbb{R}, i = 1, \dots, D.$$

This curve is smooth if the $\{x^i(t)\}$ are continuous and differentiable.

The tangent vector to the curve C at point p is then

$$V_C \equiv \dot{x}^i(0) \frac{\partial}{\partial x^i} \in \mathcal{T}_p(M)$$

where $\dot{x}^i(t) = \frac{dx^i(t)}{dt}$. This is simply the directional derivative from multivariable calculus. When we act with this tangent vector on a function f, we then get

$$V_c f = \dot{x}^i(0) \frac{\partial f(x)}{\partial x^i}|_{x=0}.$$

Now to compute the Lie algebra $L(G)$ of a Lie group G, let G be a Lie group of dimension D. Introduce coordinates $\{\theta^i\}, i = 1, \dots, D$ in some region around the identity element $e \in G$. Now we can look at the tangent space near the identity,

$$\mathcal{T}_e(G).$$

Note that $\mathcal{T}_e(G)$ is a real vector space of dimension D, and we can define a bracket

$$[,] : \mathcal{T}_e(G) \times \mathcal{T}_e(G) \to \mathcal{T}_e(G)$$

such that

$$(\mathcal{T}_e(G), [,])$$

defines a Lie algebra.

Example 5.8. The easiest case is matrix Lie groups. For instance,

$$G \subset \mathrm{Mat}_n(F)$$

for $n \in \mathbb{N}, F = \mathbb{R}$ or $\mathbb{C}$. We can turn the map from tangent vectors to matrices:

$$\rho : V^i \frac{\partial}{\partial \theta^i} \in \mathcal{T}_e(G) \mapsto V^i \frac{\partial g(\theta)}{\partial \theta^i}|_{\theta=0}$$

such that $g(\theta) \in G \subset \mathrm{Mat}_n(F)$. We will identify $\mathcal{T}_e(G)$ with the span of

$$\{\frac{\partial g(\theta)}{\partial \theta^i}|_{\theta=0}\}, i = 1, \dots D.$$

Effectively, we've parametrized elements of our group (e.g. by our local coordinate system) and then identified the tangent space with the span of the D tangent vectors which describe how our parametrized group elements change with respect to the D coordinates.

Now we have a candidate for the bracket. Let's choose

$$[X, Y] \equiv XY - YX \forall X, Y \in \mathcal{T}_e(G)$$

where XY indicates matrix multiplication. That is, the "bracket" here is really just the matrix commutator. This is clearly antisymmetric and linear, and with a little bit of algebra one can show it also obeys the Jacobi identity. But there's one other condition– the algebra must be closed under the bracket operation. It's not immediately obvious that this is true, so we'll prove it explicitly.

Let C be a smooth curve in G passing through e,

$$C : t \mapsto g(t) \in G, g(0) = I_n.$$

We require that $g(t)$ is at least C^1 smooth, $G(t) \in C^1(M), t \geq 0$. (It has at least a first derivative.) Now consider the derivative

$$\frac{dg(t)}{dt} = \frac{d\theta^i(t)}{dt} \frac{\partial g(\theta)}{\partial \theta^i}.$$

It follows that

$$\dot{g}(0) = \frac{dg(t)}{dt}|_{t=0} = \dot{\theta}^i(0) \frac{\partial g(\theta)}{\partial \theta^i}|_{\theta=0} \in \mathcal{T}_e(G).$$

This is a tangent vector to C at the point e. $\dot{g}(0) \in \mathrm{Mat}_n(F)$, but more generally this element of the tangent space need not be in the group.

Near $t = 0$ we have

$$g(t) = I_n + Xt + O(t^2), X = \dot{g}(0) \in L(G).$$

We expand our curve to first order in t near $t = 0$. For two general elements $X_1, X_2 \in L(G)$, we find curves

$$C_1 : t \mapsto g_1(t) \in G, C_2 : t \mapsto g_2(t) \in G$$

such that

$$g_1(0) = g_2(0) = I_n$$

and

$$\dot{g}_1(0) = X_1, \dot{g}_2(0) = X_2.$$

Then the maps g_1, g_2 can also be expanded to order t^2 near $t = 0$,

$$g_1(t) = I_n + X_1 t + W_1 t^2 + \dots, g_2(t) = I_n + X_2 t + W_2 t^2 + \dots$$

for some $W_1, W_2 \in \mathrm{Mat}_n(F)$. Next time, we'll show that the bracket gives a nice structure for

$$W(t) \equiv g_1^{-1}(t) g_2^{-1}(t) g_1(t) g_2(t).$$

Lecture 6.

Examples of Lie Algebras: Tuesday, October 16, 2018

Today, we'll finish the proof that the tangent space of a Lie group G at the origin, $T_e(G)$, equipped with the bracket operation $[X,Y] = XY - YX$ for $X, Y \in T_e(G)$ forms a Lie algebra. Specifically, we must prove that $L(G)$ is closed under the bracket.

The game plan is as follows. We want to show that for any two elements $X, Y \in T_e(G)$, their Lie bracket $[X,Y]$ is also in the tangent space. Therefore we will explicitly construct a curve in G out of other elements we know are in G such that our new curve has exactly the Lie bracket $[X,Y]$ as its tangent vector near $t = 0$.

Recall that last time, we considered two curves $C_1 : t \mapsto g_1(t) \in G$ and $C_2 : t \mapsto g_2(t) \in G$ which are at least twice differentiable, and by definition the tangent vectors (i.e. first derivative) of these curves give rise to two elements X_1, X_2 in the Lie algebra $L(G)$. These curves had the properties that at $t = 0$,

$$g_1(0) = g_2(0) = I_n$$

with I_n the identity matrix, and

$$\dot{g}_1(0) = X_1, \dot{g}_2(0) = X_2.$$

We proceeded to expand them to order t^2, writing

$$g_1(t) = I_n + X_1 t + W_1 t^2 + O(t^3) \text{ and } g_2(t) = I_n + X_2 t + W_2 t^2 + O(t^3).$$

Now define the element

$$h(t) \equiv g_1^{-1}(t) g_2^{-1}(t) g_1(t) g_2(t).$$

Because $h(t)$ is constructed via group multiplication in G, h is also in G. Under an appropriate reparametrization, this will be the curve we want. We can rewrite this equation as

$$g_1(t) g_2(t) = g_2(t) g_1(t) h(t).$$

Plugging in our expansions of g_1, g_2 we find that

$$g_1(t) g_2(t) = I_n + t(X_1 + X_2) + t^2(X_1 X_2 + W_1 + W_2) + O(t^3)$$

and similarly

$$g_2(t) g_1(t) = I_n + t(X_1 + X_2) + t^2(X_2 X_1 + W_1 + W_2) + O(t^3).$$

If we now expand

$$h(t) = I_n + w_1 t + w_2 t^2 + O(t^3),$$

we find that[9]

$$w_1 = 0, w_2 = X_1 X_2 - X_2 X_1 = [X_1, X_2].$$

Now let us define a new curve,

$$C_3 : s \mapsto g_3(s) = h(+\sqrt{s}) \in G$$

parametrized by some $s \in \mathbb{R}$. We need $t \geq 0$ so $s > 0, s = t^2$. Near $s = 0$, we have

$$g_3(s) = I_n + s[X_1, X_2] + O(s^{3/2}) \implies \dot{g}_3(0) = \frac{g_3(s)}{ds}|_{s=0} = [X_1, X_2] \in L(G).$$

So indeed the bracket operation $[X_1, X_2]$ corresponds to another element in the tangent space.[10] All is well and thus $L(G) = (T_e(G), [,])$ is a real Lie algebra of dimension D. ⊠

[9] It's straightforward, so I'll do it here. Explicitly, if we expand to order t we get $g_2 g_1 W(t) = I + (X_1 + X_2 + w_1)t$. But by comparison to the expression for $g_1 g_2$ we see that $w_1 = 0$. So we have to go to order t^2: $g_2 g_1 W(t) = I + (X_1 + X_2)t + (w_2 + W_1 + W_2 + X_2 X_1)$. Now comparing again we find that $w_2 + X_2 X_1 = X_1 X_2$, or equivalently $w_2 = X_1 X_2 - X_2 X_1 = [X_1, X_2]$.

[10] We might want to make sure that the tangent vector of our curve is really well-defined at $s = 0$– in particular, we might be concerned about $s < 0$. To be really thorough, we can define $\tilde{h}(t) = g_2(t)^{-1} g_1(t)^{-1} g_2(t) g_1(t)$ and by a similar process extend the curve h to negative s. This doesn't add anything to our proof but it can certainly be done and one can check that the first derivatives of h and $\tilde{h}$ match at $s = 0$.

Example 6.1. Let $G = SO(2)$. Then

$$\mathfrak{g}(t) = M(\theta(t)) = \begin{pmatrix} \cos\theta(t) & -\sin\theta(t) \\ \sin\theta(t) & \cos\theta(t) \end{pmatrix}$$

with $\theta(0) = 0$. So the tangent space is spanned by elements of the form

$$\dot{g}(0) = \begin{pmatrix} 0 & -1 \\ +1 & 0 \end{pmatrix} \dot{\theta}(0)$$

and therefore

$$L(SO(2)) = \left\{ \begin{pmatrix} 0 & -c \\ c & 0 \end{pmatrix}, c \in \mathbb{R} \right\}$$

The Lie algebra of $SO(2)$ is therefore the set of 2×2 real antisymmetric matrices.

Example 6.2. Let $G = SO(n)$. Now our curve is $g(t) = R(t) \in SO(n)$ with $R(0) = I_n$, and the defining equation of $SO(n)$ says that

$$R^T(t)R(t) = I_n \forall t \in \mathbb{R}.$$

Differentiating with respect to t (if you like, we're looking at the leading order behavior by expanding $R(0) + \dot{R}(0)t$) we find that

$$\dot{R}^T(t)R(t) + R^T(t)\dot{R}(t) = 0 \implies X^T + X = 0,$$

where as usual we let $X = \dot{R}(0) = \frac{dR(t)}{dt}|_{t=0}$. Therefore we learn that

$$X^T = -X,$$

or in other words, X is antisymmetric.

One might worry about the determinant condition, but in fact since any matrix close to the identity already has determinant 1 (recall that $O(n)$ has two connected components with $\det R = \pm 1$), the $\det R = 1$ condition does not impose an additional constraint, so moreover

$$L(O(n)) = L(SO(n)) = \{X \in \text{Mat}_n(\mathbb{R}) : X^T = -X.\}$$

The Lie algebra of $O(n)$ and $SO(n)$ is the set of real $n \times n$ antisymmetric matrices, and by counting constraints we see it has dimension $\frac{1}{2}n(n-1)$.

Example 6.3. We can play the same game with $G = SU(n)$. Let $g(t) = U(t) \in SU(n), U(0) = I_n$. THen

$$U^\dagger(t)U(t) = I_n \forall t \in \mathbb{R}.$$

Differentiating and setting $t = 0$ we find that

$$Z^\dagger + Z = 0$$

where $Z = \dot{U}(0) \in L(SU(n))$.

We also have the condition that $\det U(t) = 1 \forall t \in \mathbb{R}$.[11] Let's expand $U(t) = I_n + Zt + O(t^2)$ near $t = 0$. As an exercise, one may prove that $\det U(t) = 1 + \text{Tr}(Z)t + O(t^2)$, and so $\det U(t) = 1 \forall t \implies \text{Tr}(Z) = 0$. Thus

$$L(SU(n)) = \{Z \in \text{Mat}_n(\mathbb{C}) : Z^\dagger = -Z, \text{Tr}(Z) = 0, \}$$

the set of complex $n \times n$ antihermitian traceless matrices.

What is the dimension of $L(SU(n))$? We get $2 \times \frac{1}{2}n(n-1)$ real constraints from the entries above the diagonal, n constraints forcing the real parts of the diagonal entries to be zero, and 1 constraint from the tracelessness condition. Thus we have $n^2 + 1$ total constraints and dimension $2n^2 - (n^2 + 1) = n^2 - 1$.

[11] This didn't matter in the real case, but here we don't have the same disconnected structure as in $O(n)$. The determinant need only have unit magnitude, $|\det U|^2 = 1$, and so we get an extra constraint. Practically speaking, we see that antisymmetry already forced $X \in L(O(n))$ to be traceless, whereas this is not the case for $SU(n)$.

Example 6.4. With our results for the general $SU(n)$ in hand, we can take the specific example of $G = SU(2)$. The Lie algebra is the set of 2×2 traceless antihermitian matrices, and it should have dimension $2^2 - 1 = 3$. But we already know of three linearly independent matrices which (nearly) satisfy this property: they are the Pauli matrices from quantum mechanics.

$$\sigma_a = \sigma_a^\dagger, \mathrm{Tr}\sigma_a = 0, a = 1,2,3$$

We can define $T^a = -\frac{1}{2}i\sigma_a$ (so that T^a is antihermitian rather than hermitian). Recall the Pauli matrices obey

$$\sigma_a \sigma_b = \delta_{ab} I_2 + i\epsilon_{abc}\sigma_c,$$

so it is straightforward to compute the Lie bracket on T^a,

$$[T^a, T^b] = -\frac{1}{4}[\sigma_a, \sigma_b] = -\frac{1}{2}i\epsilon_{abc}\sigma_c = f_c^{ab} T^c$$

where

$$f_c^{ab} = \epsilon_{abc}$$

(note that indices up and down are not so important here– they are just labels and do not indicate any sort of covariant behavior as in relativity).

However, we can also compare with $SO(3)$, which we computed the Lie group for earlier. Recall that

$$L(SO(3)) = \{3 \times 3 \text{real antisymmetric matrices}\},$$

and $\dim(L(SO(3)) = \frac{1}{2}n(n-1)|_{n=3} = 3$. A convenient basis is

$$\tilde{T}^1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \tilde{T}^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \tilde{T}^3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

These are clearly linearly independent and satisfy the antisymmetry condition. More compactly, we can also write

$$\tilde{T}^a_{bc} = -\epsilon_{abc},$$

and then with respect to this basis, the Lie bracket is

$$[\tilde{T}^a, \tilde{T}^b] = f_c^{ab} \tilde{T}^c$$

where $f_c^{ab} = \epsilon_{abc}, a, b, c = 1,2,3$.

But these are exactly the same structure constants we found for $L(SU(2))$, and so we find that the Lie algebras are isomorphic:

$$L(SO(3)) \simeq L(SU(2)).$$

This is interesting since $SO(3) \not\simeq SU(2)$, i.e. the original groups are *not* isomorphic.[12] However, it will turn out that $SO(3) = SU(2)/\mathbb{Z}_2$, i.e. one can say that $SU(2)$ is the double cover of $SO(3)$.

Lecture 7.

Lost in Translation(s): Thursday, October 18, 2018

Today, we'll revisit the idea of Lie algebras from Lie groups. A Lie group is a very special type of manifold because it is equipped with a group structure, and this means that it comes with some nice maps on the manifold built-in.

Definition 7.1. In particular, for each element $h \in G$ a Lie group, we have smooth maps

$$L_h : G \to G, g \in G \mapsto hg \in G$$

and

$$R_h : G \to G, g \in G \mapsto gh \in G$$

known as *left-* and *right-translations*.

[12] One way to see this is by remembering that $SO(3)$ has the manifold structure of B_3, while $SU(2)$ has the structure of S^3.

We'll understand the meaning of this term more clearly in just a minute, but we can already see that these maps are *surjective* (their image includes every element of the group),

$$\forall g' \in G \exists g = h^{-1}g' \in G \implies L_h(g) = g'$$

and *injective* (for every element of the image, the inverse is unique), $\forall g, g' \in G, L_h(g) = L_h(g') \implies g = g'$ since

$$L_h(g) = L_h(g') \implies hg = hg' \implies g = g'$$

by the existence of unique inverses under group multiplication. ☒

Thus the *inverse map*,

$$(L_h)^{-1} = L_{h^{-1}},$$

also exists and is smooth.

Definition 7.2. We say that L_h and R_h are *diffeomorphisms* of G (i.e. an isomorphism such that both the map and its inverse are smooth).

To concretely understand how L_h acts on elements of G, we therefore introduce coordinates $\{\theta^i\}, i = 1, \ldots, D$ in some region containing the identity element e:

$$g = g(\theta) \in G, g(0) = e.$$

Let $g' = g(\theta') = L_h(g) = hg(\theta)$. A priori, g' need not be in the same coordinate patch as g, but because G is a manifold, we have some nice transition functions which will allow us to describe g' in compatible local coordinates.

To avoid these complications, let us assume for now that g and g' are in the same coordinate patch as g. In coordinates, L_h is then specified by D real functions on the coordinates θ,

$$\theta'^i = \theta'^i(\theta), i = 1, \ldots, D.$$

As L_h is a diffeomorphism, the *Jacobian matrix*

$$J^i_j(\theta) = \frac{\partial \theta'^i}{\partial \theta^j}$$

exists and is invertible (i.e. $\det J \neq 0$).

