Math 280–Statistical Mechanics
Brief Review of SU(2) Representations
10 October 2000

There are many references for representation theory in general, and the representations of $SU(2)$ in particular. Three that I used are

- A. Edmonds, *Angular Momentum in Quantum Mechanics* 1957.
- Wu-Ki Tung, *Group Theory in Physics* 1985.
- Fulton & Harris, *Representation Theory* 1991.

1 $\mathfrak{su}_2$, $\mathfrak{so}_3(\mathbb{R})$ and $\mathfrak{sl}_2(\mathbb{C})$

The group $SU(2)$ is the group of unitary 2×2 matrices with determinant 1. Every such matrix can be uniquely written as

$$\mathcal{U}(z,w) = \begin{pmatrix} z & -w \\ \overline{w} & \overline{z} \end{pmatrix}$$

for $(z,w) \in \mathbb{C}^2$, with the condition that $|z|^2 + |w|^2 = 1$. In other words, $SU(2)$ is topologically equivalent to the unit sphere in $\mathbb{C}^2$, which is the same as the real 3-sphere. $SU(2)$ is a real Lie group, meaning it is a group with a compatible structure of a real manifold.

A Lie algebra $\mathfrak{g}$ is a vector space with a bilinear form $[.,.] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called a Lie bracket, satisfying

1. $[Y,X] = -[X,Y]$,
2. $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$.

The Lie algebra $\mathfrak{su}_2$ is defined as the tangent space to $SU(2)$ at the identity. We obtain the tangent space by taking all limits

$$A(Z,W) = \lim_{\epsilon \to 0, \epsilon \in \mathbb{R}} \frac{\mathcal{U}(1+\epsilon Z, \epsilon W) - \mathcal{U}(1,0)}{\epsilon} = \begin{pmatrix} Z & -W \\ \overline{W} & \overline{Z} \end{pmatrix}$$

for those Z and W which satisfy $\det \mathcal{U}(1+\epsilon Z, \epsilon W) = 1+O(\epsilon^2)$, i.e. $\mathrm{Re}(Z) = 0$. These are the matrices in M_2 satisfying $A^* = -A$. This is a three-dimensional, real vector space with basis iS^1, iS^2, iS^3, where

$$S^1 = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}, \quad S^2 = \begin{pmatrix} 0 & -\frac{i}{2} \\ \frac{i}{2} & 0 \end{pmatrix}, \quad S^3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$$

are the spin-$\frac{1}{2}$ matrices from before. Recall the spin matrices satisfy the commutation relations $[S^1, S^2] = iS^3$, $[S^2, S^3] = iS^1$, $[S^3, S^1] = iS^2$.

The group $SU(2)$ is studied in connection with the quantization of angular momentum. One may wonder if $SO_3(\mathbb{R})$, the group of rotations in real, three-dimensional space, is a better group to study. The rotations about the three axes are given by the matrices

$$\begin{aligned} R_1(\theta) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \\ R_2(\theta) &= \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \\ R_3(\theta) &= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

The group $SO(3)$ is generated by these matrices for $\theta \in [0, 2\pi)$. It is easy to calculate the derivative of each of these matrices at zero:

$$r_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \; r_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \; r_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

These are the basis elements for $\mathfrak{so}_3(\mathbb{R})$. So $\mathfrak{so}_3(\mathbb{R})$ consists of all skew-orthogonal ($A^T = -A$), real 3×3 matrices. Moreover, the basis elements satisfy the commutation relations $[r_\alpha, r_\beta] = \varepsilon_{\alpha\beta\gamma} r_\gamma$. It is customary to define $J^\alpha = i r_\alpha$, whereupon we recover the commutation relations $[J^\alpha, J^\beta] = -\varepsilon_{\alpha\beta\gamma} J^\gamma$. So, in fact $\mathfrak{so}_3(\mathbb{R})$ is exactly the same as $\mathfrak{su}_2$.

Moreover, $SO_3(\mathbb{R})$ is not a simply connected group, while $SU(2)$ is. Indeed, $SU(2)$ is a double-cover of $SO_3(\mathbb{R})$ which can be obtained by considering the representation of $SU(2)$ on $\mathbb{R}^3$ wherein the three-dimensional vectors are actually traceless, hermitian matrices $X = x^1 S^2 + x^2 S^2 + x^3 S^3$, and $SU(2)$ acts by conjugation $X \to X' = \mathcal{U} X \mathcal{U}^*$ for $\mathcal{U} \in SU(2)$. Since X' is still hermitian, and traceless, $\mathcal{U}$ actually defines a linear transformation on this three-dimensional, real space. Note that $-\det(X) = |x^1|^2 + |x^2|^2 + |x^3|^2$, and $\det(X') = \det(X)$, so that the linear transformation is actually orthogonal. Finally, it is special because $SU(2)$ is connected (so the determinant of the image cannot take both values 1 and -1). Thus we have a map of $SU(2)$ into $SO(3)$. It is two-to-one because A and $-A$ both induce the same map.

