2 Intro to Lie Algebra

Note: Recall that the definition of the exponential of a matrix is the infinite sum (which converges for any square matrix)

$$e^X = \left(\sum_{n=0}^{\infty} \frac{X^n}{n!}\right) = I + X + \tfrac{1}{2}X^2 + \tfrac{1}{6}X^3 + \cdots$$

Infinitesimal rotations

When we have a *continuous group of matrices*, we know there are group elements arbitrarily close to the identity, and we will write

$$g(\epsilon) = I + \epsilon X$$

for ϵ an arbitrarily small real number. The matrix X is called a *generator*.

Suppose that our group is rotations of the plane (represented by $SO(2)$ or $U(1)$). Then $g(\epsilon)$ is a rotation by a very small angle $\epsilon > 0$. We can approximate rotation by a large angle θ by many repeated rotations of ϵ: $R(\theta) \simeq g(\epsilon)^k$, for $k \simeq \theta/\epsilon$. Rewriting with $\epsilon = \theta/k$ we see that

$$R(\theta) = g(\tfrac{\theta}{k})^k = \left(I + \frac{\theta}{k}X\right)^k \to e^{\theta X} \quad \text{as } k \to \infty$$

As θ is a continuous parameter we can differentiate with respect to it:

$$\frac{d}{d\theta}R(\theta) = Xe^{\theta X}$$

and then by setting $\theta = 0$, we will have that

$$X = \left.\frac{d\,R(\theta)}{d\theta}\right|_{\theta=0}$$

If we work with the unit-complex number representation of rotations, then $R(\theta) = e^{i\theta}$ so $X = i$ is the generator. If we work with the $SO(2)$ representation then

$$\begin{aligned} R(\theta) &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \\ \frac{d\,R(\theta)}{d\theta} &= \begin{pmatrix} -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix} \\ \text{so that } X &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

Exercise: Calculate X^2, X^3, X^4 when X is the 2×2 matrix above. How does this compare with i^2, i^3, i^4?

Exercise: Show that the 2×2 matrix $R(\theta) = e^{\theta X}$ by evaluating the expression for each entry of the matrix in the series expansion $e^{\theta X} = \sum \frac{(\theta X)^n}{n!}$.

The idea illustrated by this analysis of infinitesimal rotations applies to any Lie group, G, of matrix transformations. Suppose there is a parameter t so that $g(t)$ is a continuous path of group elements in G satisfying $g(0) = I$. Then there is a *generator*, X, defined by

$$X = \left.\frac{dg(t)}{dt}\right|_{t=0}, \quad \text{and } g(t) = e^{tX} \text{ for elements along this path.}$$

For a particular Lie group, G, the collection of all possible generators forms something called the *Lie algebra* and is written $\mathfrak{g}$. The group multiplication in G will induce an operation called the *Lie bracket* for combining elements of $\mathfrak{g}$. We first examine this for the case of the 3D rotation groups $SO(3)$.

Generators for the Lie Algebra of $SO(3)$

In a neighbourhood of the identity matrix, we have simple parametrisations of rotations about the three orthogonal axes, x, y, z. This is because for angles, $\alpha, \beta, \gamma = 0$, the matrices defined earlier $R_x(\alpha), R_y(\beta), R_z(\gamma)$ each become the identity. We compute the derivatives with respect to the angle coordinate in each case:

$$\frac{d}{d\alpha}R_x(\alpha) = \frac{d}{d\alpha}\begin{pmatrix}1 & 0 & 0\\ 0 & \cos(\alpha) & -\sin(\alpha)\\ 0 & \sin(\alpha) & \cos(\alpha)\end{pmatrix} = \begin{pmatrix}0 & 0 & 0\\ 0 & -\sin(\alpha) & -\cos(\alpha)\\ 0 & \cos(\alpha) & -\sin(\alpha)\end{pmatrix}$$