Definition 7.3. However, the map $L_h : G \to G$ now induces a map L_h^* from tangent vectors at g to the tangent space to $L_h(g) = hg \in G$. That is,

$$L_h^* : \mathcal{T}_g(G) \to \mathcal{T}_{hg}(G).$$

In coordinates, we see that L_h^* maps a tangent vector $V = V^i \frac{\partial}{\partial \theta^i}$ in the original coordinates:

$$L_h^* : V = V^i \frac{\partial}{\partial \theta^i} \in \mathcal{T}_g(G) \mapsto V' = V'^i \frac{\partial}{\partial \theta'^i} \in \mathcal{T}_{hg}(G)$$

with

$$V'^i = J^i_j(\theta) V^j.$$

We call this map L_h^* the *differential* of L_h.

In words, we have moved a tangent vector at g to hg by rewriting it in terms of the derivatives $\partial/\partial\theta'$ with respect to the local coordinates at hg, and the components V^i transform by multiplication by the Jacobian. This is pretty powerful– left translation lets us move tangent vectors from near the identity to anywhere we like on the group manifold! We'll see that this has consequences for the structure of the Lie algebra as well.

Definition 7.4. A *vector field* V on G specifies a tangent vector $V(g) \in \mathcal{T}_g(G)$ at each point $g \in G$. In coordinates,

$$V(\theta) = V^i(\theta) \frac{\partial}{\partial \theta^i} \in \mathcal{T}_{g(\theta)}(G).$$

We say a vector field is smooth if the component functions $V^i(\theta) \in \mathbb{R}, i = 1, \ldots, D$ are differentiable.

In fact, starting from a single tangent vector at the identity

$$\omega \in \mathcal{T}_e(G),$$

we can then define a vector field using left-translation.

$$V(g) = L_g^*(\omega)\forall g \in G.$$

So now we're leaving the tangent vector fixed and moving it all around our manifold using the differential map L_g^*. But since L_g^* is smooth and invertible, $V(g)$ is smooth and *non-vanishing*. To see this, suppose L_g^* sent some $\omega \neq 0$ to $v' = 0$. Since the components of ω transform with the Jacobian matrix, this implies that the Jacobian matrix has a zero eigenvalue (i.e $0 = J^i_j V^j$). But we said the Jacobian matrix was invertible, so this is a contradiction (otherwise J^{-1} could send the zero vector to something nonzero, ${J^{-1}}^i_j 0 = V^i, V^i \neq 0$).

Then starting from a basis $\{\omega_a\}, a = 1, \ldots, D$ for $\mathcal{T}_e(G)$, we get D *independent nowhere-vanishing vector fields on* G,

$$V_a(g) = L_g^*(\omega_a), a = 1, \ldots, D.$$

This turns out to already be a very strong constraint on what manifolds admit Lie groups.

Example 7.5. By the "hairy ball theorem," any smooth vector field on S^2 has at least two zeros.[13] Therefore $M(G) \not\simeq S^2$.

In fact, if G is compact and $\dim(G) = 2$, the only possible manifold structure is $M(G) = T^2 = S^1 \times S^1$ the torus, corresponding to the group structure $U^1 \times U^1$.

Definition 7.6. Note that $V_a(g), a \in 1, \ldots, D$ are called *left-invariant vector fields* on G. They obey

$$L_h^* V_a(g) = L_h^* \circ L_g^*(\omega_a) = L_{hg}^*(\omega_a) = V_a(hg).$$

This has some very nice consequences for the structure of the Lie algebra– for more on this, see the appendix to Prof. Dorey's notes (which I may type here later).

For matrix Lie groups, $G \subset \mathrm{Mat}_n(F), n \in \mathbb{N}, F = \mathbb{R}$ or $\mathbb{C}$, we find that $\forall h \in G, X \in L(G)$ we get a map

$$L_h^* : \mathcal{T}_e(G) \to T_h(G).$$

Recall that in general the elements in the Lie algebra are not in the Lie group itself (e.g. the elements of $U(n)$ are unitary but the elements of $L(U(n))$ are anti-hermitian). However, since L_h^* is a map on the tangent space, it turns out that L_h^* then induces a map on the elements of the Lie algebra:

$$L_h^*(X) = hX \in \mathcal{T}_h(G).$$

The proof is as follows: consider a curve

$$C : t \in \mathbb{R} \mapsto g(t) \in G$$

with $g(0) = e, \dot{g}(0) = X \in L(G)$. Near $t = 0$ we can Taylor expand,

$$g(t) \simeq I_n + tX + O(t^2).$$

Define a new curve

$$C' : t \in \mathbb{R} \mapsto h(t) = h \cdot g(t) \in G$$

with $h \in G$. Near $t = 0$, $h(t)$ then has the expansion

$$h(t) \simeq h + thX + O(t^2)$$

Therefore

$$hX \in \mathcal{T}_h(G),$$

so we can quite sensibly define a map from the Lie algebra (defined locally at the origin) to the tangent space of anywhere else we like on the manifold.

Equivalently, given any smooth curve

$$C : t \in \mathbb{R} \mapsto g(t) \in G,$$

[13] Or one that is "double zero."

with
$$\dot{g}(t) \in \mathcal{T}_{g(t)}(G) \implies g^{-1}(t)\dot{g}(t) = L^*_{g^{-1}(t)}(\dot{g}(t)) \in L(G) \forall t \in \mathbb{R}.$$
In words, we can simply take any smooth curve on G and move it back to the origin, and then its first derivative is in the tangent space at the origin, i.e. the Lie algebra $L(G)$.

Conversely, given $X \in L(G)$ we can reconstruct a curve $C_X : \mathbb{R} \to G, t \mapsto g(t)$ with
$$g^{-1}(t)\frac{dg(t)}{dt} = X \forall t \in \mathbb{R}.$$
Our goal is then to solve this ordinary differential equation with boundary condition $g(0) = I_n$. We'll define the *matrix exponential* (likely familiar from quantum mechanics). For a matrix $M \in \mathrm{Mat}_n(F)$, we use the Taylor series of the exponential to write
$$\exp(M) \equiv \sum_{l=0}^{\infty} \frac{1}{l!} M^l \in \mathrm{Mat}_n(F).$$
If we now set
$$g(t) = \exp(tX) = \sum_{l=0}^{\infty} \frac{1}{l!} t^l X^l,$$
then it's immediate that $g(0) = \exp(0) = I_n$ and
$$\begin{aligned} \frac{dg(t)}{dt} &= \sum_{l=1}^{\infty} \frac{1}{(l-1)!} t^{l-1} X^l \\ &= \exp(tX)X \\ &= g(t)X. \quad \boxtimes \end{aligned}$$
Therefore $g(t)$ solves the differential equation and we say that the exponential map takes the Lie algebra to the Lie group.

Lecture 8.

Representation Matters: Saturday, October 20, 2018

Previously, we defined the exponential map
$$g(t) = \exp(tX) = \sum_{l=0}^{\infty} \frac{1}{l!} t^l X^l.$$
In the exercises (Example Sheet 1, Q10) we'll check explicitly that for $X \in L(SU(n))$, we have $\exp(tX) \in SU(N) \forall t \in \mathbb{R}$. We'll also show in a separate question (Example Sheet 2, Q1) that for a choice of $X \in L(G)$ with G a Lie group and J an interval with $J \subset \mathbb{R}$, $S_X = \{g(t) = \exp(tX\} \forall t \in J \subset \mathbb{R}$ forms an abelian subgroup of G. We call this a one-parameter subgroup.

Now we might be interested to reconstruct G from $L(G)$. Setting $t = 1$ we get a map
$$\exp : L(G) \to G,$$
and this map is one-to-one in some neighborhood of the identity e. (We haven't proved this but it's true.) Then given $X, Y \in L(G)$ we would also like to reconstruct the group multiplication from the Lie algebra, and the solution to this will be the *Baker-Campbell-Hausdorff (BCH) formula*.

For $X, Y \in L(G)$ define
$$g_X = \exp(X), g_Y = \exp(Y)$$
and
$$g_X^\epsilon(x) = \exp(\epsilon X), g_Y^\epsilon(Y) = \exp(\epsilon Y).$$
Then their product is
$$g_X g_Y = \exp(Z) \in G, z \in L(G).$$
Expanding out, we find that
$$\left(\sum_{l=0}^{\infty} \frac{X^l}{l!}\right)\left(\sum_{l'=0}^{\infty} \frac{Y^{l'}}{l'!}\right) = \sum_{m=0}^{\infty} \frac{Z^m}{m!}$$

and one may work out the terms order by order– it looks something like this.

$$Z = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}([X,[X,Y]] - [Y,[X,Y]) + \ldots \in L(G),$$

and we know that this is in the Lie algebra since it is made up of X, Y, and brackets of X and Y which are guaranteed to be in the Lie algebra. Moreover this generalizes to Lie algebras that aren't matrix groups, since the construction only uses the vector space structure of $L(G)$ and the Lie bracket.

$L(G)$ therefore determines G in a neighborhood of the identity (up to the radius of convergence of $\exp Z$, anyway). The exponential map may *not* be globally one-to-one, however. For instance, it is not surjective when G is not connected.

Example 8.1. For $G = O(n)$,

$$L(O(n)) = \{X \in \mathrm{Mat}_n(\mathbb{R}) : X + X^T = 0\}.$$

Then $X \in L(O(n)) \implies \mathrm{Tr}X = 0$. Now let $g = \exp(X)$, $X \in L(O(n)$. We have a nice identity[14] that

$$\det(\exp X) = \exp(\mathrm{Tr}X),$$

and since $\mathrm{Tr}X = 0, \det(\exp X) = 1$. Therefore $\exp(X) \in SO(n) \subset O(n)$.

We'll mention another non-proven fact– for G compact, the image of the exp map is the connected component of the identity. This squares with what we just showed for $O(n)$.

Our map can also fail to be injective when G has a $U(1)$ subgroup.

Example 8.2. For $G = U(1)$, we have

$$L(U(1)) = \{X = ix \in \mathbb{C} : x \in \mathbb{R}\}.$$

Thus $g = \exp(X) = \exp(ix)$, but the Lie algebra elements have a degeneracy where ix and $ix + 2\pi i$ yield the same group element (by Euler's formula) under the exp map.

Let's now return to our discussion of $SU(2)$ vs. $SO(3)$. We saw that $L(SU(2)) \simeq L(SO(3))$, and so we can construct a double-covering, i.e. a globally 2:1 map $d : SU(2) \to SO(3)$ with $d : A \in SU(2) \mapsto d(A) \in SO(3)$. One can write the map explicitly as

$$d(A)_{ij} = \frac{1}{2}\mathrm{tr}_2(\sigma_i A \sigma_j A^\dagger).$$

However, d is not one-to-one since $d(A) = d(-A)$. But we'll explore the properties of this map more on Example Sheet 2. Recall that $SU(2) \simeq S^3$ the three-sphere. If we therefore quotient out by this map, this is the same as identifying antipodal points on the three-sphere. That is, this map provides an isomorphism

$$SO(3) \simeq SU(2)/\mathbb{Z}_2$$

where $\mathbb{Z}_2 = \{I_2, -I_2\}$ is the centre of $SU(2)$, which is a discrete (normal) subgroup of $SU(2)$.[15]

Put another way, $SO(3)$ is the upper hemisphere U^+ of the three-sphere S^3 with antipodal identification on the equation S^2. But the upper hemisphere U^+ is homeomorphic to the three-ball B_3, with $\partial B_3 = S^2$. So the quotient is the same thing as chopping S^3 in half, flattening out the upper hemisphere $U^+ \to B^3$ and identifying antipodal points on the equator $\partial B_3 = S^2$.

Definition 8.3. For a Lie group G, a *representation* D is a map

$$D : G \to \mathrm{Mat}_n(F) \text{ with } \det M \neq 0.$$

Equivalently we could call this a map to $GL(n, F)$. That is, a representation takes us from a Lie group to a set of invertible matrices such that the group multiplication is preserved by the map,

$$\forall g_1, g_2 \in G, D(g_1)D(g_2) = D(g_1 g_2).$$

[14]To prove this, consider a basis where X is diagonal, $X_{ij} = \delta_{ij}\lambda_i$, with λ_i the eigenvalues of X. Then powers of X are given by $X^n_{ij} = \delta_{ij}\lambda_i^n$ and the matrix exponential is simply the matrix with the exponential of each diagonal entry, $(\exp X)_{ij} = \delta_{ij}\exp(\lambda_i)$. It follows that the determinant of the exponential is $\Pi_i \exp(\lambda_i) = \exp(\sum_i \lambda_i)$, which is just the exponential of the sum of the eigenvalues.

[15]A subgroup $H \subset G$ is normal if $gHg^{-1} = H \forall g \in G$. Then we define the quotient G/H to be the original group under identification of the equivalence classes corresponding to the elements of the normal subgroup. Normal subgroups "tile" the group– they separate it into distinct cosets, so it makes good sense to quotient ("mod out") by a normal subgroup.

For a Lie group specifically, we also require that the manifold structure is preserved, so that D is a smooth map (continuous and differentiable). When the map is injective, we say that the representation is *faithful*, but in general representations may be of lower dimension (e.g. the trivial representation where we send every group element to the identity matrix).

Definition 8.4. For a Lie algebra $\mathfrak{g}$, a *representation* d is a map

$$d : \mathfrak{g} \to \mathrm{Mat}_n(F).$$

Note that the zero matrix is part of the Lie algebra since a Lie algebra has a vector space structure, so it won't make sense to require that $\det M \neq 0$. All we require is that this map d has the properties that

- it preserves the bracket operation, $[d(X_1), d(X_2)] = d([X_1, X_2])$ where $[d(X_1), d(X_2)]$ is now the matrix commutator.
- the map is linear, so it preserves the vector space structure: $d(\alpha X_1 + \beta X_2) = \alpha d(X_1) + \beta d(X_2) \forall X_1, X_2 \in \mathfrak{g}, \alpha, \beta \in F$.

The *dimension of a representation* is then the dimension n of the corresponding matrices we're using in the image of our map d or D. The matrices in the image naturally act on vectors living in a vector space $V = F^n$ (i.e. column vectors with n entries in the field F). We call this the *representation space*.

Next time, we'll show that representations of the Lie group have a natural correspondence to representations of the Lie algebra.

Lecture 9.

Representations All the Way Down: Tuesday, October 23, 2018

Last time, we started discussing representations of Lie groups. That is, a representation D is a map from a Lie group G to matrices $GL(n, F)$ over a field such that D is smooth and the group multiplication is preserved,

$$\forall g_1, g_2 \in G, D(g_1)D(g_2) = D(g_1 g_2)$$

(where the multiplication on the LHS is taken to be ordinary matrix multiplication). The field F is usually $\mathbb{R}$ or $\mathbb{C}$ and $n \in \mathbb{N}$ is called the *dimension* of the representation. In general $\dim D = n \neq \dim G$. A subtle point: the dimension of the representation is the dimension of the target space $GL(n, F)$. We'll see some examples of this shortly.

Note that this implies that

$$D(e)D(g) = D(g) \forall g \in G \implies D(e) = I_n$$

and similarly

$$D(g)D(g^{-1}) = D(gg^{-1}) = D(e) = I_n \implies D(g^{-1}) = (D(g))^{-1}.$$

Now consider a matrix Lie group,

$$G \subset \mathrm{Mat}_m(\tilde{F})$$

($\tilde{F}$ could be a different field). For each $X \in L(G)$ the Lie algebra of G, construct a curve in G,

$$C : t \in \mathbb{R} \mapsto g(t) \in G$$

such that $g(0) = I_m, \dot{g}(0) = X$. If we have a representation D of G, then $D(g(t))$ is a curve in $\mathrm{Mat}_n(F)$. Let us now define

$$d(X) \equiv \frac{d}{dt} D(g(t))|_{t=0} \in \mathrm{Mat}_n(F).$$

We claim that $d(X)$ is then a representation of the Lie algebra $L(G)$ corresponding to the representation D of the Lie group G.

Near $t = 0$, we can certainly expand $D(g(t))$ as

$$D(g(t)) = I_n + t d(X) + O(t^2).$$

Let us take $X_1, X_2 \in L(G)$ and play our usual game: we construct curves C_1, C_2 such that

$$C_1 : t \mapsto g_1(t), C_2 : t \mapsto g_2(t)$$

with

$$g_1(0) = g_2(0) = I_m, \dot{g}_1(0) = X_1, \dot{g}_2(0) = X_2.$$

We will show that multiplication of these curves in the right way produces an element corresponding to the Lie bracket.