There is a one-to-one correspondence between the representations of a Lie algebra and Lie group, when the Lie group is connected and simply connected. This means that the representations of $\mathfrak{su}_2$ and $\mathfrak{so}_3(\mathbb{R})$ are both the same as the representations of $SU(2)$, but not of $SO(3)$. This explains why we study $SU(2)$ instead of $SO(3)$.

The Lie algebra $\mathfrak{su}_2$ is a real Lie algebra. It can be thought of as a real subspace of $\mathcal{B}(\mathbb{C}^2)$. It is often useful to have a complex subspace instead. We can define the complexification of $\mathfrak{su}_2$, which is simply the complex vector space spanned by S^1, S^2, S^3. It is easy to see that this is $\mathfrak{sl}_2(\mathbb{C})$, the set of all trace-zero complex matrices in M_2. (This is the Lie algebra for the Lie group $SL_2(\mathbb{C})$: all complex 2×2 matrices with determinant 1.) We define $S^{\pm} = S^1 \pm iS^2 \in \mathfrak{sl}_2(\mathbb{C})$, which satisfy the commutation relations

$$[S^3, S^+] = S^+, \quad [S^3, S^-] = -S^-, \quad [S^+, S^-] = 2S^3$$

The matrices (S^3, S^+, S^-) generate $\mathfrak{sl}_2(\mathbb{C})$ just as well as (S^1, S^2, S^3). S^+ and S^- are the raising and lowering operators.

2 Representations

Any linear map $\rho : \mathfrak{su}_2 \to M_n$ such that $[\rho(iS^\alpha), \rho(iS^\beta)] = -\varepsilon_{\alpha\beta\gamma}\rho(iS^\gamma)$, is called an n-dimensional representation of $\mathfrak{su}_2$. Such a representation is specified by the images $\rho(iS^\alpha)$, $\alpha = 1, 2, 3$. A representation of $\mathfrak{sl}_2(\mathbb{C})$ is a linear map such that

$$[\rho(S^3), \rho(S^{\pm})] = \pm\rho(S^{\pm}), \quad [\rho(S^+), \rho(S^-)] = 2\rho(S^3)$$

Any representation of $\mathfrak{su}_2$ can be extended to a representation of $\mathfrak{sl}_2(\mathbb{C})$, and any representation of $\mathfrak{sl}_2(\mathbb{C})$ can be restricted to a representation of $\mathfrak{su}_2$.

If ρ satisfies $\rho(A^*) = \rho(A)^*$, then it is called a unitary repesentation. An important point is that if ρ is not unitary, a priori, we can always redefine the inner-product on $\mathbb{C}^n$ so as to make it unitary. This is because $SU(2)$ is a compact group, and so has a unique Haar measure, H. The unique Haar measure is characterized by the fact that $H(E) = H(\mathcal{U}E)$ for every measurable $E \subset SU(2)$ and $\mathcal{U} \in SU(2)$, and $H(SU(2)) = 1$. Whatever inner-product is on $\mathbb{C}^n$ initially, we average over $SU(2)$:

$$\langle \psi | \phi \rangle' = \int_{SU(2)} \langle \mathcal{U}\psi | \mathcal{U}\phi \rangle \, dH(\mathcal{U}) \, .$$

With the new inner-product on $\mathbb{C}^n$, ρ is unitary. From now on all representations are unitary.

A representation is irreducible if there is no proper, invariant subspace $V \subset \mathbb{C}^n$. An invariant subspace is one for which $\rho(S^\alpha)v \in V$ for every $v \in V$ and $\alpha = 1, 2, 3$. The entire list of finite-dimensional, irreducible representations was given in lecture 2. They are specified by the spin S matrices. We will not prove this here; it is proved in each of the three references above. Suppose that W is an invariant subspace of $\mathbb{C}^n$. Then the orthogonal complement $W^\perp$ is also invariant, since ρ is unitary. This proves that every finite-dimensional representation of $\mathfrak{su}_2$ (and $\mathfrak{so}_3(\mathbb{R})$ and $\mathfrak{sl}_2(\mathbb{C})$) is completely reducible; i.e. it can be decomposed into a direct sum of irreducible representations.

Suppose that $\rho_1 : \mathfrak{su}_2 \to \mathcal{B}(V_1)$ and $\rho_2 : \mathfrak{su}_2 \to \mathcal{B}(V_2)$ are two representations of $\mathfrak{su}_2$ on two f.d. complex vector spaces V_1 and V_2. Then there is a representation $\rho : \mathfrak{su}_2 \to \mathcal{B}(V_1 \otimes V_2)$ given by

$$\rho(S^\alpha) = \rho_1(S^\alpha) \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes \rho_2(S^\alpha)$$

where $\mathbb{1}_j$ is the identity operator on V_j, $j = 1, 2$. It is trivial to check that this satisfies the commutation relations, since for $A_1, B_1 \in V_1$, $A_2, B_2 \in V_2$:

- $[\mathbb{1}_1 \otimes A_2, B_1 \otimes \mathbb{1}_2] = 0$,
- $[\mathbb{1}_1 \otimes A_2, \mathbb{1}_1 \otimes B_2] = \mathbb{1}_1 \otimes [A_2, B_2]$,
- $[A_1 \otimes \mathbb{1}_2, B_1 \otimes \mathbb{1}_2] = [A_1, B_1] \otimes \mathbb{1}_2$.