$$\frac{d}{d\beta}R_y(\beta) = \frac{d}{d\beta}\begin{pmatrix}\cos(\beta) & 0 & \sin(\beta)\\ 0 & 1 & 0\\ -\sin(\beta) & 0 & \cos(\beta)\end{pmatrix} = \begin{pmatrix}-\sin(\beta) & 0 & \cos(\beta)\\ 0 & 0 & 0\\ -\cos(\beta) & 0 & -\sin(\beta)\end{pmatrix}$$

$$\frac{d}{d\gamma}R_z(\gamma) = \frac{d}{d\gamma}\begin{pmatrix}\cos(\gamma) & -\sin(\gamma) & 0\\ \sin(\gamma) & \cos(\gamma) & 0\\ 0 & 0 & 1\end{pmatrix} = \begin{pmatrix}-\sin(\gamma) & -\cos(\gamma) & 0\\ \cos(\gamma) & -\sin(\gamma) & 0\\ 0 & 0 & 0\end{pmatrix}.$$

Evaluating each of these at $\alpha = \beta = \gamma = 0$ we have

$$\frac{d}{d\alpha}R_x(\alpha)\big|_{\alpha=0} = \begin{pmatrix}0 & 0 & 0\\ 0 & 0 & -1\\ 0 & 1 & 0\end{pmatrix} := X_1$$

$$\frac{d}{d\beta}R_y(\beta)\big|_{\beta=0} = \begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ -1 & 0 & 0\end{pmatrix} := X_2$$

$$\frac{d}{d\gamma}R_z(\gamma)\big|_{\gamma=0} = \begin{pmatrix}0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} := X_3.$$

As for the rotations of $SO(2)$, we now have the result that $R_x(\alpha) = e^{\alpha X_1}$, $R_y(\beta) = e^{\beta X_2}$, and $R_z(\gamma) = e^{\gamma X_3}$.

Now consider a rotation by θ about an arbitrary fixed axis, $\mathbf{n}$. We know that the matrix, $R_n(\theta)$, for this transformation is an element of $SO(3)$, and it will have a generator X (whose form depends on n) so that $R_n(\theta) = e^{\theta X}$. We don't have a general formula for writing down the matrix elements of $R_n(\theta)$, we only have expressions for it as a product of the coordinate axis rotation matrices. So we approach things a little indirectly and look at what other properties $R_n(\theta)$ has.

The fact that $R_n(\theta) \in SO(3)$ means that $R_n(\theta)^T R_n(\theta) = I$, and it is also the case that $(e^X)^T = e^{X^T}$. Putting these together we have

$$R_n(\theta)^T R_n(\theta) = e^{\theta X^T} e^{\theta X} = \left(\sum_{n=0}^{\infty}\frac{(\theta X^T)^n}{n!}\right)\left(\sum_{n=0}^{\infty}\frac{(\theta X)^n}{n!}\right) = I.$$

The only way for this matrix product to be I for $\theta > 0$ is if $X^T + X = 0$. In other words, X must be *anti-symmetric*: $X_{ji} = -X_{ij}$. This also implies that $X_{ii} = 0$, so that $\mathrm{tr}(X) = 0$. So the matrix for the generator of rotation about an arbitrary axis must have the form

$$X = \begin{pmatrix}0 & -c & b\\ c & 0 & -a\\ -b & a & 0\end{pmatrix} = aX_1 + bX_2 + cX_3$$

This tells us X is a linear combination of the three generators X_1, X_2, X_3, and in fact these three matrices form a basis for the Lie algebra $\mathfrak{so}(3)$ associated with the Lie group $SO(3)$.