Consider the curve

$$h(t) = g_1^{-1}(t)g_2^{-1}(t)g_1(t)g_2(t) \in G.$$

Previously, we expanded g_1 and g_2 and showed that $h(t)$ can be written as

$$h(t) = I_m + t^2[X_1, X_2] + O(t^3).$$

Suppose we now pass this curve to the representation of G and calculate $D(h(t))$. Since a representation preserves group multiplication, we get

$$D(h(t)) = D(g_1)^{-1}D(g_2)^{-1}D(g_1)D(g_2).$$

But we can also use our map on the Taylor expansion of h.

$$\begin{aligned} D(h) &= D(I_m + t^2[X_1, X_2] + O(t^3)) \\ &= D(I_m) + t^2\left(\frac{d}{dt^2}D(h(t))|_{t=0}\right) + O(t^3) \\ &= I_n + t^2 d([X_1, X_2]) + O(t^3) \end{aligned}$$

where we have used the fact that $[X_1, X_2]$ is the coefficient for t^2 in $h(t)$ (if you like, you can think of h as a function of t^2, or reparametrize h as we did when initially constructing the Lie algebra from the tangent space of G) so that $\frac{d}{dt^2}D(h(t^2))|_{t^2=0} = d([X_1, X_2])$.

Expanding the individual terms in the group multiplication we get

$$D(g_1) = D(I_m + tX_1 + \ldots) = I_n + td(X_1) + O(t^2)$$

and

$$D(g_1)^{-1} = [I_m + td(X_1) + O(t^2)]^{-1} = I_n - td(X_1) + O(t^2).$$

If we multiply it all out, we get that

$$D(g_1)^{-1}D(g_2)^{-1}D(g_1)D(g_2) = I_m + t^2[d(X_1), d(x_2)].$$

So indeed the bracket is preserved under the representation map:

$$d([X_1, X_2]) = [d(X_1), d(X_2)]. \quad \boxtimes$$

Exercise 9.1. Given a representation d of $L(G)$, show that in some neighborhood of the identity e, $g = \exp(X), X \in L(G)$, show that

$$D(g) = D[\exp X] = \exp(d(X)).$$

Show that $g_1 = \exp(X_1), g_2 \exp(X_2), X_1, X_2 \in L(G)$, the group multiplication is preserved by D, $D(g_1g_2) = D(g_1)D(g_2)$.

We'll now consider representations of Lie algebras in more depth. One of the nice features of the Cartan classification of Lie groups is that it will also classify their representations.

Let $\mathfrak{g}$ be a Lie algebra of dimension D. Here are some representations of $\mathfrak{g}$.

Definition 9.2. The *trivial representation* d_0 maps all elements of $\mathfrak{g}$ to the number 0:

$$d_0(X)\forall X \in \mathfrak{g} \implies \dim(d_0) = 1.$$

Trivial representations correspond to invariants– all elements of the algebra are mapped to zero and by the exponential map, all group elements are the identity.

Definition 9.3. If $\mathfrak{g} = L(G)$ for some matrix Lie group, $G \subset \mathrm{Mat}_n(F)$, we have the *fundamental representation* d_f with

$$d_f(X) = X\forall X \in \mathfrak{g} \implies \dim(d_f) = n.$$

That is, we just take the element of the Lie algebra and represent it by itself.

Definition 9.4. All Lie algebras have an *adjoint representation*, d_{Adj}, with

$$\dim(d_{Adj}) = \dim(\mathfrak{g}) = D$$

(where D is the dimension of the Lie algebra).

For all $X \in \mathfrak{g}$, we define a linear map

$$ad_X : \mathfrak{g} \to \mathfrak{g}$$

by

$$Y \in \mathfrak{g} \mapsto ad_X(Y) = [X, Y] \in \mathfrak{g}.$$

Since ad_X is a linear map between vector spaces of dimension D, it is equivalent to a $D \times D$ matrix. Choosing a basis

$$B = \{T^a, a = 1, \ldots, D\}$$

for $\mathfrak{g}$ and setting $X = X_a T^a, Y = Y_a T^a$, we get

$$[X, Y] = X_a Y_b [T^a, T^b] = X_a Y_b f_c^{ab} T^c.$$

In this basis, we therefore have the explicit form of the adjoint map:

$$[ad_X(Y)]_c = (R_X)_c^b Y_b$$

with

$$(R_X)_c^b \equiv X_a f_c^{ab},$$

where R_X is a $D \times D$ matrix.

We can then define the *adjoint representation* by

$$d_{adj}(X) = ad_X \forall X \in \mathfrak{g},$$

or with respect to a basis,

$$[d_{adj}(X)]_c^b = (R_X)_c^b \forall X \in \mathfrak{g}, b, c = 1, \ldots, D.$$

We can then check the defining properties of a representation.

i)

$$\forall X, Y \in \mathfrak{g}, [d_{Adj}(X), d_{Adj}(Y)] = d_{adj}([X, Y]).$$

Proof: $d_{Adj}(X) = ad_X, d_{Adj}(Y) = ad_Y$. The $\forall Z \in \mathfrak{g},$, composing the ad maps gives us

$$(d_{adj}(X) \circ d_{Adj}(Y))(Z) = [X, [Y, Z]]$$

and in the other order,

$$(d_{adj}(Y) \circ d_{Adj}(X))(Z) = [Y, [X, Z]].$$

Evaluating the RHS of our expression, we have

$$d_{Adj}([X, Y])(Z) = ad_{[X,Y]} Z = [[X, Y], Z].$$

Subtracting the LHS from the RHS, we can rewrite as

$$\begin{aligned}
(\text{LHS} - \text{RHS})(Z) &= [X, [Y, Z]] - [Y, [X, Z]] - [[X, Y, Z] \\
&= [X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] \\
&= 0
\end{aligned}$$

using the antisymmetry property of the bracket and the Jacobi identity.

ii) $\forall X, Y \in \mathfrak{g}, \alpha, \beta \in F$ we have

$$d_{Adj}(\alpha X + \beta Y) = \alpha d_{Adj}(X) + \beta d_{Adj}(Y),$$

which holds due to the linearity of ad_X, ad_Y.

Lecture 10.

Representation Theory of SU(2): Thursday, October 25, 2018

Today, we'll consider the consequences of some specific representations and their structures.

Definition 10.1. Two representations R_1 and R_2 are isomorphic if $\exists$ a matrix S such that

$$R_2(X) = SR_1(X)S^{-1}\forall X \in \mathfrak{g}.$$

Note this must be the same matrix S: that is, R_2 and R_1 are related by a change of basis. If so, we denote this as

$$R_1 \cong R_2.$$

Definition 10.2. A representation R with representation space V has an *invariant subspace* $U \subset V$ if

$$R(X)u \in U \forall X \in \mathfrak{g}, u \in U.$$

(This is equivalent to our ideals in Lie algebras and normal subgroups in group theory.)

Any representation has two trivial invariant subspaces: they are the vector $U = \{0\}$ and $U = V$ the whole representation space.

Definition 10.3. An *irreducible representation* (irrep) of a Lie algebra has no non-trivial invariant subspaces.

With these definitions in hand, let's look at the representation theory of $L(SU(2))$. It's useful to us to write down a basis for the Lie algebra $L(SU(2))$:

$$\{T^a = -\frac{1}{2} i\sigma_a, a = 1, 2, 3\}$$

with σ_a the Pauli matrices. We calculated the structure constants:

$$[T^a, T^b] = f^{ab}_c T^c$$

with $f^{ab}_c = \epsilon_{abc}$ (the alternating tensor/symbol) and $a, b, c = 1, 2, 3$. Let's do something kind of strange– we'll write a new complex basis,

$$\begin{aligned} H &\equiv \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ E_+ &\equiv \frac{1}{2}(\sigma_1 + i\sigma_2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ E_- &\equiv \frac{1}{2}(\sigma_1 - \sigma_2) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \end{aligned}$$

This is really a basis for a somewhat bigger space, the complexified Lie algebra

$$L_{\mathbb{C}}(SU(2)) = \mathrm{Span}_{\mathbb{C}}\{T^a, a = 1, 2, 3\}.$$

For now, we'll simply note that for $X \in L(SU(2))$, we can certainly rewrite X as

$$X = X_H H + X_+ E^+ + X_- E^-,$$

where $X_H \in i\mathbb{R}$ and $X_+ = (\bar{X}_-)$. This is called the Cartan-Weyl basis for $L(SU(2))$.[16] In this basis, a general element takes the form

$$X = \begin{pmatrix} X_H & X_+ \\ X_- & -X_H \end{pmatrix}.$$

This is certainly traceless, and $X \in L(SU(2)) \iff X$ is antihermitian, i.e. $X_H \in i\mathbb{R}, X_+ = -(\bar{X}_-)$.

This basis has some nice properties. For instance, we see that

$$\begin{aligned} [H, E_\pm] &= \pm 2E_\pm \\ \text{and } [E_+, E_-] &= H. \end{aligned}$$

[16] We've been writing $L(G)$ to distinguish the Lie algebra from the corresponding Lie group G, but other texts may use the convention of writing $su(2)$ using lowercase letters or the Fraktur script $\mathfrak{su}(2)$ for the Lie algebra. Just a convention to be aware of.

Hence the ad map takes a very simple form: $ad_H(E_\pm) = \pm 2E_\pm, ad_H(H) = 0$. We also have $ad_H(X) = [H, X] \forall X \in L_\mathbb{C}(SU(2))$. This describes a general X, but note that in this basis, our basis vectors $\{E_+, E_-, H\}$ are eigenvectors of

$$ad_H : L(SU(2)) \to L(SU(2)).$$

That is, we have chosen a basis that diagonalizes the ad map, and its eigenvalues $\{+2, -2, 0\}$ are called *roots*.

Definition 10.4. Consider a representation R of $L(SU(2))$ with a representation space V. We *assume* that $R(H)$ is also diagonalizable. Then the representation space V is spanned by eigenvectors of $R(H)$, with

$$R(H)v_\lambda = \lambda v_\lambda : \lambda \in \mathbb{C}.$$

The eigenvalues λ are called *weights* of the representation R.

Definition 10.5. For such a representation, we call $E_\pm$ the *step operators* (cf. the ladder operators from quantum mechanics).

In particular,

$$\begin{aligned} R(H)R(E_\pm)v_\lambda &= (R(E_\pm)R(H) + [R(H), R(E_\pm)])v_\lambda \\ &= (\lambda \pm 2)R(E_\pm)v_\lambda. \end{aligned}$$

We see that the vectors $R(E_\pm)v_\lambda$ we got from acting on eigenvectors of $R(H)$ with the step operators are also eigenvectors of $R(H)$ with new eigenvalues $\lambda \pm 2$.

Note that a finite dimensional representation R of $L(SU(2))$ must have a highest weight $\Lambda \in \mathbb{C}$, or else we could just keep acting with the raising operator E_+ to get more linearly independent vectors. (We can play a similar trick assuming only a lowest weight– this is what led us to the ladder of harmonic oscillator states.) If there is a highest state, we have

$$\begin{aligned} R(H)v_\Lambda &= \Lambda v_\Lambda \\ R(E_+)v_\Lambda &= 0 \end{aligned}$$

If R is irreducible, then all the remaining basis vectors of V can be generated by acting with $R(E_-)$ (that is, there is only one ladder of states to construct). We get

$$V_{\Lambda-2n} = (R(E_-))^n v_\Lambda, n \in \mathbb{N}.$$

What happens if we now try to raise the lowered states back up? The result is as nice as we could have hoped– we will get back our old states, up to some normalization.

$$\begin{aligned} R(E_+)v_{\Lambda-2n} &= R(E_+)R(E_-)v_{\Lambda-2n+2} \\ &= (R(E_-)R(E_+) + [R(E_+), R(E_-)])v_{\Lambda-2n+2} \\ &= R(E_-)R(E_+)v_{\Lambda-2n+2} + (\Lambda - 2n + 2)v_{\Lambda-2n+2}. \end{aligned}$$

where we have used the fact that the representation preserves the bracket structure.

Looking at the lowest-n cases, we can now take $n = 1$ to find

$$R(E_+)v_{\Lambda-2} = \Lambda v_\Lambda$$

and then for $n = 2$,

$$\begin{aligned} R(E_+)v_{\Lambda-4} &= R(E_-)R(E_+)v_{\Lambda-2} + (\Lambda - 2)v_{\Lambda-2} \\ &= \Lambda R(E_-)v_\Lambda + (\Lambda - 2)v_{\Lambda-2} \\ &= (2\Lambda - 2)V_{\Lambda-2}. \end{aligned}$$

Proceeding by induction, we find that we can always use the relations for lower n to eliminate the $R(E_+)$s at any n we like and write the final result in terms of the next state up. That is,

$$R(E_+)v_{\Lambda-2n} = r_n v_{\Lambda-2n+2}.$$

Plugging this into our general equation for $R(E_+)v_{\Lambda-2n}$, we get a recurrence relation[17]:

$$r_n = r_{n-1} + \Lambda - 2n + 2$$

with the single boundary condtion that $R(E_+)v_\Lambda = 0$. This implies that $r_0 = 0$, so we use this to find that our recurrence relation takes the form

$$r_n = (\Lambda + 1 - n)n.$$

In addition, a finite-dimensional representation must also have a lowest weight $\Lambda - 2N$ (recall N is the dimension of the representation). That is, we have some lowest weight vector $v_{\Lambda-2N} \neq 0$ such that

$$R(E_-)v_{\Lambda-2N} = 0 \implies v_{\Lambda-2N-2} = 0 \implies r_{N+1} = 0.$$

But that vanishing means that

$$(\Lambda - N)(N+1) = 0 \implies \Lambda = N \in \mathbb{Z}_{\geq 0}.$$

This completes the characterization of the representation theory of $L(SU(2))$. We conclude that a finite dimensional irrep R_Λ of $L(SU(2))$ can be described totally by a highest weight $\Lambda \in \mathbb{Z}_{\geq 0}$ and it comes with a remaining set of weights

$$S_{R_\Lambda} = \{-\Lambda, -\Lambda + 2, \ldots \Lambda - 2, \Lambda\} \subset \mathbb{Z},$$

where

$$\dim(R_\Lambda) = \Lambda + 1.$$

Example 10.6. Let's take some explicit cases. R_0 has dimension 1 (d_0, the trivial representation), R_1 has dimension 2 (d_f, the fundamental representation), and R_2 has dimension 3 (d_{Adj}, the adjoint representation).

This is precisely equivalent to the theory of angular momentum in quantum mechanics but with a different normalization– in QM, our spin states had single integer steps but with $j_{max} = n/2, n \in \mathbb{N}$. This happens because the angular momentum operators obey the same bracket structure (i.e. fail to commute) in exactly the same way as the basis elements of the Lie algebra $L(SU(2))$.

Lecture 11.

Direct Sums and Tensor Products: Saturday, October 27, 2018

Some remarks from last time. In terms of the original basis vectors, our basis for $SU(2)$ was

$$H = 2iT^3, \quad E_\pm = i(T^1 \pm iT^2),$$

with $T^a \equiv -\frac{1}{2}i\sigma_a, a = 1,2,3$. Conversely, we can invert these relationships to find $T^3 = H/2i, T^1 = \frac{1}{2i}(E_+ + E_-), T^2 = -\frac{1}{2}(E_+ - E_-)$.

It follows that a representation R of the complexified Lie algebra $L_\mathbb{C}(SU(2))$ (i.e. a set of linear maps $R(H), R(E_+), R(E_-)$) induces a representation of the original Lie algebra $L(SU(2))$, which we get by applying our representation to the original basis elements, e.g.

$$R(T^1) = \frac{1}{2i}R(E_+ + E_-) = \frac{1}{2i}(R(E_+) + R(E_-)).$$

Today we'll consider the $SU(2)$ representation from $L(SU(2))$ representations. That is, we'll look at the connection between the representation of a Lie algebra $L(G)$ and the representation of the original Lie group G.

The punchline from last time was that finite-dimensional irreps of $L(SU(2))$ can be labeled by the highest weight $\Lambda \in \mathbb{Z}_{\geq 0}$, with a weight set

$$S_\Lambda = \{-\Lambda, \Lambda + 2, \ldots, \Lambda - 2, +\Lambda\} \subset \mathbb{Z}.$$

We had $\dim(R_\Lambda) = \Lambda + 1$.

