MATH-F410 – REP. DES GROUPES & APP. A LA PHYS.

Solutions des exercices 2023-2024

Séance 6 : rappels sur $SU(2)$ et sur l'exponentielle d'une matrice

1. $\mathfrak{su}(2)$ algebra

Section a)

Rotations are orthogonal matrices, so $R^T R = I$. If we do an expansion around the identity, $R = I + \epsilon J$, to first order in ϵ this equation is

$$I + \epsilon(J^T + J) + \mathcal{O}(\epsilon^2) = I \ ,$$

from which it follows $J^T = -J$ (J is antisymmetric).

Section b)

Rotation matrices by an angle θ around the three Euclidean axes are:

$$R_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \ , \quad R_2 = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \ , \quad R_3 = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \ .$$

A linear expansion for $\theta = \epsilon$ gives the following three generators:

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \ , \qquad J_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \ , \qquad J_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ .$$

One can check the algebra from these forms by taking products, e.g., $J_1 J_2 - J_2 J_1 = J_3$. Alternatively, looking at the form of the J_i one realizes $(J_i)_{jk} = -\epsilon_{ijk}$, where indices outside the parentheses denote matrix elements and ϵ_{ijk} is the Levi-Civita symbol. Using identities of this symbol then, and adopting Einstein summation convention for repeated indices:

$$\begin{aligned} ([J_i, J_j])_{lm} &= \epsilon_{ilk}\epsilon_{jkm} - \epsilon_{jlk}\epsilon_{ikm} = -\left(\delta_{ij}\delta_{lm} - \delta_{im}\delta_{jl}\right) + \left(\delta_{ij}\delta_{lm} - \delta_{il}\delta_{jm}\right) \\ &= \delta_{im}\delta_{jl} - \delta_{il}\delta_{jm} = \epsilon_{ijk}\epsilon_{mlk} = -\epsilon_{ijk}\epsilon_{klm} = \epsilon_{ijk}(J_k)_{lm} \ . \end{aligned}$$

Section c)

$$g\left([X,Y], Z\right) = \mathrm{Tr}\left(XYZ - YXZ\right) = \mathrm{Tr}\left(XYZ - XZY\right) = g\left(X, [Y,Z]\right) \ .$$

Section d)

We have to compute the metric and invert it to define $g^{ij} J_i J_j = J^2$:

$$\left(J_i J_j\right)_{lm} = \epsilon_{ilk}\epsilon_{jkm} = \delta_{im}\delta_{jl} - \delta_{ij}\delta_{lm} \ ,$$

which upon taking $l \to m$ and summing over m gives $\mathrm{Tr}(J_i J_j) = -2\delta_{ij}$. Thus, $g_{ij} = -\frac{1}{2}\mathrm{Tr}(J_i J_j) = \delta_{ij}$, and the inverse is just $g^{ij} = \delta^{ij}$. The Casimir is then

$$(J^2)_{lm} = \delta^{ij}(J_i J_j)_{lm} = -2\delta_{lm} \ .$$

It is therefore proportional to the identity, $J^2 = -2I$, and trivially commutes with any of the operators in the algebra.

Section e)

From now on, we turn to generators $\tilde{J}_k = iJ_k$, so that they are Hermitian: $\tilde{J}_k^\dagger = -iJ_k^T = iJ_k = \tilde{J}_k$ (we drop tildes since we will always be dealing with the Hermitian generators). The algebra becomes $[J_i, J_j] = i\epsilon_{ijk}J_k$. The commutation relations presented in the question are just trivial applications of this basic ones:

$$\begin{aligned}
[J_3, J^\pm] &= \frac{1}{\sqrt{2}}\left([J_3, J_1] \pm i[J_3, J_2]\right) = \frac{1}{\sqrt{2}}\left(iJ_2 \pm J_1\right) = \pm J^\pm \ , \\
[J^+, J^-] &= \frac{1}{2}\left(-i[J_1, J_2] + i[J_2, J_1]\right) = J^3 \ , \\
(J^+)^\dagger &= \frac{1}{\sqrt{2}}\left(J_1 + iJ_2\right)^\dagger = \frac{1}{\sqrt{2}}\left(J_1 - iJ_2\right) = J^- \ .
\end{aligned}$$

Section f)

Let $|m\rangle$ be an eigenvector of J_3 with eigenvalue m. Then

$$J_3 J^\pm |m\rangle = \left([J_3, J^\pm] + J^\pm J_3\right)|m\rangle = (\pm 1 + m)\, J^\pm |m\rangle \ ,$$

so $J^\pm |m\rangle$ is an eigenvector of J_3 with eigenvalue $m \pm 1$ (unless it vanishes).