Now we investigate what the matrix product in $SO(3)$ tells us about combinations of generators in $\mathfrak{so}(3)$. Take two group elements $g, h \in SO(3)$, with $g = e^X, h = e^Y$ where $X = (aX_1 + bX_2 + cX_3)$, and $Y = (dX_1 + eX_2 + fX_3)$ are two generators for $SO(3)$. We must have $e^X e^Y \in SO(3)$, and if X, Y were numbers, or elements of an Abelian group then $e^X e^Y = e^{X+Y}$. But X and Y are matrices and their product does not commute in general, so we must use the famous Baker-Campbell-Hausdorff formula:

$$\begin{aligned} e^X e^Y &= \left(\sum_{n=0}^{\infty} \frac{X^n}{n!}\right)\left(\sum_{n=0}^{\infty} \frac{Y^n}{n!}\right) \\ &= \exp\left[X + Y + \tfrac{1}{2}[X,Y] + \tfrac{1}{12}([X,[X,Y]] + [Y,[Y,X]]) + \cdots\right] \end{aligned}$$

Because we need to have $e^X e^Y = e^Z \in SO(3)$ for some other $Z \in \mathfrak{so}(3)$ we see that the Lie algebra must be closed with respect to the *commutator*, $[X,Y] := XY - YX$. In other words, the 'natural' group product for the Lie algebra is the commutator, NOT the regular matrix product.

In fact it is easy to check that $X_i X_j$ is not anti-symmetric for any choice of $i, j \in \{1,2,3\}$. And

$$[X_1, X_2] = X_3, \quad [X_2, X_3] = X_1, \quad \text{and } [X_3, X_1] = X_2.$$

So this implies that if $X = (aX_1 + bX_2 + cX_3)$ and $Y = (dX_1 + eX_2 + fX_3)$, then $[X,Y]$ is also a linear combination of the X_i.

Abstract definition of Lie Algebra

The construction we made above built the Lie algebra as a tangent space to the Lie group at the identity. Lie algebras can also be defined independently as an interesting and useful algebraic object in their own right, as described here.

A *real coefficient Lie algebra* of dimension n is a vector space, $\mathfrak{g}$, together with an operation called a *Lie product*, or *Lie bracket* written $[a,b]$, which must satisfy the four conditions below.

First of all, the vector space structure means that there is a basis of n elements $\{a_1, \ldots, a_n\}$ whose linear combinations span $\mathfrak{g}$. This means every element $a \in \mathfrak{g}$ can be written uniquely as

$$a = \sum_{p=1}^{n} \alpha_p a_p \quad \text{for some } \alpha_p \in \mathbb{R}.$$

Definition of Lie bracket

1. *Closure.* Given $a, b \in \mathfrak{g}$, we must have $[a,b] \in \mathfrak{g}$.
2. *Linearity.* Given $a, b, c \in \mathfrak{g}$, and $\alpha, \beta \in \mathbb{R}$, we have $[\alpha a + \beta b, c] = \alpha[a,c] + \beta[b,c]$.
3. *Anti-symmetry.* $[a,b] = -[b,a]$ for all $a, b \in \mathfrak{g}$.
4. *Jacobi's identity.* Given $a, b, c \in \mathfrak{g}$,
$$[a,[b,c]] + [b,[c,a]] + [c,[a,b]] = 0.$$

If the elements of the Lie algebra are square matrices, we define the Lie product as the commutator: $[a,b] = ab - ba$. With this defintion, the properties 2,3,4 automatically follow.

Asides:

- Even when the matrices $a \in \mathfrak{g}$ have complex entries, we work with the real-coefficient Lie algebra, only allowing real coefficients in linear combinations of the basis.
- Note that the Lie bracket properties are also satisfied by the *Poisson bracket* of Hamiltonian mechanics, for example. So a Lie algebra can consist of objects other than matrices.

- Ado's theorem: Any finite-dimensional abstract Lie algebra satisfying the above conditions will be isomorphic to a matrix Lie algebra with $[,]$ being the standard matrix commutator. This justifies our focus on matrix Lie groups and algebras.