To a physicist, this is simply a complicated way of expressing angular momentum in quantum mechanics. Recall that the total angular momentum is orbital + spin angular momentum. We had our $\mathbf{J}$ operator,

$$\mathbf{J} = (J_1, J_2, J_3)$$

[17] Clearly, the left side of our original recurrence relation just becomes $r_n v_{\Lambda-2n+2}$. On the right side, we've left out a few steps. $R(E_-)R(E_+)v_{\Lambda-2n+2} = R(E_-)R(E_+)v_{\Lambda-2(n-1)} = R(E_-)r_{n-1}v_{\Lambda-2n+4} = r_{n-1}v_{\Lambda-2n+2}$. Pull out the $v_{\Lambda-2n+2}$s everywhere and you're left with the recurrence relation.

with eigenstates (e.g. of J_3) labeled by $j \in \mathbb{Z}/2, j \geq 0$. We then had

$$m \in \{-j, j+1, \ldots, +j\}$$

such that

$$\hat{J}_3 |j, m\rangle = m |j, m\rangle$$

and the total angular momentum J^2 with

$$\hat{J}^2 |j, m\rangle = j(j+1) |j, m\rangle .$$

We can set up the correspondence

$$J_3 = \frac{1}{2} R(H)$$

and

$$J_\pm = J_1 \pm i J_2 = R(E_\pm)$$

so that the highest weight Λ of the representation corresponds to

$$\Lambda = 2j \in \mathbb{Z}$$

and a general weight $\lambda \in S(R)$ corresponds to the angular momentum along a particular axis,

$$\lambda = 2m \in \mathbb{Z}.$$

The eigenvector v_Λ thus corresponds to $v_\Lambda \sim |j, j\rangle$ and similarly $v_\lambda \sim |j, m\rangle$. This explains in an algebraic context why m ranges from $-j$ to j in integer steps, with j a positive half-integer. Fixing Λ is equivalent to choosing the total angular momentum, and fixing λ is then choosing the angular momentum along a particular axis (e.g. J_3).

Recall that locally we can parametrize group elements $A \in SU(2)$ using the exponential map,

$$A = \exp(X), X \in L(SU(2)).$$

Starting from the irreducible representations R_Λ of $L(SU(2))$ defined above, we can then define the representation

$$D_\Lambda(A) \equiv \exp(R_\Lambda(X)), \Lambda \in \mathbb{Z}_{\geq 0}.$$

Recall that $SU(2)$ and $SO(3)$ have the same Lie algebra. In general this will yield a valid representation of $SU(2)$ but *not* of $SO(3) \simeq SU(2)/\mathbb{Z}_2$. For this to be a representation of $SO(3)$, we must further require that it is well-defined on the quotient by the center of the group $\{I_2, -I_2\}$, i.e.

$$D_\Lambda(-I_2) = D_\Lambda(I_2) \iff D_\Lambda(-A) = D_\Lambda(A) \forall A \in SU(2).$$

Let's check explicitly if that holds. First note that we can write

$$-I_2 = \exp(i\pi H)$$

(from the explicit form of H– check this). Now we'll pass to the representation D_Λ:

$$D_\Lambda(-I_2) = D_\Lambda(\exp(i\pi H)) = \exp(i\pi R_\Lambda(H)).$$

But $R_\Lambda(H)$ has eigenvalues $\lambda \in \{-\Lambda, \Lambda+2, \ldots, +\Lambda\}$, so the matrix on the left $D_\Lambda(-I_2)$ must have the same eigenvalues (after exponentiation)

$$\exp(i\pi\lambda) = \exp(i\pi\Lambda) = (-1)^\Lambda$$

since λ goes in steps of 2. Therefore we find that

$$D_\Lambda(-I_2) = D_\Lambda(I_2) = (I)_{(\Lambda+1)\times(\Lambda+1)} \iff \Lambda \in 2\mathbb{Z}.$$

That is, Λ must be even. In this case, $\Lambda \in 2\mathbb{Z} \implies D_\Lambda$ is a representation of $SU(2)$ and $SO(3)$, whereas $\Lambda \in 2\mathbb{Z}+1 \implies D_\Lambda$ is a representation of $SU(2)$ but *not* of $SO(3)$. Sometimes we call this a "spinor representation" (i.e. half-integer spin) of $SO(3)$, but these aren't really representations of $SO(3)$– really, they're representations of the double cover $SU(2)$.

This reveals something a bit interesting– the true rotation group of the physical world we live in is not $SO(3)$ but $SU(2)$. The particles which see these complex rotations are exactly the particles with half-integer spin.

New representations from old

Definition 11.1. If R is a representation of a real Lie algebra $\mathfrak{g}$, we define a *conjugate representation* by

$$\bar{R}(X) = R(X)^* \forall X \in \mathfrak{g}.$$

It's an exercise to check that this really is a representation– see example sheet 2. Note that sometimes $\bar{R} \simeq R$, so the new representation is isomorphic to the old one.

Definition 11.2. Given representations R_1, R_2 of a Lie algebra with corresponding representation spaces V_1, V_2 and dimensions d_1, d_2, we may define the *direct sum* of the representations, denoted

$$R_1 \oplus R_2.$$

The direct sum acts on the direct sum of the vector spaces,

$$V_1 \oplus V_2 = \{v_1 \oplus v_2 : v_1 \in V_1, v_2 \in V_2\}.$$

The dimension of the new representation space is simply $\dim(V_1 \oplus V_2) = d_1 + d_2$. The direct sum is then defined very simply by

$$(R_1 \oplus R_2)(X) \cdot (v_1 \oplus v_2) = (R_1(X)v_1) \oplus (R_2(X)v_2) \in V_1, V_2.$$

There's probably a nice commuting diagram for this in category theory. To make this more concrete, one can write $(R_1 \oplus R_2)(X)$ as a block diagonal matrix with $R_1(X)$ in the upper left, $R_2(X)$ in the lower right.

$$R(X) = \left(\begin{array}{c|c} R_1(X) & \\ \hline & R_2(X) \end{array}\right)$$

This is known as a *reducible* representation, i.e. a representation which can be written as the direct sum of two (or more) representations.

Definition 11.3. Given vector spaces V_1, V_2 with dimensions d_1, d_2, we define the *tensor product space* $V_1 \otimes V_2$. It's got a different structure than the more familiar Cartesian product– our tensor product space is spanned by basis elements

$$v_1 \otimes v_2, v_1 \in V_1, v_2 \in V_2.$$

In particular, the tensor product space has dimension $\dim(V_1 \otimes V_2) = d_1 \times d_2$. This is already different from the Cartesian (direct) product, which has dimension $\dim(V_1 \times V_2) = d_1 + d_2$.

Moreover, the addition structure on the tensor product space is special. In a Cartesian product, it makes sense to add terms like $(0,1) + (1,0) = (1,1)$ (i.e. term-wise addition). But a tensor product is a formal product. In a tensor product, $|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle$ cannot be simplified any further. It only makes sense to add terms which have one of the elements from the original spaces in common, e.g. something like $(0,1) + (1,1) = (1,1)$.

Tensor products are of particular interest in physics because when we consider the Hilbert space of a multi-particle state, it can be represented not as a direct product but a tensor product of the individual single-particle states.[18]

Given two linear maps $M_1 : V_1 \to V_1, M_2 : V_2 \to V_2$, we can define the *tensor product map* $(M_1 \otimes M_2) : V_1 \otimes V_2 \to V_1 \otimes V_2$ such that

$$(M_1 \otimes M_2)(v_1 \otimes v_2) = (M_1 v_1) \otimes (M_2 v_2) \in V_1 \otimes V_2,$$

which may be extended naturally to all elements of $V_1 \otimes V_2$ by linearity (since we have defined it on all basis vectors).

[18] For a simple example, consider two particle spin states taking discrete values. $|a\rangle, |b\rangle \in \{|0\rangle, |1\rangle\}$. Then the two-particle states are described by the tensor product space $|a\rangle \otimes |b\rangle$ (sometimes $|a\rangle |b\rangle$ or simply $|ab\rangle$), which is spanned by $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. It's obvious in this notation that it doesn't make sense to add states like $|10\rangle + |01\rangle$– addition is only well-defined when at least one of the original basis states matches, e.g. $|10\rangle + |11\rangle = |1\rangle (|0\rangle + |1\rangle)$. It's also clear that the tensor product space is "bigger" than the direct product space. To make contact with quantum mechanics, it's the tensor product structure which lets us prepare entangled states like $|00\rangle + |11\rangle$ which have no natural projection onto the original one-particle states.

Definition 11.4. Suppose we have two representations R_1, R_2 of a Lie algebra $\mathfrak{g}$ acting on representation spaces V_1, V_2. By definition, for $X \in \mathfrak{g}$ we have

$$R_1(X) : V_1 \to V_1, R_2(X) : V_2 \to V_2.$$

Then we can define a new representation, the *tensor product representation* $(R_1 \otimes R_2)$ such that for each $X \in \mathfrak{g}$, we get

$$(R_1 \otimes R_2)(X) : V_1 \otimes V_2 \to V_1 \otimes V_2,$$

given explicitly by

$$(R_1 \otimes R_2)(X) \equiv R_1(X) \otimes I_{V_2} + I_{V_1} \otimes R_2(X).$$

Here, I've denoted I_{V_1} as the identity on V_1 and the same is true for I_{V_2}. We'll talk more about what this looks like and why it's defined this way next time.

Lecture 12.

Reducibility and Remainders: Tuesday, October 30, 2018

Last time, we defined the tensor product of two representations. Suppose we have a Lie algebra $\mathfrak{g}$ and two representations R_1, R_2 with representation spaces V_1, V_2 respectively and dimensions $\dim(R_1) = d_1, \dim(R_2) = d_2$. Then the tensor product of these two representations acts on the representation space $V_1 \otimes V_2$ and is defined such that $\forall X \in \mathfrak{g}$,

$$(R_1 \otimes R_2)(X) = R_1(X) \otimes I_2 + I_1 \otimes R_2(X).$$

Here, I_1, I_2 are the identity maps on V_1 and V_2. Note also that

$$(R_1 \otimes R_2)(X) \neq R_1(X) \otimes R_2(X),$$

since this would be quadratic rather than linear in X and would therefore fail to be a representation.

To make this more concrete, let us choose bases

$$\begin{aligned} B_1 &= \{v_1^j; j = 1, \ldots, d_1\} \\ B_2 &= \{v_2^\alpha; \alpha = 1, \ldots, d_2\}. \end{aligned}$$

Thus a basis for $V_1 \otimes V_2$ is

$$B_{1\otimes 2} = \{v_1^j \otimes v_2^\alpha; j = 1, \ldots, d_1, \alpha = 1, \ldots, d_2\}.$$

The dimension of the new representation is therefore $\dim(R_1 \otimes R_2) = d_1 d_2$ (i.e. it is spanned by $d_1 d_2$ tensor products of the d_1 basis vectors of V_1 and the d_2 basis vectors of V_2).

The tensor product representation $R_1 \otimes R_2$ is then given in a basis by

$$(R_1 \otimes R_2)(X)_{a\alpha, j\beta} = R_1(X)_{ij} \underbrace{I_{\alpha\beta}}_{d_2 \times d_2} + \underbrace{I_{ij}}_{d_1 \times d_1} R_2(X)_{\alpha\beta}$$

where the identity matrices have the dimensions indicated.

Definition 12.1. We say that a representation R with representation space V has an *invariant subspace* $U \subset V$ if

$$R(X)u \in U \forall X \in \mathfrak{g}, u \in U.$$

Every representation space has two trivially invariant subspaces, $U = \{0\}$ and $U = \{V\}$. We then say that if V has no non-trivial invariant subspaces, we call the corresponding representation an *irreducible representation* or *irrep* of $\mathfrak{g}$.

If R has an invariant subspace U, we may find a basis such that for all $X \in \mathfrak{g}$, the representation matrices take the block matrix form

$$R(X) = \left(\begin{array}{c|c} A(X) & B(X) \\ \hline 0 & C(X) \end{array}\right)$$

where the elements of U now take the form

$$\left(\begin{array}{c} U \\ \hline 0 \end{array}\right).$$

Definition 12.2. Moreover, a *fully reducible representation* can be written as a direct sum of irreps, i.e. in some basis, R takes a block diagonal form

$$R(X) = \left(\begin{array}{c|c|c|c} R_1(X) & & & \\ \hline & R_2(X) & & \\ \hline & & \ddots & \\ \hline & & & R_n(X) \end{array}\right)$$

It's an important fact that if $R_i, i = 1, \ldots, m$ are finite-dimensional irreps of a simple Lie algebra then

$$R_1 \otimes R_2 \otimes \ldots \otimes R_m$$

is fully reducible as some direct sum

$$R_1 \otimes R_2 \otimes \ldots \otimes R_m \cong \tilde{R}_1 \oplus \tilde{R}_2 \oplus \ldots \oplus \tilde{R}_{\tilde{m}}.$$

Practically speaking, let's consider tensor products of $L(SU(2))$ representations. Let $R_\Lambda, R_{\Lambda'}$ be two irreps of $L(SU(2))$ with highest weights Λ, Λ' and representation spaces $V_\Lambda, V_{\Lambda'}$, where $\Lambda, \Lambda' \in \mathbb{Z}_{\geq 0}$. We defined these last time– these are just the spin states of particles with total spin $\Lambda/2, \Lambda'/2$. They have dimension

$$\dim(R_\Lambda) = \Lambda + 1, \dim(R_{\Lambda'}) = \Lambda' + 1.$$

We can then form the tensor product representation $R_\Lambda \otimes R_{\Lambda'}$ with representation space spanned by the tensor products of basis vectors:

$$V_\Lambda \otimes V_{\Lambda'} = \mathrm{span}_{\mathbb{R}}\{v \otimes v'; v \in V_\Lambda, v' \in V_{\Lambda'}\}.$$

Now $\forall X \in L(SU(2))$ we have

$$(R_\Lambda \otimes R_{\Lambda'})(X) \cdot (V \otimes v') = (R_\Lambda(X)v) \otimes v' + v \otimes (R_{\Lambda'}(X)v').$$

Since $L(SU(2))$ is simple, this gives a fully reducible representation of $L(SU(2))$ of dimension

$$\dim(R_\Lambda \otimes R_{\Lambda'}) = (\Lambda + 1)(\Lambda' + 1)$$

Then we can rewrite the tensor product as a direct product:

$$R_\Lambda \otimes R_{\Lambda'} = \bigoplus_{\Lambda'' \in \mathbb{Z}_{\geq 0}} L_{\Lambda,\Lambda'}^{\Lambda''} R_{\Lambda''}$$

for some non-negative integers $L_{\Lambda,\Lambda'}^{\Lambda''}$, which we call "Littlewood-Richardson coefficients." That is, the various irreps appear in the direct sum with some multiplicity given by these coefficients.

Now recall that V_Λ has a basis $\{v_\lambda\}$ where λ specifies the weights,

$$\lambda \in S_\Lambda = \{-\Lambda, -\Lambda + 2, \ldots, +\Lambda\}$$

where these v_λ are eigenvectors of $R_\Lambda(H)$ such that

$$R_\Lambda(H)v_\lambda = \lambda v_\lambda.$$

Similarly $V_{\Lambda'}$ is equipped with a basis $\{v'_{\lambda'}\}$ where

$$\lambda' \in S_{\Lambda'} = \{-\Lambda', -\Lambda' + 2, \ldots, +\Lambda'\}$$

and

$$R_{\Lambda'}(H)v'_{\lambda'} = \lambda' v'_{\lambda'}.$$

Therefore a basis for the representation space $V_\Lambda \otimes V_{\Lambda'}$ is given by

$$B_{\Lambda \otimes \Lambda'} = \{v_\lambda \otimes v'_{\lambda'}; \lambda \in S_\lambda, \lambda' \in S_{\Lambda'}\}.$$

Acting on a particular basis vector, we find that

$$\begin{aligned} (R_\Lambda \otimes R_{\Lambda'})(H)(v_\lambda \otimes v'_{\lambda'} &= (R_\Lambda(H)v_\lambda) \otimes v'_{\lambda'} + v_\lambda \otimes (R_{\Lambda'}(H)v'_{\lambda'}) \\ &= (\lambda + \lambda')(v_\lambda \otimes v'_{\lambda'}). \end{aligned}$$

What we find is that the possible weights are therefore just sums of the individual λ, λ'. That is, the weight set of $R_\Lambda \otimes R_{\Lambda'}$ is simply

$$S_{\Lambda,\Lambda'} = \{\lambda + \lambda' : \lambda \in S_\Lambda, \lambda' \in S_{\Lambda'}\}.$$

Note that elements of this set can have degeneracy– the same number can appear more than once![19] However, it's also true that if we look for the highest weight of the new tensor product representation, it is exactly $\Lambda + \Lambda'$, appearing with multiplicity one:

$$L_{\Lambda,\Lambda'}^{\Lambda+\Lambda'} = 1.$$

Thus we may write

$$R_\Lambda \otimes R_{\Lambda'} = R_{\Lambda+\Lambda'} \oplus \tilde{R}_{\Lambda,\Lambda'}$$

where we have written the tensor product in terms of a new irrep $R_{\Lambda+\Lambda'}$ and also a *remainder* $\tilde{R}_{\Lambda,\Lambda'}$. This has a new weight set $\tilde{S}_{\Lambda,\Lambda'}$ such that

$$S_{\Lambda,\Lambda'} = S_{\Lambda+\Lambda'} \cup \tilde{S}_{\Lambda,\Lambda'}.$$

Equivalently $\tilde{S}_{\Lambda,\Lambda'} = S_{\Lambda,\Lambda'} \setminus S_{\Lambda+\Lambda'}$.