2. Representations of $\mathfrak{su}(2)$

Before starting, keep in mind that the J_i now are not the explicit 3×3 matrices of the previous section. They are a representation of the algebra we found, so they are generic (Hermitian) operators satisfying the same commutation relations.

Section a)

The space of greatest eigenvalues of J_3 is spanned by the $|j, \alpha\rangle$, $\alpha = 1, \ldots, k$ (α is labeling different eigenvectors with the same, highest eigenvalue j). There are two useful results for this question that we will often use:

$$\begin{aligned}
J^+ J^- &= \frac{1}{2}\left(J_1^2 + J_2^2 - i[J_1, J_2]\right) = \frac{1}{2}\left(J^2 - J_3^2 + J_3\right) \ , \\
J^- J^+ &= \frac{1}{2}\left(J_1^2 + J_2^2 + i[J_1, J_2]\right) = \frac{1}{2}\left(J^2 - J_3^2 - J_3\right) \ .
\end{aligned}$$

Furthermore, since $|j, \alpha\rangle$ is an eigenvector with maximum eigenvalue of J_3, it must be $J^+ |j, \alpha\rangle = 0$ (because otherwise J^+ would increase the eigenvalue). Thus,

$$0 = J^- J^+ |j, \alpha\rangle = \frac{1}{2}\left(J^2 - j^2 - j\right)|j, \alpha\rangle \ ,$$

from which we conclude that $J^2 |j,\alpha\rangle = j(j+1)|j,\alpha\rangle$. Since J^2 commutes with all the operators J_i, any state obtained by applying J^- from these ones will have the same eigenvalue for J^2. We can now start from one of the $|j,\alpha\rangle$ (i.e., for fixed α) and decrease the eigenvalue with respect to J_3 in steps of one unit by applying J^-: $J^-|m,\alpha\rangle = \lambda_m |m-1,\alpha\rangle$. The norm of the state after applying J^- is obtained by means of the previous formulas, using also $(J^-)^\dagger = J^+$:

$$\langle m,\alpha|J^+J^-|m,\alpha\rangle = \frac{1}{2}\left(j(j+1) - m(m-1)\right) ,$$

assuming $|m,\alpha\rangle$ has unit norm. There are two consequences of this:

- The right-hand side vanishes for $m = j+1$ or for $m = -j$, signalling that in those cases $J^-|m,\alpha\rangle = 0$. The first case is not relevant because $m \leq j$ always by assumption (j is the greatest eigenvalue). In a finite-dimensional representation, the chain of eigenvectors must terminate: the second case tells us that it does after reaching $m = -j$.
- If we want to have normalized states $|m-1,\alpha\rangle$, where $J^-|m,\alpha\rangle = \lambda_m |m-1,\alpha\rangle$, we must pick (up to a phase) $\lambda_m = \frac{1}{\sqrt{2}}\sqrt{j(j+1)-m(m-1)}$. Then, the previous result shows that $|m-1,\alpha\rangle$ has unit norm.

The previous steps should be though as a recursive procedure. We start from $|j,\alpha\rangle$, which is assumed to be normalized, and then we use the previous formula for $m = j$ and define λ_j so that $|j-1,\alpha\rangle$ is normalized. We then go to $|j-2,\alpha\rangle$, $|j-3,\alpha\rangle, \ldots$, until we reach $|-j,\alpha\rangle$, where the chain stops ($J^-|-j,\alpha\rangle = 0$). We used m above to encompass all these steps at once. Notice that, in the subspace of fixed α states, $V_\alpha = \text{span}\{|m,\alpha\rangle, m = -j, -j+1, \ldots, j-1, j\}$, the operators $J_1 = \frac{1}{\sqrt{2}}(J^+ + J^-)$, $J_2 = \frac{i}{\sqrt{2}}(J^- - J^+)$ and J^3 act irreducibly (they do not leave any subspace of V_α invariant, and they map V_α to itself, because neither J_3 nor $J^\pm$ change the value of α).[1] If the representation of the algebra is to be irreducible, there must be a single value of α, so $k = 1$ in the notation of the question. Otherwise we would have separate subspaces which are invariant under the action of the algebra operators.