Physics versus Mathematics defintions of Lie Algebras

The above study of rotation groups uses the matrix exponential function $e^{\theta X}$ to map elements of the Lie algebra to elements of the Lie group $SO(3)$. This means all the coefficients and matrix entries have real values, and we saw that X is an anti-symmetric matrix with $\mathrm{tr}(X) = 0$.

Physicists use a different definition to map elements from a Lie algebra to the Lie group: $e^{i\theta J}$. The elements of the group must be the same, so this means the entries of J are now imaginary, with $J = -iX$. We find that the complex-transpose, $J^\dagger = iX^T = -iX = J$, so that J is now a *Hermitian* matrix. The advantage here is that in quantum-mechanical settings, J often corresponds to a physical observable and will have real eigenvalues. The disadvantage is that this notation muddies the distinction between real-coefficient and complex-coefficient Lie algebras. In mathematics texts, there is a careful distinction between Lie algebras with real coefficients and those with complex coefficients as the vector spaces can be quite different, as we will see in later examples. However, we will now follow the physics convention in these notes.

Exercise: Write out the matrices J_1, J_2, J_3 for the physics definition of the Lie algebra $\mathfrak{so}(3)$.

Structure constants

Let $\{T_1, \ldots, T_m\}$ be a basis for an m-dimensional real-coefficient Lie algebra. Every $T \in \mathfrak{g}$ is uniquely written as $T = \sum_{a=1}^m \theta_a T_a$, with $\theta_a \in \mathbb{R}$. This means the *commutator* of two basis elements can be written as a linear combination:

$$[T_a, T_b] = i \sum_{c=1}^{m} f_{abc} T_c.$$

The numbers f_{abc} are called the *stucture constants*. They tell us exactly how to combine elements of the Lie algebra. If two Lie algebras have the same dimension and same structure constants then these Lie algebras have the same abstract structure and are said to be *isomorphic*.

At first glance it appears that there are up to m^3 numbers required to specify the structure constants, but the properties of the Lie bracket impose some constraints. Anti-symmetry implies $f_{bac} = -f_{abc}$, and Jacobi's identity gives

$$\sum_{s=1}^{m} [f_{asr} f_{bcs} + f_{bsr} f_{cas} + f_{csr} f_{abs}] = 0$$

Exercise: Write out the structure constants for $\mathfrak{so}(3)$ with respect to the (physics) basis J_1, J_2, J_3.

The Lie algebra of SU(2)

Recall that $SU(2)$ is the group of 2×2 matrices with complex elements, such that $U^\dagger U = I$ and $\det(U) = 1$. We saw in the tutorial that $SU(2)$ has three free parameters, and it will therefore have three generators in the basis for its Lie algebra, $\mathfrak{su}(2)$. The next few steps derive conditions on the generators and construct a basis for the Lie algebra, using the indirect method.

Suppose $U = e^{i\theta A}$. The condition that U is unitary tells us that

$$(e^{i\theta A})^\dagger e^{i\theta A} = e^{-i\theta A^\dagger} e^{i\theta A} = e^{i\theta A} e^{-i\theta A^\dagger} = I = e^0$$

It follows that A is a *Hermitian* matrix: $A^\dagger = A$. Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{where } a, b, c, d \in \mathbb{C}$$

Then the property of being Hermitian tells us that

$$a^* = a, \quad d^* = d, \quad c^* = b$$

So a, d are real and the off-diagonal elements are $b = x - iy$, $c = x + iy$. This means we have a matrix form:

$$A = \begin{pmatrix} a & x - iy \\ x + iy & d \end{pmatrix} \quad \text{where } a, d, x, y \in \mathbb{R}.$$

Now we use the condition that $U \in SU(2)$ means $\det(U) = 1$, and the matrix identity $\det(e^X) = e^{\mathrm{tr}(X)}$.

$$1 = \det(e^{i\theta A}) = e^{\mathrm{tr}(i\theta A)}, \quad \text{implying} \quad \mathrm{tr}(i\theta A) = 0.$$

Applying this to the matrix form for A above, we see that in fact $a = -d = z$ and the general form for a generator is

$$A = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix} \quad \text{where } x, y, z \in \mathbb{R}.$$

A possible basis for the Lie algebra $\mathfrak{su}(2)$ is therefore

$$s_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad s_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad s_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

You might recognise these as the Pauli matrices.