What does this look like? Consider the case $\Lambda = \Lambda' = 1$. Then

$$S_1 = \{-1, +1\},$$

so the weight set of the tensor product is

$$\begin{aligned} S_{1\otimes 1} &= \{(-1) + (-1), (-1) + 1, 1 + (-1), 1 + 1\} \\ &= \{-2, 0, 0, +2\} \\ &= \{-2, 0, 2\} \cup \{0\}. \end{aligned}$$

It follows that

$$R_1 \otimes R_1 = R_2 \oplus R_0,$$

which is the sophisticated version of the fact from undergrad quantum mechanics that a system of two spin 1/2 particles can behave like a spin 1 particle or a spin 0 particle:

$$\text{spin } 1/2 \otimes \text{spin } 1/2 = \text{spin } 1 \oplus \text{spin } 0.$$

Lecture 13.

Thursday, November 1, 2018

Last time, we finished discussing the representation theory of $L(SU(2))$. In particular, we defined the tensor product representation and showed that we can usually express the tensor product of two representations in terms of the direct product of many copies of R_Λ:

$$R_\Lambda \otimes R_{\Lambda'} = \bigoplus_{\Lambda'' \in \mathbb{Z}_{\geq 0}} L_{\Lambda,\Lambda'}^{\Lambda''} R_{\Lambda''}$$

with

$$L_{\Lambda,\Lambda'}^{\Lambda''} \in \mathbb{Z}_{\geq 0}.$$

We also described an algorithm to work out the direct product representation, namely writing the tensor product as a direct sum of the representation $R_{\Lambda+\Lambda'}$ and some remainder term $\tilde{R}_{\lambda,\Lambda'}$. It's an exercise (sheet 2, Q3) to work out that

$$R_N \otimes R_M = R_{|N-M|} \oplus R_{|N-M|+2} \oplus \ldots \oplus R_{N+M}.$$

Tensor products are important because multi-particle spaces are described in general by tensor products, not direct products (this leads to the phenomenon of entanglement).

Let us now define something called the *Killing form.*

[19] Consider $\lambda = 2, \lambda' = 0$ and $\lambda = 0, \lambda' = 2$. Both of these will appear as terms in the set so the weight 2 can appear twice. We'll see this concretely in a minute.

Definition 13.1. Given a vector space V over F $(= \mathbb{R}, \mathbb{C})$ an *inner product* i is a symmetric bilinear map

$$i : V \times V \to F.$$

In particular, i is *non-degenerate* if for every $v \in V$ $(v \neq 0)$, there is a $w \in V$ such that

$$i(v, w) \neq 0.$$

That is, there is no vector that is orthogonal to all the others under the inner product, or equivalently it has no zero eigenvalues considered as a linear map.

Question: is there a "natural" inner product on a Lie algebra $\mathfrak{g}$? The answer is yes– it is called the *Killing form*, an inner product κ with

$$\kappa : \mathfrak{g} \times \mathfrak{g} \to F.$$

We'll define the formula first and then explore why it makes sense.

Definition 13.2. The Killing form K is defined such that $\forall X, Y \in \mathfrak{g}$,

$$\kappa(X, Y) \equiv \mathrm{Tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y).$$

That is, K is the trace of the linear map

$$\mathrm{ad}_X \circ \mathrm{ad}_Y : \mathfrak{g} \to \mathfrak{g}$$

which takes

$$Z \in \mathfrak{g} \mapsto [X, [Y, Z]] \in \mathfrak{g}.$$

Why is this a sensible choice? Suppose we choose a basis $\{T^a\}, a = 1, \dots D$ for $\mathfrak{g}$ with dimension D. Then

$$X = X_a T^a, \quad Y = Y_a T^a, \quad Z = Z_a T^a.$$

We also have some structure constants associated to the basis,

$$[T^a, T^b] = f^{ab}_c T^c.$$

Thus the composition of the ad maps is some $D \times D$ matrix, and we can work out in this basis the components of this matrix.

$$\begin{aligned}
[X, [Y, Z]] &= X_a Y_b Z_c [T^a, [T^b, T^c]] \\
&= X_a Y_b Z_c [T^a, f^{bc}_d T^d] \\
&= X_a Y_b Z_c f^{ad}_e f^{bc}_d T^e \\
&= M(X, Y)^c_e Z_c T^e
\end{aligned}$$

with

$$M(X, Y)^c_e \equiv X_a Y_b f^{ad}_e f^{bc}_d.$$

The matrix $M(X, Y)$ is therefore the linear map $\mathrm{ad}_X \circ \mathrm{ad}_Y : \mathfrak{g} \to \mathfrak{g}$, and all that remains is to take the trace to get the Killing form.

$$\begin{aligned}
\kappa^{ab} X_a Y_b &= \mathrm{Tr}_D[M(X, Y)] \\
&= M(X, Y)^c_c \\
&= X_a Y_b f^{ad}_c f^{bc}_d.
\end{aligned}$$

Therefore the Killing form in terms of structure constants is explicitly

$$\kappa^{ab} = f^{ad}_c f^{bc}_d.$$

The indices c and d are summed over, so we get the two free indices a, b as desired.

Now what do we mean by saying that the Killing form is a "natural" inner product on a Lie algebra? It is the property that κ is invariant under the adjoint action of $\mathfrak{g}$,

$$\kappa([Z, X], Y) + \kappa(X, [Z, Y]) = 0$$

for all $Z \in \mathfrak{g}, X, Y \in \mathfrak{g}$. This is the equivalent of invariance under a conjugation by a Lie algebra element, gXg^{-1}.

Let's show that this property holds for this inner product.

$$\begin{aligned}\kappa([Z,X],Y) &= \mathrm{Tr}[\mathrm{ad}_{[Z,X]} \circ \mathrm{ad}_Y] \\ &= \mathrm{Tr}[(\mathrm{ad}_Z \circ \mathrm{ad}_X - \mathrm{ad}_X \circ \mathrm{ad}_Z) \circ \mathrm{ad}_Y] \\ &= \mathrm{Tr}[\mathrm{ad}_Z \circ \mathrm{ad}_X \circ \mathrm{ad}_Y] - \mathrm{Tr}[\mathrm{ad}_X \circ \mathrm{ad}_Z \circ \mathrm{ad}_Y]\end{aligned}$$

where in going from the first to the second line, we have used the fact that the ad map is also a representation and can therefore be rewritten by linearity in its argument $[Z, X]$. Similarly,

$$\kappa(X,[Z,Y]) = \mathrm{Tr}[\mathrm{ad}_X \circ \mathrm{ad}_Z \circ \mathrm{ad}_Y] - \mathrm{Tr}[\mathrm{ad}_X \circ \mathrm{ad}_Y \circ \mathrm{ad}_Z].$$

However, if we now compare these two expressions we see that by the cyclic property of the trace (i.e. interpreting the ad maps as matrices on the vector space), their sum vanishes[20], and so

$$\kappa([Z,X],Y) + \kappa(X,[Z,Y]) = 0. \quad \boxtimes$$

We may next ask under what conditions κ is non-degenerate, i.e. the map κ^{ab} is invertible.

Theorem 13.3. *(Cartan) The Killing form κ on a Lie algebra $\mathfrak{g}$ is non-degenerate $\iff$ $\mathfrak{g}$ is semi-simple.*

In the specific case, if $\mathfrak{g}$ is simple, then the Killing form κ is the unique invariant inner product on $\mathfrak{g}$ up to an overall scalar multiple.

Definition 13.4. A Lie algebra is *semi-simple* if it has no abelian ideals. (This is a little weaker than simple, clearly.)

Exercise 13.5. From Example Sheet 2, Question 9b: Show that a finite dimensional semi-simple Lie algebra can be written as the direct sum of a finite number of simple Lie algebras,

$$\mathfrak{g} = g_1 \oplus g_2 \oplus \ldots \oplus g_l, \quad g_i \text{ simple.}$$

Note that a direct product $\mathfrak{g} \oplus \mathfrak{f}$ of Lie algebras $\mathfrak{g}, \mathfrak{f}$ is defined such that $\forall X \in \mathfrak{g}, Y \in \mathfrak{f}, [X,Y] = 0$.

Let us prove the forward direction of Cartan's theorem. First note that

$$\kappa \text{ non-degenerate} \implies \mathfrak{g} \text{ is semi-simple}$$

is equivalent to proving the contrapositive,

$$\mathfrak{g} \text{ not semi-simple} \implies \kappa \text{ is degenerate.}$$

Suppose $\mathfrak{g}$ is not semi-simple. Then $\mathfrak{g}$ has an abelian ideal $\mathfrak{j}$. Let $\dim(\mathfrak{g}) = D$ and suppose the ideal has dimension $\dim(\mathfrak{j}) = d$. WLOG we can choose a basis B for $\mathfrak{g}$ such that

$$B = \{T^a\} = \{\underbrace{T^i; i = 1, \ldots, d}_{\text{span } \mathfrak{j}}\} \cup \{T^\alpha; \alpha = d+1, d+2, \ldots, D\},$$

i.e. a subset T^i of the basis vectors span the ideal $\mathfrak{j}$. Since $\mathfrak{j}$ is abelian,

$$[T^i, T^j] = 0 \quad \forall i,j = 1, \ldots, d.$$

Therefore the structure constants are constrained by

$$f^{ij}_a = 0, \quad i,j = 1, \ldots, d, a = 1, \ldots, D.$$

That is, the bracket vanishes for all pairs of elements in the abelian ideal, so all structure constants with ij indices up are zero.

Moreover, $\mathfrak{j}$ is an ideal, so the bracket of a basis element for $\mathfrak{j}$ with a general basis element is still in $\mathfrak{j}$. That is,

$$[T^\alpha, T^j] = f^{\alpha j}_k T^k \in j \implies f^{\alpha j}_\beta = 0, \beta = d+1, \ldots, D.$$

We'll use these facts next time to complete the proof of Cartan's theorem in one direction.

[20] Explicitly, this means that $\mathrm{Tr}[\mathrm{ad}_X \circ \mathrm{ad}_Y \circ \mathrm{ad}_Z] = \mathrm{Tr}[\mathrm{ad}_Z \circ \mathrm{ad}_X \circ \mathrm{ad}_Y]$.

Lecture 14.

Saturday, November 3, 2018

Today we'll complete our initial discussion of Killing forms and begin the Cartan classification of finite-dimensional simple complex Lie algebras.

Last time, we showed that if $\mathfrak{g}$ is not semi-simple, then some of the structure constants must vanish. Let's see what how this implies that κ is degenerate (i.e. there exists some $v \in V$ such that its Killing form vanishes, $i(v,w) = 0$ for all $w \in V$).

Theorem 14.1 (Cartan's theorem)**.** *For a Lie algebra $\mathfrak{g}$, if its Killing form κ is nondegenerate, then $\mathfrak{g}$ is semi-simple.*

Proof. We started by separating the basis vectors T^a into a set $\{T^i, i = 1,\ldots,d\}$ spanning the ideal $\mathfrak{j}$ and the rest of the basis vectors $\{T^\alpha, \alpha = d+1,\ldots,D\}$. Since $\mathfrak{j}$ is abelian, we found that

$$[T^i, T^j] = 0 \quad \forall i,j = 1,\ldots,d \implies f^{ij}_a = 0, \tag{14.2}$$

and since $\mathfrak{j}$ is an ideal

$$[T^\alpha, T^j] = f^{\alpha j}_k T^k \in \mathfrak{j} \implies f^{\alpha j}_\beta = 0. \tag{14.3}$$

Now consider a general element of the Lie algebra,

$$X = X_a T^a \in \mathfrak{g}$$

and a general element of the ideal,

$$Y = Y_i T^i \in \mathfrak{j}.$$

Then

$$\kappa(X,Y) = \kappa^{ai} X_a Y_i, \text{ with } \kappa^{ai} \equiv f^{ad}_c f^{ic}_d.$$

Let's take this carefully.

$$\begin{aligned} \kappa^{ai} &= f^{ae}_c f^{ic}_e \text{ by definition} \\ &= f^{ae}_\alpha f^{i\alpha}_e \text{ by 14.2} \\ &= f^{aj}_\alpha f^{i\alpha}_j \text{ by 14.3.} \end{aligned}$$

To go from the first line to the second, we have used the fact that if $c = 1,\ldots,d$, then f^{ic}_e vanishes, so $f^{ic}_e = f^{i\alpha}_e$. To go from the second line to the third, we have then used the fact that if $e = d+1,\ldots,D$ then f^{ae}_α vanishes, so $f^{ae}_\alpha = f^{aj}_\alpha$. Now separate the sum over $a = 1,\ldots,D$ into $k = 1,\ldots,d$ and $\beta = d+1,\ldots,D$. Thus

$$\kappa^{ai} = \underbrace{f^{\beta j}_\alpha}_{\text{zero by } f^{\alpha j}_\beta = 0} f^{i\alpha}_j + \underbrace{f^{kj}_\alpha}_{\text{zero by} f^{kj}_a = 0} f^{i\alpha}_k = 0.$$

Therefore

$$\kappa[X,Y] = 0 \quad \forall Y \in \mathfrak{j}, \forall X \in \mathfrak{g} \implies \kappa \text{ is degenerate.}$$

Taking the contrapositive, we conclude that κ is nondegenerate $\implies$ $\mathfrak{g}$ is semi-simple. ⊠

In Hugh Osborn's notes, he proves the other direction, so this turns out to be an if and only if. That is, κ is nondegenerate $\iff$ $\mathfrak{g}$ is semi-simple.

Cartan classification Cartan proved in 1894 that one can fully classify all finite dimensional, simple, complex Lie algebras. Happily, these are often the ones which are of most use to us in physics. Simple Lie groups come with non-degenerate inner products (a fortiori, since simple implies semi-simple), which is a nice property. Moreover we will often look at complex Lie algebras since the field $\mathbb{C}$ is *algebraically closed*–polynomials with complex coefficients have in general complex solutions, whereas the same is not true for polynomials with real coefficients (which can have complex solutions).

Recall that when we did the representation theory of $L(SU(2))_\mathbb{C}$, we defined a Cartan-Weyl basis,

$$\{H, E_\pm\}$$

where H is diagonal and $E_\pm$ moves us between eigenvectors. The brackets turned out to be

$$[H,H] = 0, \quad [H, E_\pm] = \pm 2E_\pm$$

What this tells us is that the ad map ad_H (defined by $\mathrm{ad}_H(X) = [H, X]$, in case you forgot) is diagonal, and it has eigenvalues $0, \pm 2$. We would now like to generalize this principle.

Definition 14.4. We say that $X \in \mathfrak{g}$ is *ad-diagonalizable* (AD) if

$$\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$$

is diagonalizable.

For a matrix, this meant that we could write the map as a diagonal matrix by a similarity transformation. More generally, a map is diagonal if we can construct a complete basis for the space out of eigenvectors of that map.

Definition 14.5. A *Cartan subalgebra* $\mathfrak{h}$ of $\mathfrak{g}$ is a maximal abelian subalgebra containing only AD elements.

Unpacking this definition, a Cartan subalgebra therefore has the following properties.

i) $H \in \mathfrak{h} \implies H$ is AD.
ii) $H, H' \in h \implies [H, H'] = 0$.
iii) If $X \in \mathfrak{g}$ and $[X, H] = 0 \forall H \in \mathfrak{h}$ then $X \in \mathfrak{h}$ (this is what we mean by maximal).

Definition 14.6. The dimension of the Cartan subalgebra,

$$r \equiv \dim[\mathfrak{h}],$$

is known as the *rank*. It turns out that all possible Cartan subalgebras $\mathfrak{h} \subset \mathfrak{g}$ have the same dimension, so it makes sense to say that r is the rank of $\mathfrak{g}$.

Example 14.7. In $L_{\mathbb{C}}(SU(2))$, we have $H = \sigma_3$ and $E_\pm = \frac{1}{2}(\sigma_1 \pm i\sigma_2)$. We explicitly wrote down the eigenvalues and eigenvectors of the ad map of H, so H is ad-diagonalizable. We may choose $\mathfrak{h} = \mathrm{span}_{\mathbb{C}}\{H\}$. We could have chosen σ_1 or σ_2 as our element of the Cartan subalgebra (cooking up combinations of the other two Pauli matrices so that ad_{σ_1} or ad_{σ_2} is diagonal). However, *we could not have chosen E_+ as the element of our Cartan subalgebra.* This is apparent when we write down E_+ as a matrix:

$$E_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

is clearly not diagonalizable and therefore not ad-diagonalizable.