Section b)

Say we need to apply J^- n times to reach eigenvalue $m = -j$ starting from the largest one, $m = j$. Then, $j - n = -j$. It follows that $j = n/2$ for $n \in \mathbb{N}$.

Section c)

This part of the question changes notation a bit. We drop α (we assume we are in an irrep, so there is a single value of α), and we indicate explicitly j in the vectors (j can be thought as the greatest J_3 eigenvalue, or as giving the eigenvalue of J^2 through $j(j+1)$).

[1] Technically, we did not check how J^+ acts on the states $|m,\alpha\rangle$. Let us do it now: $J^+|m-1,\alpha\rangle = \frac{1}{\lambda_m}J^+J^-|m,\alpha\rangle = \frac{1}{2\lambda_m}(j(j+1)-m(m-1))|m,\alpha\rangle = \lambda_m|m,\alpha\rangle$. So it does not change α.

Other than that, we have all the results needed to compute the inner products (notice that vectors with different values of m are orthogonal, because they have different J_3 eigenvalues):

$$\begin{aligned}
\langle j, m'|J^2|j, m\rangle &= j(j+1)\delta_{m',m} \ , \\
\langle j, m'|J_3|j, m\rangle &= m\delta_{m',m} \ , \\
\langle j, m'|J^+|j, m\rangle &= \lambda_{m+1} \langle j, m'|j, m+1\rangle = \frac{1}{\sqrt{2}}\sqrt{j(j+1) - m(m+1)}\delta_{m',m+1} \ , \\
\langle j, m'|J^-|j, m\rangle &= \lambda_m \langle j, m'|j, m-1\rangle = \frac{1}{\sqrt{2}}\sqrt{j(j+1) - m(m-1)}\delta_{m',m-1} \ .
\end{aligned}$$

Matrix elements of J_1 and J_2 are obtained from those of $J^{\pm}$.

3. Tensor product representations

It is useful to understand why the tensor product representation of a Lie algebra has the form $\mathbf{J} = \mathbf{J}_1 \otimes \mathbf{1} + \mathbf{1} \otimes \mathbf{J}_2$ (a priori, one could expect something like $\mathbf{J}_1 \otimes \mathbf{J}_2$). This is a consequence of how the tensor product representation of the corresponding group is defined (and it is essential to have the desired bracket structure in the tensor product representation of the algebra). Let $T_i(g) = \mathbf{1} + \epsilon\, t_i(x)$ be a representative of a group element infinitesimally close to the identity, with $t_i(x)$ the representative of the Lie algebra element. The tensor product representation of T_1 and T_2 assigns to g the operator

$$T_1(g) \otimes T_2(g) = (\mathbf{1} + \epsilon\, t_1(x)) \otimes (\mathbf{1} + \epsilon\, t_2(x)) = \mathbf{1} + \epsilon\ (t_1(x) \otimes \mathbf{1} + \mathbf{1} \otimes t_2(x)) + \dots \ .$$

Thus, the tensor product representation of the algebra is $t(x) = t_1(x) \otimes \mathbf{1} + \mathbf{1} \otimes t_2(x)$.