Our next task is to find the structure constants from the commutators of the basis elements.

Exercise: Show that

$$[s_1, s_2] = 2is_3, \quad [s_2, s_3] = 2is_1, \quad [s_3, s_1] = 2is_2.$$

Exercise: Show that if we set $A_k = \frac{1}{2}s_k$, then

$$[A_1, A_2] = iA_3, \quad [A_2, A_3] = iA_1, \quad [A_3, A_1] = iA_2.$$

These are exactly the same commutation relations as we found for the generators J_1, J_2, J_3 of $\mathfrak{so}(3)$. What is going on here? How are 2x2 complex matrices in $SU(2)$ related to rotations of three-dimensional space?

Let's use a correspondence between generators to construct a mapping between the two Lie algebras: $f : \mathfrak{su}(2) \to \mathfrak{so}(3)$ defined by $f(A_i) = J_i$ and extended linearly so that for any A,

$$\begin{aligned} A &\in \mathfrak{su}(2), & A &= aA_1 + bA_2 + cA_3, \\ f(A) &\in \mathfrak{so}(3), & f(A) &= aJ_1 + bJ_2 + cJ_3 \end{aligned}$$

This mapping f is an *isomorphism*: it defines a one-to-one correspondence between the elements of the two Lie algebras, that respects the Lie bracket (commutator).

What does this tell us about the Lie groups $SU(2)$ and $SO(3)$? We know that $SO(3)$ is related to the unit quaternions in a fairly natural way, so let's see if we can find a relationship (mapping) between $SU(2)$ and the unit quaternions.

SU(2) and the unit quaternions

The defining property of $U \in SU(2)$ is that $U^\dagger U = I$ and $\det(U) = 1$. This implies the matrices U have the following form:

$$U = \begin{pmatrix} a+di & b+ci \\ -b+ci & a-di \end{pmatrix} \quad \text{with } a^2+d^2+b^2+c^2 = 1$$

Exercise: Check that $U^\dagger U = I$.

We define a mapping g from the unit quaternions to $SU(2)$ using the following correspondence, and extending by linearity

$$g(\mathbf{1}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad g(\mathbf{i}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad g(\mathbf{j}) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad g(\mathbf{k}) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$
$$g(q) = g(a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) = \begin{pmatrix} a+di & b+ci \\ -b+ci & a-di \end{pmatrix}$$

The condition $\det(U) = 1$ is equivalent to q being a unit quaternion.

Exercise: Check that $g(q_1 q_2) = g(q_1)g(q_2)$, i.e., the mapping respects the product operation. (This means g is a group homomorphism.)

The mapping g is one-to-one and reaches all elements of $SU(2)$, so g is an isomorphism between the unit quaternions and $SU(2)$. Now we have

$$SU(2) \overset{\text{one-to-one}}{\longleftrightarrow} U(1,Q) \overset{\text{two-to-one}}{\longrightarrow} SO(3)$$

so we say that the Lie group $SU(2)$ is a double cover of the Lie group $SO(3)$.

Formal definitions

So far we have described Lie groups just as some type of continuous group. The modern mathematical definition is quite involved.

Definition: A *Lie group* is a group, G, with product $\circ$. Elements of the group form a differentiable manifold M, and the product induces a differentiable map f_g of the manifold onto itself

$$f_g(h) = g \circ h \quad \text{is differentiable with respect to } h.$$

The map of M taking a group element to its inverse must also be differentiable, $v : M \to M$ with $v(h) = h^{-1}$. If we have coordinates for M, then this means the coordinates of $g \circ h$ are differentiable functions of the coordinates for h, and this must hold for all $g \in G \simeq M$.