Example 14.8. Consider $\mathfrak{g} = L_{\mathbb{C}}(SU(n))$, the set of traceless complex $n \times n$ matrices. A natural basis set is the pairs of diagonal elements

$$(H^i)_{\alpha,\beta} = \delta_{\alpha i}\delta_{\beta i} - \delta_{\alpha(i+1)}\delta_{\beta(i+1)}.$$

For instance, H^1 looks like

$$H^1 = \begin{pmatrix} 1 & 0 & & \\ 0 & -1 & & \\ & & 0 & \\ & & & \ddots \end{pmatrix}.$$

We now claim that the diagonal elements $H^i, i = 1, \ldots, n-1$ are generators of the Cartan subalgebra, i.e.

$$\mathfrak{h} = \mathrm{span}_{\mathbb{C}}\{H^i, i = 1, \ldots n-1\}.$$

It's clear that the diagonal elements H^i commute and therefore have vanishing bracket,

$$[H^i, H^j] = 0 \forall i, j = 1, \ldots, r.$$

But passing to the adjoint representation, this means that

$$(\mathrm{ad}_{H^i} \circ \mathrm{ad}_{H^j} - \mathrm{ad}_{H^j} \circ \mathrm{ad}_{H^i}) = 0,$$

so our basis elements H^i naturally define r linear maps

$$\mathrm{ad}_{H^i} : \mathfrak{g} \to \mathfrak{g}$$

which are simultaneously diagonalizable (i.e. we can find a single set of eigenvectors which are compatible with all the linear maps). Therefore $\mathfrak{g}$ is spanned by the simultaneous eigenvectors of ad_{H^i}. What may we conclude from this? Well, the ad map ad_{H^i} has some zero eigenvalues:

$$\mathrm{ad}_{H^i}(H^j) = [H^i, H^j] = 0 \forall i, j = 1, \dots, r$$

so the ad map has r zero eigenvalues.

The map also has non-zero eigenvalues which correspond to some set of eigenvectors

$$\{E^\alpha, \alpha \in \Phi\}$$

with Φ some set of eigenvalues. The ad map acts on these E^α by

$$\mathrm{ad}_{H^i}(E^\alpha) = [H^i, E^\alpha] = \alpha^i E^\alpha$$

where the $\alpha^i \in \mathbb{C}, i = 1, \dots, r$ are not all zero. (If they were all zero, E^α would be in $\mathfrak{h}$ by the maximality condition.) In $L(SU(2))_{\mathbb{C}}$, these eigenvectors were just the elements $E_\pm$.

Definition 14.9. These values α^i define a *root* α of $\mathfrak{g}$. That is, a root α can be thought of as an abstract label on the eigenvectors E^α defining its eigenvalues under the ad map ad_{H^i}. We'll see another way to think of roots shortly, as objects in their own right (namely, linear maps) which act on the elements $H^i \in \mathfrak{h}$. More on this next time.

Lecture 15.

Tuesday, November 6, 2018

Last time, we started discussing the Cartan classification of finite-dimensional simple complex Lie algebras. We defined the Cartan subalgebra, the maximal abelian subalgebra containing only ad-diagonalizable elements. We defined the rank of the Cartan subalgebra of a Lie algebra $\mathfrak{g}$ as

$$\mathrm{Rank}[\mathfrak{g}] = \dim \mathfrak{h} = r.$$

Now the idea is that if $H^i, i = 1, \dots, r$ is a basis for the Cartan subalgebra, then $[H_i, H_j] = 0$ and all the H_i considered as matrices can be simultaneously diagonalized in some basis. We conclude that $\mathfrak{g}$ is spanned by simultaneous eigenvectors of H_i.

Now there also exist eigenvectors $E^\alpha \in \mathfrak{g}$ such that

$$\mathrm{ad}_{H^i}(E^\alpha) = [H^i, E^\alpha] = \alpha^i E^\alpha,$$

with $\alpha^i \in \mathbb{C}, i = 1, \dots, r$. A general element of the Cartan subalgebra $H \in \mathfrak{h}$ can be written

$$H = e_i H^i$$

where e_i represents the components of H. Now we can write

$$[H, E^\alpha] = \alpha(H) E^\alpha$$

where

$$\alpha(H) : \mathfrak{h} \to \mathbb{C}, H = e_i H^i \mapsto e_i \alpha^i$$

is now a (multi)linear function from H (specifically the components e_i of H) to the complex numbers $\mathbb{C}$.

A root (i.e. a set of values $\alpha = \{\alpha^i\}$) therefore defines a linear map $\mathfrak{h} \to \mathbb{C}$. We think of roots as elements of the *dual vector space* $\mathfrak{h}^*$ of the Cartan subalgebra $\mathfrak{h}$. One can further prove that the roots are non-degenerate (see Fuchs and Schweigert, pg. 87). For our purposes, we will simply assume this is true.

Then we have a set of roots Φ consisting of $d - r$ distinct elements of $\mathfrak{h}^*$ (that is, $\dim(\mathfrak{g}) - \dim(\mathfrak{h})$).

Definition 15.1. We define the Cartan-Weyl basis for $\mathfrak{g}$ to be the set

$$B = \{H^i, i = 1, \dots, r\} \cup \{E^\alpha, \alpha \in \Phi\}$$

with $|\Phi| = d - r$.

Recall now that by the Cartan theorem, $\mathfrak{g}$ simple $\implies$ the Killing form is non-degenerate. The Killing form is the natural choice of inner product on the Lie algebra, and it is defined by

$$K(X,Y) = \frac{1}{N}\text{Tr}[\text{ad}_X \circ \text{ad}_Y]$$

(where we have WLOG chosen a normalization constant $N \in \mathbb{R}^*$).

We'll need a few properties of Lie algebras to move forward here. First, the bracket satsifies the Jacobi identity,

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0 \tag{15.2}$$

for all $X, Y, Z \in \mathfrak{g}$. We have the property of the adjoint representation that taking the ad map commutes nicely with taking the bracket,

$$\text{ad}_{[X,Y]} = \text{ad}_X \circ \text{ad}_Y - \text{ad}_Y \circ \text{ad}_X \tag{15.3}$$

for all $X, Y \in \mathfrak{g}$. Finally, we have the invariance of the Killing form,

$$K([Z,X],Y) + K(X,[Z,Y]) = 0 \tag{15.4}$$

for all $X, Y, Z \in \mathfrak{g}$.

We'd like to prove the following two statements:

i) $\forall H \in \mathfrak{h}, \alpha \in \Phi$, we have

$$K(H, E^\alpha) = 0.$$

ii) $\forall \alpha, \beta \in \Phi, \alpha + \beta \neq 0$,

$$K(E^\alpha, E^\beta) = 0.$$

Let's prove this. $\forall H' \in \mathfrak{h}$ and any $H \in \mathfrak{h}$, we can write

$$\begin{aligned}\alpha(H')K(H,E^\alpha) &= K(H,[H',E^\alpha]) \\ &= -K([H',H],E^\alpha) \text{ by 15.4} \\ &= -K(0,E^\alpha) = 0,\end{aligned}$$

where in the first line we have simply moved the $\alpha(H')$ into the inner product. Hence $\alpha \in \phi, \alpha \neq 0 \implies K(H, E^\alpha) = 0$. ⊠

The second statement is proved as follows. $\forall H' \in \mathfrak{h}$, we write

$$(\alpha(H') + \beta(H'))K(E^\alpha, E^\beta) = K([H', E^\alpha], E^\beta) + K(E^\alpha, [H', E^\beta])].$$

But by15.4, this whole expression vanishes. Therefore $\forall \alpha, \beta \in \Phi, \alpha + \beta \neq 0 \forall H' \in \mathfrak{h}$,

$$\implies K(E^\alpha, E^\beta) = 0 \text{ if } \alpha + \beta \neq 0. \quad \boxtimes$$

Let's prove one more lemma.

iii) $\forall H \in \mathfrak{h}, \exists H' \in \mathfrak{h}$ such that $K(H, H') \neq 0$.

The proof is as follows. For some $H \in \mathfrak{h}$, assume that no such H' exists. Then

$$K(H, H') = 0 \forall H' \in \mathfrak{h}.$$

But from i) above, we know that

$$K(H, E^\alpha) = 0 \forall \alpha \in \Phi.$$

Since the matrices $H_i \in \mathfrak{h}$ and $E^\alpha \in \mathfrak{g}$ form a basis for $\mathfrak{g}$, this means that

$$K(H, X) = 0 \forall X \in \mathfrak{g} \implies K \text{ is degenerate},$$

which contradicts our assumption that K was non-degenerate. Therefore $\exists H' \in \mathfrak{h}$ with $K(H, H') \neq 0$. ⊠

Therefore it is not only the case that K is a non-degenerate inner product on $\mathfrak{g}$; in fact, we have proven the stronger result that K is non-degenerate even when restricted to the Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$.

In components, we write

$$K(H, H') = K^{ij} e_i e'_j$$

where $K^{ij} = K(H^i, H^j)$. Thus iii) implies that K is invertible as an $r \times r$ matrix– it has no zero eigenvalues. That is,

$$\exists (K^{-1})_{ij} \text{ such that } (K^{-1})_{ij} K^{jk} = \delta_i^k.$$

Why is this useful? Precisely because K^{-1} now induces a non-degenerate inner product on $\mathfrak{h}^*$, the dual space. Thus with $\alpha, \beta \in \Phi$, we get

$$[H^i, E^\alpha] = \alpha^i E^\alpha \text{ and } [H^i, E^\beta] = \beta^i E^\alpha.$$

Definition 15.5. We therefore define the inner product on elements of the dual space,

$$(\alpha, \beta) = (K^{-1})_{ij} \alpha^i \beta^j.$$

Next lecture, we'll see why this is a natural choice of inner product.

We shall now prove the following statement.

iv) With $\alpha \in \Phi \implies -\alpha \in \Phi$ with $K(E^\alpha, E^{-\alpha}) \neq 0$.

From i) we have $K(E^\alpha, H) = 0 \forall H \in \mathfrak{h}$, and from ii) we have $K(E^\alpha, E^\beta) = 0 \forall \beta \in \Phi, \alpha \neq -\beta$. Suppose $-\alpha \notin \Phi$. Then we would have

$$K(E^\alpha, X) = 0 \quad \forall X \in \mathfrak{g},$$

which would contradict the non-degeneracy of K on $\mathfrak{g}$. Therefore there must be another basis vector $E^{-\alpha}$ such that $K(E^\alpha, E^{-\alpha}) \neq 0$. ⊠

Now we've almost completely characterized the algebra in the Cartan-Weyl basis. We've written

$$[H^i, H^j] = 0 \quad \forall i, j = 1, \ldots, r$$
$$[H^i, E^\alpha] = \alpha^i E^\alpha \quad \forall i = 1, \ldots, r \forall \alpha \in \Phi.$$

But there's one more set of brackets we must compute: the bracket of the step elements with themselves, $[E^\alpha, E^\beta]$. Fortunately, the computation is not too bad. We can do it with the Jacobi identity:

$$\begin{aligned}[H^i, [E^\alpha, E^\beta]] &= -[E^\alpha, [E^\beta, H^i]] - [E^\beta, [H^i, E^\beta]] \\ &= (\alpha^i + \beta^i)[E^\alpha, E^\beta],\end{aligned}$$

where we have freely used the antisymmetry of the brackets to switch the order of E^β, E^α, and also made use of the known commutation relation of $[H^i, E^\alpha]$. Thus for $\alpha + \beta \neq 0$, we conclude that

$$[E^\alpha, E^\beta] = \begin{cases} N_{\alpha,\beta} E^{\alpha+\beta} & \text{if } \alpha + \beta \in \Phi \\ 0 & \text{if } \alpha + \beta \notin \Phi. \end{cases}$$

That is, we've found that when $\alpha + \beta \neq 0$, the element $[E^\alpha, E^\beta]$ is actually proportional to the step element with root $\alpha + \beta$ (if it exists), with some undetermined constants $N_{\alpha,\beta}$.

Lecture 16.

Thursday, November 8, 2018

Today, we will continue our study of the Cartan-Weyl basis. Recall that we introduced the idea of a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$ in the last two lectures. With a little work, we produced a set of basis vectors E^α where $\alpha \in \mathfrak{h}^*$ are the set of roots, and they form a basis for the dual vector space $\mathfrak{h}^*$.

We found that since the Cartan subalgebra is an abelian subalgebra, the commutator of two elements in it vanishes by definition,

$$[H^i, H^j] = 0.$$

The step operators E^α are eigenvectors of the ad map,

$$[H^i, E^\alpha] = \alpha^i E^\alpha.$$

And the bracket of two step operators is either zero or proportional to another step operator,

$$[E^\alpha, E^\beta] = \begin{cases} N_{\alpha,\beta} E^{\alpha+\beta} & \text{if } \alpha + \beta \in \Phi \\ 0 & \text{if } \alpha + \beta \notin \Phi \end{cases}$$

for some (as yet undetermined) constants $N_{\alpha,\beta}$ (and with the caveat that $\alpha + \beta \neq 0$).

Now, our proof relied on the computation

$$[H^i, [E^\alpha, E^\beta]] = (\alpha^i + \beta^i)[E^\alpha, E^\beta],$$

so $[E^\alpha, E^\beta]$ is only an eigenvector with nonzero eigenvalue if $\alpha+\beta \neq 0$. Otherwise, for the case $\alpha+\beta = 0$ let us consider the inner product

$$\kappa([E^\alpha, E^{-\alpha}], H)$$

and claim it can be written as

$$\kappa([E^\alpha, E^{-\alpha}], H) = \alpha(H)\kappa(E^\alpha, E^{-\alpha}). \tag{16.1}$$

The proof is straightforward: using the invariance of the inner product, we can rewrite

$$\begin{aligned}\kappa([E^\alpha, E^{-\alpha}], H) &= \kappa(E^\alpha, [E^{-\alpha}, H]) \\ &= \kappa(E^\alpha, -[e_i H^i, E^{-\alpha}]) \\ &= \kappa(E^\alpha, -e_i(-\alpha^i E^{-\alpha})) \\ &= \alpha(H)\kappa(E^\alpha, E^{-\alpha}),\end{aligned}$$

recalling from last time that if $H = e_i H^i$ in some basis, then $\alpha(H) = e_i\alpha^i$. ⊠

But by iv from the previous lecture, we found that

$$\kappa(E^\alpha, E^{-\alpha}) \neq 0,$$

so let us now define

$$H^\alpha \equiv \frac{[E^\alpha, E^{-\alpha}]}{\kappa(E^\alpha, E^{-\alpha})}.$$

If we write this as an expression for $[E^\alpha, E^{-\alpha}]$ and substitute into Eqn. 16.1, then by the linearity of the inner product we see that

$$\kappa(H^\alpha, H) = \alpha(H) \forall h \in \mathfrak{h}.$$

This gives us a linear equation on the components e_i of H. Note first that we previously computed $[H^i, [E^\alpha, E^\beta]] = (\alpha^i + \beta^i)[E^\alpha, E^\beta]$. If we set $\alpha = -\beta$, we find that the element $[E^\alpha, E^{-\alpha}$ commutes with all the generators H^i of the Cartan subalgebra. By the maximality assumption, this means that $[E^\alpha, E^{-\alpha}] \in \mathfrak{h}$, so it makes good sense to expand H^α (which is nothing more than a rescaled version of $[E^\alpha, E^{-\alpha})$ in a basis for $\mathfrak{h}$.

Now writing H^α and H in a basis for $\mathfrak{h}$,

$$H^\alpha = \rho_i^\alpha H^i, \quad H = e_i H^i \in \mathfrak{h},$$

the equation becomes

$$K^{ij}\rho_i^\alpha e_j = \alpha^j e_j$$

or equivalently

$$K^{ij}\rho_i^\alpha = \alpha^j,$$

so we can solve for the components ρ_i^α of H^α in terms of the roots $'alpha^j$:

$$\rho_i^\alpha = (K^{-1})_{ij}\alpha^j \implies H^\alpha = \rho_i^\alpha H^i = (K^{-1})_{ij}\alpha^j H^i.$$

Therefore we find that

$$[E^\alpha, E^\beta] = \begin{cases} N_{\alpha,\beta}E^{\alpha+\beta} & \text{if } \alpha+\beta \in \Phi \\ \kappa(E^\alpha, E^{-\alpha} H^\alpha & \text{if } \alpha+\beta = 0 \\ 0 & \text{otherwise.}\end{cases}$$

What properties does this $H^\alpha \in \mathfrak{h}$ have? $\forall \alpha, \beta \in \Phi$, we see that

$$\begin{aligned}[H^\alpha, E^\beta] &= (\kappa^{-1})_{ij}\alpha^i [H^j, E^\beta] \\ &= (\kappa^{-1})_{ij}\alpha^i\beta^j E^\beta \\ &= (\alpha, \beta)E^\beta,\end{aligned}$$

where we see that as promised, κ^{-1} has a natural interpretation as an inner product on elements α, β in the dual space. Now for all $\alpha \in \Phi$ we shall define

$$e^\alpha = \sqrt{\frac{2}{(\alpha,\alpha)\kappa(E^\alpha, E^{-\alpha})}}E^\alpha$$

and

$$h^{\alpha} = \frac{2}{(\alpha,\alpha)} H^{\alpha}.$$

Note that we require $(\alpha,\alpha) \neq 0$ for these expressions to be sensible– see e.g. Fuchs and Schweigert pg. 87 for the proof.