Section a)

Recalling that operators on different tensor factors act independently,

$$\begin{aligned}
[J_i, J_j] &= [J_{1,i} \otimes \mathbf{1} + \mathbf{1} \otimes J_{2,i}\ ,\ J_{1,j} \otimes \mathbf{1} + \mathbf{1} \otimes J_{2,j}] \\
&= [J_{1,i}, J_{1,j}] \otimes \mathbf{1} + \mathbf{1} \otimes [J_{2,i}, J_{2,j}] = i\epsilon_{ijk} J_k \ .
\end{aligned}$$

Section b)

The square of $\mathbf{J}$ (understood as sum of the squares of the components) is

$$\mathbf{J}^2 = \mathbf{J}_1^2 \otimes \mathbf{1} + \mathbf{1} \otimes \mathbf{J}_2^2 + 2\sum_i J_{1,i} \otimes J_{2,i} \ .$$

The rest are just consequences of the commutation relations:

$$\begin{aligned}
[\mathbf{J}^2, \mathbf{J}_1^2 \otimes \mathbf{1}] &= 2\sum_i \left[J_{1,i} \otimes J_{2,i}, \mathbf{J}_1^2 \otimes \mathbf{1}\right] = 0 \ , \\
[\mathbf{J}^2, J_{1,z} \otimes \mathbf{1}] &= 2\sum_i [J_{1,i} \otimes J_{2,i}, J_{1,z} \otimes \mathbf{1}] = 2\left(-iJ_{1,y} \otimes J_{2,x} + iJ_{1,x} \otimes J_{2,y}\right) \ .
\end{aligned}$$

4. Exponential of a matrix

Section a)

Assume X and Y commute. The only tricky step is how to organize the double sum,

$$e^X e^Y = \sum_{n,m=0}^{\infty} \frac{1}{n!m!} X^n Y^m = \sum_{N=0}^{\infty} \sum_{n=0}^{N} \frac{1}{n!(N-n)!} X^n Y^{N-n} = \sum_{N=0}^{\infty} \frac{1}{N!} (X+Y)^N = e^{X+Y} \ .$$

In the second step we computed first the sum at fixed $N = n+m$, and then summed over all values of N. In the third step, it is essential that the matrices commute to use the binomial theorem (for non-commuting matrices it is not true, you can convince yourself of this in a simple example like $(X+Y)^2$ for non-commuting X and Y).

Section b)

Since $[X, -X] = 0$, the previous section shows $e^X e^{-X} = e^0 = I$. Thus, $(e^X)^{-1} = e^{-X}$.

Section c)

The Hermitian conjugate of a sum is the sum of the Hermitian conjugates, thus in the series definition of the exponential we can just apply Hermitian conjugation to each term to obtain $(e^X)^\dagger = e^{X^\dagger}$.

Section d)

Taking derivatives in the series expansion:

$$\frac{d}{dt} e^{tX} = \frac{d}{dt} \sum_{n=0}^{\infty} \frac{1}{n!} t^n X^n = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} t^{n-1} X^n = \sum_{n=0}^{\infty} \frac{1}{n!} t^n X^{n+1} = e^{tX} X = X e^{tX} \ .$$

Section e)

This is easy for X diagonalizable. In that case, we can write $D = PXP^{-1}$, with $D = \mathrm{diag}\,(\lambda_1, \dots, \lambda_n)$ the diagonal matrix containing the eigenvalues. Then, it is clear that $e^D = \mathrm{diag}\left(e^{\lambda_1}, \dots, e^{\lambda_n}\right)$ and using $e^X = e^{P^{-1}DP} = P^{-1} e^D P$ (easy to see from the series expansion),

$$\det\left(e^X\right) = \det\left(P^{-1} e^D P\right) = \det\left(e^D\right) = e^{\lambda_1 + \dots + \lambda_n} = e^{\mathrm{Tr}(D)} = e^{\mathrm{Tr}\left(PXP^{-1}\right)} = e^{\mathrm{Tr}(X)} \ .$$

For generic matrices, diagonalization is not always possible. We can however find a Jordan normal form, in which the D in the previous argument is upper triangular, with the eigenvalues along the diagonal. Taking powers of D, for instance D^k, produces along the diagonal the numbers $(\lambda_1^k, \dots, \lambda_n^k)$. Furthermore, D^k is also upper triangular. The argument then is analogous, because e^D is an upper triangular matrix with $(e^{\lambda_1}, \dots, e^{\lambda_n})$ along the diagonal, which therefore satisfies $\det\left(e^D\right) = e^{\lambda_1 + \dots + \lambda_n} = e^{\mathrm{Tr}(D)}$.