Example: The group $SO(2)$ has the unit circle S^1 as its manifold, and the product map

$$R(\phi)R(\theta) = \begin{pmatrix} \cos(\phi+\theta) & -\sin(\phi+\theta) \\ \sin(\phi+\theta) & \cos(\phi+\theta) \end{pmatrix} = R(\phi+\theta)$$

is then $f_\phi(\theta) = \phi + \theta$, which clearly maps S^1 back onto itself and is a differentiable function of θ. The inverse map is

$$R(\phi)^{-1} = R(\phi)^T = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} = \begin{pmatrix} \cos(-\phi) & -\sin(-\phi) \\ \sin(-\phi) & \cos(-\phi) \end{pmatrix} = R(-\phi)$$

So $v(\phi) = -\phi$ is clearly a continuous mapping of S^1.

If we take this geometric perspective of a Lie group, then the Lie algebra is seen to be the tangent space to the Lie group at its identity. The fact that $SU(2)$ and $SO(3)$ have the same structure for their Lie algebras is due to the fact that they have the same local manifold and group product structure around their identity elements. We have also seen that they do not have the same global structure: the manifold for $SU(2)$ is the three-sphere, S^3, while the manifold for $SO(3)$ is $\mathbb{R}P^3$, the three-sphere with antipodal points identified.

This illustrates one of the deep theorems in Lie theory: every finite-dimensional Lie algebra $\mathfrak{g}$ has a unique simply connected Lie group G associated to it, with elements $g \in G$ defined by exponentiating elements $X \in \mathfrak{g}$: $g = e^X$. This unique simply-connected Lie group, G, is also called the universal covering group. Another theorem tells us that there is a continuous k-to-one map from G onto any other group G' that has the same Lie algebra, $p : G \to G'$ with p being a local isomorphism.

Physics interpretation

The group $SO(3)$ is a natural object for us to study because its matrices represent rotations of space. However, we've seen that it is closely related to the group $SU(2)$, which has a simpler topology (the simply-connected 3-sphere), and so it is the largest possible group associated with their common Lie algebra. Formally, $SU(2)$ is the universal covering group of all Lie groups that have a Lie algebra isomorphic to the one generated as

$$[A_1, A_2] = iA_3, \quad [A_2, A_3] = iA_1, \quad [A_3, A_1] = iA_2$$

We now take a look at how this Lie algebra appears in a physical setting: namely in the quantum theory of orbital angular momentum. Component-wise, the orbital angular momentum operator is

$$\begin{aligned}
\hat{L}_x &= -i\hbar(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}) \\
\hat{L}_y &= -i\hbar(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}) \\
\hat{L}_z &= -i\hbar(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x})
\end{aligned}$$

and has commutation relations

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \quad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y.$$

We have the familiar pattern of commutators again, demonstrating that the *algebraic structure* of orbital angular momentum and $\mathfrak{su}(2)$ are closely related.

3 Representations of Lie groups and algebras

Symmetries of a quantum Hamiltonian operator

Suppose that H is invariant with respect to a *unitary* transformation T, i.e., $T^\dagger HT = H$. T being unitary implies $[H, T] = 0$, and the set of all unitary transformations that commute with H forms a group, G.

Now consider an eigenfunction for the Hamiltonian, $H\psi = E\psi$. Then

$$H(T\psi) = (HT)\psi = (TH)\psi = T(H\psi) = TE\psi = E(T\psi)$$

meaning that $T\psi$ is another eigenfunction for H with the same eigenvalue E.

Quantum operators are *linear* and their eigenfunctions span a Hilbert space. Suppose the eigenfunctions with identical eigenvalue E span a d-dimensional space with basis $\{\psi_1, \ldots, \psi_d\}$. We know that for any of the unitary transformations, $T\psi_a$ must also have eigenvalue E. So for each a, $T\psi_a = \sum_b t_{ab}\psi_b$.