Supposing these elements are well-defined, we now get a similar set of brackets in this basis (written in terms of the roots α). That is,

$$\begin{aligned}
[h^{\alpha}, h^{\beta}] &= 0 \\
[h^{\alpha}, e^{\beta}] &= \frac{2(\alpha,\beta)}{(\alpha,\alpha)} e^{\beta} \\
[e^{\alpha}, e^{\beta}] &= \begin{cases} n_{\alpha,\beta} e^{\alpha+\beta} & \alpha+\beta \in \Phi \\ h^{\alpha} & \alpha+\beta = 0 \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Let's look at a specific example. Consider $L_{\mathbb{C}}(SU(2))$ subalgebras. We have

$$\alpha \in \Phi \implies -\alpha \in \Phi.$$

Now for each pair $\pm\alpha \in \Phi$, we get a subalgebra of $L_{\mathbb{C}}(SU(2))$ spanned by the set

$$\{e^{\alpha}, e^{-\alpha}, e^{\alpha}\}.$$

Our brackets therefore tell us that

$$\begin{aligned}
[h^{\alpha}, e^{\pm\alpha}] &= \pm 2 e^{\pm\alpha} \\
[e^{+\alpha}, e^{-\alpha}] &= h^{\alpha},
\end{aligned}$$

so we immediately recover the subalgebra structure we saw before. Let us label these subalgebras by our choice of root α and call the corresponding subalgebra $sl(2)_{\alpha}$.

Then as a consequence, we get what are called root strings.

Definition 16.2. For $\alpha, \beta \in \Phi$, define the *α-string passing through β* as the set of roots of the form $\beta + \rho\alpha, \rho \in \mathbb{Z}$. That is,

$$S_{\alpha,\beta} = \{\beta + \rho\alpha \in \Phi, \rho \in \mathbb{Z}\}.$$

Now there is a corresponding vector subspace of $\mathfrak{g}$ which we can obtain by exponentiating the root string:

$$V_{\alpha,\beta} = \operatorname{span}_{\mathbb{C}}\{e^{\beta+\rho\alpha}; \beta + \rho\alpha \in S_{\alpha,\beta}\}.$$

Now consider the action of $sl(2)_{\alpha}$ on $V_{\alpha,\beta}$. We see that

$$\begin{aligned}
[h^{\alpha}, e^{\beta+\rho\alpha}] &= \frac{2(\alpha, \beta+\rho\alpha)}{(\alpha,\alpha)} e^{\beta+\rho\alpha} \in V_{\alpha,\beta} \\
&= \left(\frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2\rho \right) e^{\beta+\rho\alpha}.
\end{aligned}$$

By a similar computation, we find that

$$[e^{\pm\alpha}, e^{\beta+\rho\alpha}] \propto e^{\beta+(\rho\pm 1)\alpha} \text{ if } \beta + (\rho \pm 1)\alpha \in \Phi,$$

and it is zero otherwise. Therefore $V_{\alpha,\beta}$ is a representation space for a representation R of $sl(2)_{\alpha}$. In particular,

$$R(h^{\alpha}) = \operatorname{ad}_{h^{\alpha}} \text{ and } R(e^{\pm\alpha}) = \operatorname{ad}_{e^{\pm\alpha}}.$$

We see that R has a weight set given by

$$S_R = \{\frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2\rho; \beta + \rho\alpha \in \Phi\}.$$

Now the representation R has some direct sum representation:

$$R = R_{\Lambda_1} \oplus \ldots \oplus R_{\Lambda_L}, \Lambda_l \in \mathbb{Z}_{\geq 0}.$$

The total weight set is of course the union of all the individual weight sets of the elements of the direct product:

$$S_R = S_{\Lambda_1} \cup \ldots \cup S_{\Lambda_L}.$$

It's also true that $\forall \Lambda \in \mathbb{Z}_{\geq 0}$, we have a weight set which can be written

$$S_\Lambda = \{-\Lambda, -\Lambda + 2, \ldots, +\Lambda\}.$$

But recall that our set of roots is non-degenerate– each $\alpha \in \Phi$ appears once and only once. So the non-degeneracy of the roots of $\mathfrak{g}$ means that the weights of our representation R are also non-degenerate. Therefore

$$S_R = S_\Lambda = \{-\Lambda, -\Lambda + 2, \ldots, +\Lambda\}.$$

Lecture 17.

Saturday, November 10, 2018

In general, if we consider a real Lie algebra $\mathfrak{g}_\mathbb{R}$, such that $\dim(\mathfrak{g}_\mathbb{R}) = D$, its Killing form κ^{ab} will also in general be real. Treating κ as a $D \times D$ matrix, one can then ask about the *signature* of the matrix, i.e. how many positive and negative eigenvalues κ has when interpreted as a matrix, and more generally whether they are all of the same sign.

Definition 17.1. A real Lie algebra $\mathfrak{g}_\mathbb{R}$ is of *compact type* if $\exists$ a basis in which

$$\kappa^{ab} = -K\delta^{ab}, K \in \mathbb{R}^+.$$

If we have a basis for $\mathfrak{g}_\mathbb{R}$ with $\mathfrak{g}_\mathbb{R}$ the real span of the generators $\{T^a, a = 1, \ldots, D\}$ then we can take the *complexification* of $\mathfrak{g}_\mathbb{R}$, which is defined to be the complex span of the same basis vectors,

$$\mathfrak{g}_\mathbb{C} = \mathrm{Span}_\mathbb{C}\{T^a, a = 1, \ldots, D\}.$$

However, going back to a real Lie algebra from a complex one is harder– there may be several real Lie algebras which have the same complexification. Instead, we say that $\mathfrak{g}_\mathbb{R}$ is a *real form* of $\mathfrak{g}_\mathbb{C}$.

Theorem 17.2. *Every complex semi-simple Lie algebra of finite dimension has a real form of compact type.*

This may be helpful on the final question of Example Sheet 2.

Continuing with our discussion of root strings, we previously defined

$$S_{\alpha,\beta} = \{\beta + \rho\alpha \ \in \Phi, \rho \in \mathbb{Z}\}.$$

We then argued that for $V_{\alpha,\beta}$ the representation space of a repn R of $sl(2)_\alpha$, there was a weight set

$$S_R = \{\frac{2(\alpha,\beta)}{\alpha,\alpha} + 2\rho; \beta + \rho\alpha \in \Phi\}.$$

We then noted that if we conisder the irreps R_Λ of $sl(2)_\alpha$ for some $\Lambda \in \mathbb{Z}_{\geq 0}$, we get

$$S_R = S_\Lambda = \{-\Lambda, -\Lambda + 2, \ldots, +\Lambda\}.$$

The allowed values of ρ are then $\rho = n \in \mathbb{Z}$ such that $n_- \leq n \leq n_+, n\pm \in \mathbb{Z}$ are some bounding values. By comparing the expression for the weight set with the minimum and maximum weights $\pm\Lambda$, we see that

$$-\Lambda = \frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n_-$$

$$+\Lambda = \frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n_+$$

$$\implies \frac{2(\alpha,\beta)}{(\alpha,\alpha)} = -(n_+ + n_-) \in \mathbb{Z}.$$

However, we also know that the allowed set of roots form an unbroken string,

$$S_{\alpha,\beta} = \{\beta + n\alpha; n \in \mathbb{Z}, n_- \leq n \leq n_+\}.$$

So this places a constraint on what the roots can be. This inner product constraint would be a lot stronger if we could guarantee the roots were real.

Let's pass for a moment to the Cartan-Weyl basis,

$$[H^i, E^\delta] = \delta^i E^\delta$$

where $i = 1, \ldots, r \forall \delta \in \Phi$. Then we write the Killing form as

$$\kappa^{ij} = \kappa(H^i, H^j) = \frac{1}{N} \operatorname{Tr}[\mathrm{ad}_{H^i} \circ \mathrm{ad}_{H^j}]$$

Now we remark that it would be very nice if these ad maps were mutually diagonal, since for

$$A = \begin{pmatrix} \lambda_1^A & & \\ & \ddots & \\ & & \lambda_n^A \end{pmatrix}, B = \begin{pmatrix} \lambda_1^B & & \\ & \ddots & \\ & & \lambda_n^B \end{pmatrix},$$

the trace is given simply by

$$\operatorname{Tr}[AB] = \sum_{i=1}^n \lambda_i^A \lambda_i^b.$$

So let us rewrite the ad maps in terms of the roots (which are of course just the eigenvalues when we diagonalize both maps):

$$\begin{aligned} \kappa^{ij} &= \frac{1}{N} \operatorname{Tr}[\mathrm{ad}_{H_i} \circ \mathrm{ad}_{H_j}] \\ &= \frac{1}{N} \sum_{\delta \in \Phi} \delta^i \delta^j. \end{aligned}$$

Moreover we know that

$$(\alpha, \beta) = \alpha^i \beta^j (\kappa^{-1})_{ij} = \frac{1}{N} \sum_{\delta \in \Phi} \alpha_i \delta^i \delta^j \beta_j,$$

where

$$\alpha_i \equiv (\kappa^{-1})_{ij} \alpha^j.$$

Now since $\alpha_i \delta^i = (\alpha, \delta)$, we see that

$$(\alpha, \beta) = \frac{1}{N} \sum_{\delta \in \Phi} (\alpha, \delta) \beta, \delta).$$

Thus the quantity

$$R_{\alpha,\beta} = \frac{2(\alpha, \beta)}{(\alpha, \alpha)} \in \mathbb{Z}.$$

Moreover

$$\frac{2}{(\beta, \beta)} R_{\alpha,\beta} = \frac{1}{N} \sum_{\delta \in \Phi} R_{\alpha,\delta} R_{\beta,\delta} \in \mathbb{R} \implies (\beta, \beta) \in \mathbb{R} \forall \beta \in \Phi.$$

We conclude that the inner product of two roots is always real,

$$(\alpha, \beta) \in \mathbb{R} \quad \forall \alpha, \beta \in \Phi.$$

Real Geometry of Roots Now that we know that the roots have real inner products, it makes good sense to discuss the real geometry of the dual space $\mathfrak{h}^*$. Let us now claim that the roots $\alpha \in \Phi$ are not only elements of $\mathfrak{h}^*$ but indeed span the dual space,

$$\mathfrak{h}^* = \operatorname{Span}_{\mathbb{C}}\{\alpha \in \Phi\}.$$

Proof. If the roots α do not span $\mathfrak{h}^*$, then $\exists \lambda \in \mathfrak{h}^*$ with

$$(\lambda, \alpha) = (\kappa^{-1})_{ij} \lambda^i \alpha^j = \kappa^{ij} \lambda_i \alpha_j = 0 \quad \forall \alpha \in \Phi,$$

i.e. another element λ which is orthogonal to all the roots. Thus we can construct

$$H_\lambda = \lambda_i H^i \in \mathfrak{h},$$

and we can compute some brackets now:

$$[H_\lambda, H] = 0 \quad \forall H \in \mathfrak{h}$$

and

$$[H_\lambda, E^\alpha] = (\lambda, \alpha)E^\alpha = 0 \quad \forall \alpha \in \Phi.$$

But this is very strange, because this means that $[H_\lambda, X] = 0 \forall X \in \mathfrak{g}$. This means that $\mathfrak{g}$ has a non-trivial ideal, namely

$$\mathfrak{j} = \mathrm{Span}_\mathbb{C}\{H_\lambda\}.$$

But this would mean that $\mathfrak{g}$ is not simple, so we have reached a contradiction. ☒

Therefore the r roots form a basis for $\mathfrak{h}^*$ (with complex coefficients), and we may then define the *real subspace*,

$$\mathfrak{h}^*_\mathbb{R} = \mathrm{Span}_\mathbb{R}\{\alpha_{(i)}; i = 1, \ldots, r\},$$

which is the *real* span of the roots.

Since the roots span $\mathfrak{h}^*$, any root $\beta \in \Phi$ can be written as

$$\beta = \sum_{i=1}^r \beta^i \alpha_{(i)}$$

with $\beta \in \mathbb{C}$ generically. However, if we take the inner product of β with each of the $\alpha_{(j)}$s, we find that

$$(\beta, \alpha_{(j)}) = \sum_{i=1}^r \beta^i (\alpha_{(i)}, \alpha_{(j)}.$$

However, we know that the inner products of the roots α are real, so $(a_{(i)}, \alpha_{(j)})$ considered as an $r \times r$ matrix is real, and $(\beta, \alpha_{(j)}$ considered as a vector of length r is also real. Therefore all the coefficients β^i must also be real, which means that all the roots live in the real subspace:

$$\beta \in \mathfrak{h}^*_\mathbb{R} \quad \forall \beta \in \Phi.$$

Lecture 18.

Tuesday, November 13, 2018

For a simple Lie algebra of dimension d, we defined the set of r roots α in some set Φ of size $|\Phi| = d - r$. In particular, we showed that the roots α not only lie in the dual space to the Cartan subalgebra $\mathfrak{h}^*$ but indeed they form a basis for $\mathfrak{h}^*$.

That is, a basis for $\mathfrak{h}^*$ is given by

$$B = \{\alpha_{(i)}, i = 1, \ldots, r, \alpha_{(i)} \in \Phi\}.$$

Now for a generic element $\beta \in \Phi$, it can be decomposed into its components

$$\beta = \sum_{i=1}^r \beta^i \alpha_{(i)}$$

where the coefficients are in general complex, $\beta^i \in \mathbb{C}$. We further reasoned that

$$(\beta, \alpha_{(j)}) = \sum_{i=1}^r \beta^i \underbrace{(\alpha_{(i)}, \alpha_{(j)})}_{(\kappa^{-1})_{ij}}.$$

But since the entries of κ^{-1} are real and $(\alpha, \beta) \in \mathbb{R} \forall \alpha, \beta \in \mathbb{R}$, this tells us that

$$\forall \beta \in \Phi, \beta \in \mathfrak{h}^*_\mathbb{R} = \mathrm{Span}_\mathbb{R}\{\alpha_{(i)}, i = 1, \ldots, r\}.$$

That is, a generic element of $\mathfrak{h}^*$ lies in the real span of the roots.

Not consider the inner product of two general elements of the dual space, $\lambda, \mu \in \mathfrak{h}^*_{\mathbb{R}}$.

$$\lambda = \sum_{i=1}^{r} \lambda^i \alpha_{(i)} \in \mathfrak{h}^*_{\mathbb{R}}$$
$$\mu = \sum_{i=1}^{r} \mu^i \alpha_{(i)} \in \mathfrak{h}^*_{\mathbb{R}},$$

with $\lambda^i, \mu^i \in \mathbb{R}$ (since these elements are in $\mathfrak{h}^*_{\mathbb{R}}$ and therefore in the real span).

But this means that their inner product also lies in the real span of the roots,

$$(\lambda, \mu) = \sum_{i,j=1}^{r} \lambda^i \mu^i (\alpha_{(i)}, \alpha_{(j)}) \in \mathbb{R}.$$

However, we also recall that the inner product of a vector with itself can be written as

$$(\lambda, \lambda) = \frac{1}{N} \sum_{\delta \in \Phi} \lambda^i \delta^i \delta^j \lambda_j = \frac{1}{N} \sum_{\delta \in \Phi} (\lambda, \delta)^2.$$

But now we see that since this inner product is a sum of squares of real numbers, $(\lambda, \lambda) \geq 0$ with equality iff $(\lambda, \delta) = 0 \forall \delta \in \Phi$. But by the non-degeneracy of the inner product, we see that there can be no element that is orthogonal to all other elements of Φ unless that element is the zero element, i.e.

$$(\lambda, \lambda) = 0 \iff \lambda = 0.$$

To summarize, we've recovered many of the nice properties of Euclidean space on the roots. The roots $\alpha \in \Phi$ live in a real vector space $\mathfrak{h}^*_{\mathbb{R}} \simeq \mathbb{R}^r$, such that $\forall \lambda, \mu \in \mathfrak{h}^*_{\mathbb{R}}$, the following properties hold:

(a) $(\lambda, \mu) \in \mathbb{R}$
(b) $(\lambda, \lambda) \geq 0$
(c) $(\lambda, \lambda) = 0 \iff \lambda = 0$,

which means that $\mathfrak{h}^*_{\mathbb{R}}$ admits a Euclidean inner product $(\cdot, \cdot)$.