The coefficients t_{ab} form a d-dimensional matrix representation for the group element $T \in G$, with respect to the vector space with basis $\{\psi_1, \ldots, \psi_d\}$. On this vector space, H acts as a multiple of the identity matrix since $H\psi_a = E\psi_a$.

A different eigenvalue for H may have a different multiplicity of states, and so this eigenspace will provide a different set of matrices for the group elements T.

This example serves to illustrate and motivate the next topic, *representation theory.*

Basic definitions for Representations

We have seen that groups come in many different guises, with different operations, as abstract entities, or as transformations acting on a space. For example, the Lie algebra $\mathfrak{su}(2)$ can be built, or realised, using 2×2 matrices, 3×3 matrices or differential operators. So even though we started with a definition of $\mathfrak{su}(2)$ as 2×2 matrices, this is not intrinsic to its algebraic structure.

Representation theory is a body of mathematics that studies how to find all the possible ways a group can be built using matrices from $GL(n, \mathbb{C})$.

Definition: Given a group G, a *representation of dimension* n is a homomorphism $\rho : G \to GL(n, \mathbb{C})$. This means ρ has the following properties:

- It maps the identity $e \in G$ to the identity matrix $I_n \in GL(n, \mathbb{C})$: $\rho(e) = I_n$.
- Inverses map to matrix inverses: $\rho(g^{-1}) = [\rho(g)]^{-1}$.
- The group product in G maps to matrix multiplication: $\rho(gh) = \rho(g).\rho(h)$.

Definition: A representation ρ is called *faithful* if the only element that maps to the identity matrix is the identity: $\rho(g) = I_n$ if and only if $g = e \in G$.

Definition: The *trivial* representation (of dimension 1) is the map that sends every element to the identity: $\rho(g) = 1 = e^0$. This is clearly not a faithful representation (unless you start with the trivial group).

Example: Representations of $SO(2)$

We start with our one-dimensional Lie group of rotations, and use it as a simple illustration of some of the more complicated definitions required in general representations. We already know that $SO(2) \simeq U(1) = \{e^{i\theta}, \theta \in S^1\}$. But in fact the group $SO(2)$ has 1-dimensional representations as $\rho_{1,k}(R_\theta) = e^{ik\theta}$, for k an integer.

- Identity: $\rho_{1,k}(R_0) = e^0 = 1$; and $\rho_{1,k}(R_{2\pi}) = e^{ik2\pi} = 1$ provided k is an integer.
- Inverse: $\rho_{1,k}(R_\theta^{-1}) = e^{-ik\theta} = [e^{ik\theta}]^{-1}$;

· Product: $\rho_{1,k}(R_\theta R_\phi) = e^{ik(\theta+\phi)} = e^{ik\theta}e^{ik\phi} = \rho_{1,k}(R_\theta)\rho_{1,k}(R_\phi)$.

Note that only $k = \pm 1$ give us faithful representations, and $k = 0$ is the trivial representation.

Higher dimensional examples of reps for $SO(2)$ include a 2-dimensional representation as the standard rotation matrices, and a 3-dimensional rep as $\rho(R_\theta) = R_{z,\theta}$. When we work in $GL(2, \mathbb{C})$ we can diagonalise the standard (real) rotation matrix using a change-of-basis transformation. First of all we find the eigenvalues and eigenvectors of R_θ are

$$\lambda_1 = e^{i\theta},\, v_1 = \begin{pmatrix} i \\ 1 \end{pmatrix} \qquad \lambda_2 = e^{-i\theta},\, v_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$$

Now set

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}$$

The change-of-basis (or similarity) transformation converts the standard matrix for R_θ into $D_\theta = S^{-1}R_\theta S$

$$D_\theta = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} \begin{pmatrix} -i & 1 \\ i & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

Using the formal definition, we have two representations $\rho_2 : SO(2) \to GL(2, \mathbb{C})$ defined as

$$\rho_2(R_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

and $\rho_d : SO(2) \to GL(2, \mathbb{C})$ with

$$\rho_d(R_\theta) = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} = S^{-1}\rho_2(R_\theta)S$$

The really important point here is that S is independent of θ, so the similarity transformation makes the two *representations* equivalent, not just two matrices. We won't count ρ_2 and ρ_d as different representations.