Since $(\alpha, \alpha) > 0 \forall \alpha \in \Phi$, we can therefore define a *length* of a root defined as

$$|\alpha| \equiv +(\alpha, \alpha)^{1/2} > 0,$$

and thus an "angle" ϕ between two roots, which takes the standard form

$$(\alpha, \beta) = |\alpha||\beta| \cos \phi$$

with $\phi \in [0, \pi] \forall \alpha, \beta \in \Phi$.

However, we also recall that there was an integer quantization condition on the inner products, which we may write as

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = \frac{2|\beta|}{|\alpha|} \cos \phi \in \mathbb{Z} \tag{18.1}$$
$$\frac{2(\beta, \alpha)}{(\alpha, \alpha)} = \frac{2|\alpha|}{|\beta|} \cos \phi \in \mathbb{Z}. \tag{18.2}$$

We can of course multiply these two RHS expressions together to get another integer, and we find that

$$4 \cos^2 \phi \in \mathbb{Z},$$

which tells us that

$$\cos \phi = \pm \frac{\sqrt{n}}{2}$$

where $n \in \{0, 1, 2, 3, 4\}$.

Thus we have $\phi = 0, \pi/2, \pi$ corresponding to $\alpha = \beta, (\alpha, \beta) = 0, \alpha = -\beta$. We also have other possible values for ϕ: when $\phi = \pi/6, \pi/4, \pi/3$, then $(\alpha, \beta) > 0$, while when $\phi = 2\pi/3, 3\pi/4, 5\pi/6$, then $(\alpha, \beta) < 0$. To recap, not only can we define angles between two roots α, β, but these angles are also constrained to a finite set.

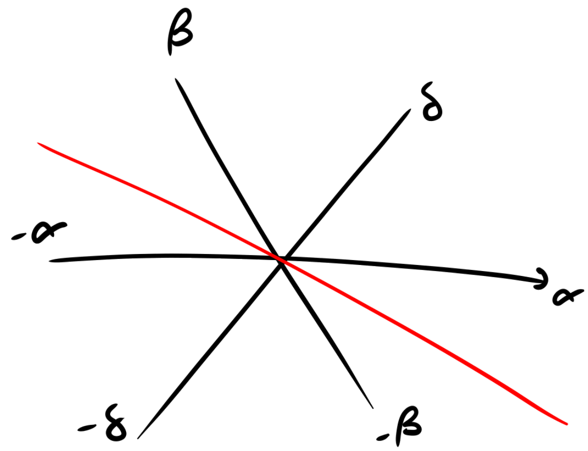

FIGURE 4. An illustration of the division of simple roots into two sets, Φ_+ and Φ_-. The red line represents the hyperplane, and α, β, δ are roots in Φ. Φ_+ lies above the red line and Φ_- below.

Simple roots Let us divide the roots $\alpha \in \Phi$ into positive and negative by a hyperplane in $\mathfrak{h}^*$. This hyperplane divides our set of roots Φ into

$$\Phi = \Phi_+ \cup \Phi_-$$

but is otherwise arbitrary. See Fig. 4 for an illustration. We can always construct such a plane by picking any plane that is not coplanar with any root (if it is, just move the plane a bit) and labeling all the roots on one side to be the $+$ roots and all the roots on the other to be $-$. Therefore $\forall \alpha, \beta \in \Phi$, we get the following nice properties:

(a) $\alpha \in \Phi_+ \iff -\alpha \in \Phi_-$
(b) $\alpha, \beta \in \Phi_+$ and $\alpha + \beta \in \Phi \implies \alpha + \beta \in \Phi_+$ (and similarly $\alpha, \beta \in \Phi_-, \alpha + \beta \in \Phi \implies \alpha + \beta \in \Phi_-$).

Definition 18.3. We now say that a *simple root* is a positive root which cannot be written as the sum of two positive roots, i.e.

$$\delta \in \Phi_S \text{ (is simple)} \iff \delta \in \Phi_+, \delta \neq \alpha + \beta \text{ for any } \alpha, \beta \in \Phi_+.$$

Simple roots have some good properties.

i) if $\alpha, \beta \in \Phi_S$, then $\alpha = \beta$ is not a root. For suppose $\alpha - \beta \in \Phi$. Then either
  - $\alpha - \beta \in \Phi_+$. Then $\alpha = \alpha - \beta + \beta \implies \alpha$ not simple.
  - $\alpha - \beta \in \Phi_-$. Then $\beta = \beta - \alpha + \alpha \implies \beta$ not simple.

  Either way, we reach a contradiction, so $\alpha - \beta \notin \Phi$.

ii) If $\alpha, \beta \in \Phi_S$, then the α-string through β ($\alpha \neq \beta$) takes the form

$$l_{\alpha,\beta} = 1 - \frac{2(\alpha, \beta)}{(\alpha, \alpha)} \in \mathbb{N}.$$

Proof. The string consists of roots

$$S_{\alpha,\beta} = \{\beta + n\alpha; n \in \mathbb{Z} \quad n_- \leq n \leq n_+\}.$$

But this set certainly contains at least one element– β, with $n = 0$. Therefore $n_+ \geq 0$ and $n_- \leq 0$. However, we also know that the sum of the bounds n_+, n_- is given by

$$(n_+ + n_-) = -\frac{2(\alpha, \beta)}{(\alpha, \alpha)} \in \mathbb{Z}.$$

Since α, β are simple roots, $\beta - \alpha \notin \Phi \implies n_- = 0$. Thus

$$n_+ = -\frac{2(\alpha, \beta)}{(\alpha, \alpha)} \in \mathbb{Z}_{\geq 0}.$$

We conclude that the root string takes the form

$$l = n_+ - n_- + 1 = n_+ + 1 = 1 - \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{N}.$$

⊠

iii) $\forall \alpha, \beta \in \Phi_S, \alpha \neq \beta$, we have $(\alpha, \beta) \leq 0$. (This follows from the previous property about the root string and the positivity of (α, α).)

iv) Any positive root $\beta \in \Phi_+$ can be written as a linear combination of simple roots with positive integer coefficients.

Proof. This is trivially true if $\beta \in \Phi_S$ is itself a simple root. If $\beta \notin \Phi_S$, then we can write $\beta = \beta_1 + \beta_2$ where $\beta_1, \beta_2 \in \Phi_+$. If $\beta_1, \beta_2 \in \Phi_S$, then we are done. Otherwise, we can split $\beta_1 = \beta_3, \beta_4$ where $\beta_3, \beta_4 \in \Phi_+$. The set of roots is of finite dimension, so this process has to terminate eventually with a full decomposition of β into simple roots. ⊠

v) Simple roots are linearly independent.

Proof. Consider vectors $\lambda \in \mathfrak{h}^*_{\mathbb{R}}$ which can be written in terms of the $\alpha_{(i)}$ basis elements. Thus

$$\lambda = \sum_{i \in J} c_i \alpha_{(i)}$$

with J a set of indices, $\alpha_{(i)} \in \Phi_S$, and $c_i \in \mathbb{R}$. We can split J into $J = J_+ \cup J_-$ where

$$J_+ = \{i \in J : c_i > 0\}, J_- = \{i \in J : c_i < 0\}$$

We'll finish the proof next time. ⊠

Lecture 19.

Thursday, November 15, 2018

Let's finish the proof from last time. We want to prove that the simple roots are linearly independent. The goal is this– we will write a general element in the dual space as a sum of the simple roots, and show that that element is nonzero unless all the coefficients of the basis vectors are zero.

Consider vectors $\lambda \in \mathfrak{h}^*_{\mathbb{R}}$, which we can write in terms of the basis elements $\alpha_{(i)}$. Thus

$$\lambda = \sum_{i \in \mathcal{I}} c_i \alpha_{(i)}$$

with $\mathcal{I}$ a set of indices, $\alpha_{(i)} \in \Phi_S$, and $c_i \in \mathbb{R}$ some real coefficients. We can split the set of indices $\mathcal{I}$ into $\mathcal{I} = \mathcal{I}_+ \cup \mathcal{I}_-$, where

$$\mathcal{I}_+ = \{i \in \mathcal{I} : c_i > 0\}, \mathcal{I}_- = \{i \in \mathcal{I} : c_i < 0\}.$$

We then define

$$\lambda_+ = \sum_{i \in \mathcal{I}_+} c_i \alpha_{(i)}$$
$$\lambda_- = -\sum_{i \in \mathcal{I}_-} c_i \alpha_{(i)}$$

so that

$$\lambda = \lambda_+ - \lambda_-,$$

and we take λ_+, λ_- to not both be zero.

$$\begin{aligned} \lambda &= \lambda_+ - \lambda_- \\ \implies (\lambda, \lambda) &= (\lambda_+, \lambda_+) + (\lambda_-, \lambda_-) - 2(\lambda_+, \lambda_-) \\ &> -2(\lambda_+, \lambda_-) \\ &= +2 \sum_{i \in \mathcal{I}_+} \sum_{j \in \mathcal{I}_-} c_i c_j (\alpha_{(i)}, \alpha_{(j)}) > 0. \end{aligned}$$

In going from the second to third lines, we have used the fact that the inner product of any nonzero element (specifically, λ_+ and λ_-) with itself is positive, and in the last line we have used the property that $c_i c_j < 0$ and $(\alpha_{(i)}, \alpha_{(j)}) < 0$. Thus $(\lambda, \lambda) > 0 \implies \lambda \neq 0$, so the simple roots are linearly independent. (That is, there is no general combination λ of the simple roots α which is zero.) ⊠

Since the simple roots are linearly independent, we now see that since $|\Phi_S| \leq r$, by iv) from last time we find that $\forall \beta \in \Phi_+, \beta = \sum c_i \alpha_{(i)}$ with $\alpha_{(i)} \in \Phi_S, c_i \in \mathbb{Z}_{\geq 0}$. If $\beta \in \Phi_-$, then $-\beta \in \Phi_+$ and $\beta = \sum \tilde{c}_i \alpha_{(i)}$. Either way, the *simple* roots entirely span the entire set $|\Phi|$, so $|\Phi_S| = r$.

Now let us choose the simple roots as a basis for $\mathfrak{h}^*_{\mathbb{R}}$. Thus the set

$$\begin{aligned} B &= \{\alpha \in \Phi_S\} \\ &= \{\alpha_{(i)} : i = 1, \ldots, r\} \end{aligned}$$

completely spans $\mathfrak{h}^*_{\mathbb{R}}$.

Definition 19.1. We now define the *Cartan matrix* A^{ij}, given by

$$A^{ij} = \frac{2(\alpha_{(i)}, \alpha_{(j)})}{(\alpha_{(j)}, \alpha_{(j)})} \in \mathbb{Z}, \tag{19.2}$$

with $i, j = 1, \ldots, r$. We know the elements of A are in $\mathbb{Z}$ by our previous results about root strings.

Now we will relate this back to subalgebras. For each $\alpha_{(i)} \in \Phi_S$, we get an $sl(2)_{\alpha_{(i)}}$ subalgebra spanned by

$$\{h^i = h^{\alpha_{(i)}}, e^i_\pm = e^{\pm\alpha_{(i)}}\}.$$

We call this the *Chevally basis*. How does this compare to our old subalgebras? Let's write down some brackets.

$$\begin{aligned} [h^i, h^j] &= 0 \quad \forall i, j = 1, \ldots, r \\ [h^i, e^j_\pm] &= \pm A^{ji} e^j_\pm \end{aligned}$$

Note that repeated indices are *not* summed over here. The bracket $[e^i_+, e^j_-]$ takes some care to compute. When $i = j$ it is just h^i, but when $i \neq j$, we have

$$[e^i_+, e^j_-] = n_{ij} e^{\alpha_{(i)} - \alpha_{(j)}} \text{ if } \alpha_{(i)} - \alpha_{(j)} \in \Phi,$$

and it is zero otherwise. However, since $\alpha_{(i)}, \alpha_{(j)}$ are simple roots by assumption, it must be that $\alpha_{(i)} - \alpha_{(j)} \notin \Phi$, so this bracket is *always zero*. We conclude that

$$[e^i_+, e^j_-] = \delta_{ij} h^i.$$

Here's another bracket:

$$\begin{aligned} [e^i_+, e^j_+] &= \mathrm{ad}_{e^i_+} e^j_+ \\ &\propto e^{\alpha_{(i)} + \alpha_{(j)}} \text{ if } \alpha^{(i)} + \alpha^{(j)} \in \Phi \\ &= 0 \text{ otherwise.} \end{aligned}$$

Indeed, we could repeat the ad map n times to get that

$$\begin{aligned} (\mathrm{ad}_{e^i_+})^n e^j_+ &\propto e^{n\alpha_{(i)} + \alpha_{(j)}} \text{ if } n\alpha^{(i)} + \alpha^{(j)} \in \Phi \\ &= 0 \text{ otherwise.} \end{aligned}$$

But the $\alpha_{(i)}$ root string through $\alpha_{(j)}$ has length

$$l_{ij} = 1 - \frac{2(\alpha_{(i)}, \alpha_{(j)})}{(\alpha_{(i)}, \alpha_{(i)})} = 1 - A^{ji}.$$

Therefore we derive the *Serre relation*:

$$(\mathrm{ad}_{e^i_\pm})^{1-A^{ji}} e^j_\pm = 0. \tag{19.3}$$

That is, if we apply the ad map enough times, we will eventually exhaust the elements of the root string.

What we've proved is that a finite-dimensional simple complex Lie algebra $\mathfrak{g}$ is completely determined by its Cartan matrix.

The Cartan matrix comes with some constraints. Recall it's defined as

$$A^{ij} \equiv \frac{2(\alpha_{(i)}, \alpha_{(j)})}{(\alpha_{(j)}, \alpha_{(j)})} \in \mathbb{Z},$$

Then it satisfies the following.

(a) $A^{ij} = 2$ with $i = 1, \ldots, r$.
(b) $A^{ij} = 0 \iff A^{ji} = 0$ (by the symmetry of the inner product in the numerator).
(c) $A^{ij} \in \mathbb{Z}_{\geq 0}$ if $i \neq j$ (since $\alpha_{(i)} \neq \alpha_{(j)} \in \Phi_S \implies (\alpha_{(i)}, \alpha_{(j)}) \leq 0$).
(d) $\det A > 0$.

Proof. Note that the Cartan matrix is proportional to κ^{-1} because of the inner product in the numerator, so we can write A as $A^{ij} = S^{ik} D^j_k$, where $S^{ik} = (\alpha_{(i)}, \alpha_{(k)}) = (\kappa^{-1})_{ik}$ represents the κ^{-1} part and $D^j_k = \frac{2}{(\alpha_{(j)}, \alpha_{(j)})} \delta^j_k$ is diagonal. It's true that $\det D > 0$, so we only need to prove that $\det S > 0$ and then our result is true by the property that determinants multiply. Note that κ^{-1} is a real symmetric matrix, so we can diagonalize it: $S = \kappa^{-1} = O \Lambda O^T$ with $\Lambda = \mathrm{diag}\{\rho_1, \ldots, \rho_r\}$ for $\rho_i \in \mathbb{R}$ an eigenvalue of κ^{-1}. But then for any eigenvector v^i_ρ, we have

$$(\kappa^{-1})_{ij} v^j_\rho = \rho \delta_{ij} v^j_\rho$$

so

$$(v_\rho, v_\rho) = (\kappa^{-1})_{ij} v^i_\rho v^j_\rho = \rho \sum_{i=1}^r (v^i_\rho)^2.$$

But $(v_\rho, v_\rho) = |v_\rho|^2 > 0$, so we conclude that each eigenvalue ρ of S is > 0. Thus, $\det S > 0$, which implies $\det A > 0$. ☒

Example 19.4. If we take $r = 2$, then the Cartan matrix looks like

$$A = \begin{pmatrix} 2 & -m \\ -n & 2 \end{pmatrix}.$$

Here, $m, n \in \mathbb{Z}_{\geq 0}$. Since $\det A > 0$ we know that $mn < 4$, and so we find that the possible values of m and n are very constrained:

$$\{m, n\} = \{0, 0\}, \{1, 2\}, \{1, 3\}, \{1, 1\}.$$

v Semi-simple Lie algebras correspond to reducible Cartan matrices (i.e. matrices which can be written in block diagonal form), so if the Lie algebra is simple then A is irreducible.

This allows us to exclude the boring solution of $\{m, n\} = \{0, 0\}$ (which would make A diagonal), leaving us with three possibilities for A.

Let us also note that relabeling the roots will in general permute the rows and columns of A. We don't want to think of these as "different" Cartan matrices, so next time we'll introduce a sort of diagram known as a Dynkin diagram which allows us to keep track of which As are related by such relabelings.