The final thing to note is that the diagonal elements of $\rho_d(R_\theta)$ are copies of 1-dimensional reps. This means we have shown that ρ_d is a *reducible* representation: it is built from smaller representations acting on independent, orthogonal subspaces of $\mathbb{C}^2$. We write this as a direct sum: $\rho_d = \rho_{1,1} \oplus \rho_{1,-1}$.

Equivalent representations

As we saw above, two n-dimensional representations of a group G, say ρ_1, ρ_2, are equivalent or similar if there is a (single) linear transformation $S \in GL(n, \mathbb{C})$ so that $\rho_2(g) = S^{-1}\rho_1(g)S$ for every $g \in G$.

Reducible and irreducible representations

An n-dimensional representation of a group G is *reducible* if it is equivalent to a representation ρ for which every linear transformation $\rho(g) \in GL(n, \mathbb{C})$ has a fixed upper-triangular block form, e.g.,

$$\rho(g) = \begin{pmatrix} R_{11}(g) & R_{12}(g) \\ 0 & R_{22}(g) \end{pmatrix}$$

where $R_{11}(g)$ are $k \times k$ matrices, $R_{22}(g)$ are $(n-k) \times (n-k)$ matrices and $R_{12}(g)$ are $k \times (n-k)$. This block form means that the first k coordinates and the last $(n-k)$ coordinates form invariant subspaces, so R_{11}, R_{22} are k- and $(n-k)$-dimensional representations for G respectively.

An n-dimensional representation of a group G is *completely reducible* if it is equivalent to a representation ρ for which every linear transformation $\rho(g) \in GL(n, \mathbb{C})$ has a fixed *diagonal* block form, e.g.,

$$\rho(g) = \begin{pmatrix} R_{11}(g) & 0 & \cdots & 0 \\ 0 & R_{22}(g) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & R_{rr}(g) \end{pmatrix}$$

Where each R_{kk} is an *irreducible* representation for the group G. This means the coordinates belonging to each block form invariant and independent subspaces.

Aside: It can be shown that if ρ is a finite-dimensional *unitary* representation of a matrix Lie group then ρ is completely reducible.

For example, the group $SO(2)$ has a 3-dimensional representation as the group of rotations fixing the z-axis:

$$\rho(R_\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

These matrices already have a block form of $(2 \times 2) \oplus (1 \times 1)$. The 1 block is just the identity; the third coordinate of any point is unchanged by the action of this matrix. The 2×2 block acts only on the first and second coordinates, but mixes them up. However the 2×2 block can be further reduced to two 1×1 blocks by a similarity $S \in GL(3, \mathbb{C})$ built from the eigenvectors for $\rho(R_\theta)$ because they are independent of θ.

$$S^{-1}\rho(R_\theta)S = \begin{pmatrix} e^{i\theta} & 0 & 0 \\ 0 & e^{-i\theta} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

In this basis of eigenvectors, each (complex) coordinate is transformed separately; we write $\rho = \rho_{1,1} \oplus \rho_{1,-1} \oplus \rho_{1,0}$.

An *irreducible* representation is defined as one that is not reducible.

In other words, an irreducible representation has no invariant subspaces other than the trivial one and all of $\mathbb{C}^n$. What this means is that all representations are either irreducible, or they are built up in a block-matrix fashion from smaller irreducible parts. The irreducible representations are the fundamental building blocks in representation theory. Our challenge now is to find all the irreducible representations (shortened to *irreps*) for a given group.