It is also trivial to check that the tensor product of two unitary representations is again unitary. In general it is not true that if V_1 and V_2 are irreducible representations then the tensor product $V_1 \otimes V_2$ is also irreducible. It is therefore a natural question to ask how $V_1 \otimes V_2$ decomposes into irreducibles. This is the Clebsch-Gordon problem, which we will now discuss.

3 Clebsch-Gordon coefficients

Instead of considering arbitrary representations V_1, V_2 we will consider irreducible representations. The reason is that we already know V_1 and V_2 can de decomposed into irreducibles

$$V_1 = V_{1,1} \oplus \cdots \oplus V_{1,r_1}, \quad V_2 = V_{2,1} \oplus \cdots \oplus V_{2,r_2}\,.$$

Since $\otimes$ is distributive w.r.t. $\oplus$, we see that in general

$$V_1 \otimes V_2 = \bigoplus_{i_1=1}^{r_1} \bigoplus_{i_2=1}^{r_2} V_{1,i_1} \otimes V_{2,i_2}\,.$$

If we can say what each $V_{1,i_1} \otimes V_{2,i_2}$ is in terms of irreducibles, then we can determinde the direct sum decomposition of $V_1 \otimes V_2$. Since V_{1,i_1} and V_{2,i_2} are each irreducible, it suffices to solve the Clebsch-Gordon problem for V_1 and V_2 both irreducible.

The entire list of irreps of $\mathfrak{su}_2$ was given in Lecture 2, as the spin S matrices. Let us suppose we have $S = j_1$ and $S = j_2$. Therefore, $V_1 = \mathbb{C}^{2j_1+1}$ and $V_2 = \mathbb{C}^{2j_2+1}$. We will refer to these representations as $\mathcal{D}^{(j_1)}$ and $\mathcal{D}^{(j_2)}$, following standard practice. The states of V_α will be labeled by $\psi_\alpha(j_\alpha, m_\alpha)$ where $m_\alpha = -j_\alpha, -j_\alpha + 1, \cdots, j_\alpha$, $(\alpha = 1, 2)$. This is chosen so that

$$S^3_\alpha \psi_\alpha(j_\alpha, m_\alpha) = m_\alpha \psi_\alpha(j_\alpha, m_\alpha)\,.$$

We define the Casimir operator $|\mathbf{S}_\alpha|^2 = (S^1_\alpha)^2 + (S^2_\alpha)^2 + (S^3_\alpha)^2$. Observe that by the derivation property of $[.,.]$ (namely, $[AB, C] = A[B, C] + [A, C]B$), we have

$$[A, B^2] = \{[A, B], B\}$$

where $\{X, Y\} = XY + YX$ is the anticommutator. Therefore

$$[S^3_\alpha, (S^1_\alpha)^2] = \{[S^3_\alpha, S^1_\alpha], S^1_\alpha\} = i\{S^2_\alpha, S^1_\alpha\}\,.$$

Similarly,

$$[S^3_\alpha, (S^2_\alpha)^2] = \{[S^3_\alpha, S^2_\alpha], S^2_\alpha\} = -i\{S^1_\alpha, S^2_\alpha\}\,.$$

Finally, $[S^3_\alpha, (S^3_\alpha)^2] = 0$ for obvious reasons. Since the anticommutator is symmetric, we see that $[S^3_\alpha, |\mathbf{S}_\alpha|^2] = 0$. By permutation symmetry, we also have $[S^1_\alpha, |\mathbf{S}_\alpha|^2] = [S^2_\alpha, |\mathbf{S}_\alpha|^2] = 0$. Therefore $[S^\pm_\alpha, |\mathbf{S}_\alpha|^2] = 0$, and we see that $|\mathbf{S}_\alpha|^2$ acts on V_α as a constant times the identity matrix. To see just what constant, we calculate it on $\psi_\alpha(j_\alpha, j_\alpha)$. Note that $|\mathbf{S}_\alpha|^2 = \frac{1}{2}(S_\alpha^+ S_\alpha^- + S_\alpha^- S_\alpha^+) + (S_\alpha^3)^2$. Therefore, since $S_\alpha^+ \psi_\alpha(j_\alpha, j_\alpha) = 0$,

$$\begin{aligned} |\mathbf{S}_\alpha|^2 \psi_\alpha(j_\alpha, j_\alpha) &= \frac{1}{2} S_\alpha^+ S_\alpha^- \psi_\alpha(j_\alpha, j_\alpha) + j_\alpha^2 \psi_\alpha(j_\alpha, m_\alpha) \\ &= j_\alpha(j_\alpha + 1)\psi_\alpha(j_\alpha, m_\alpha) \end{aligned}$$

An important fact is that this Casimir operator can distinguish vectors in different irreps because its eigenvalue is an injective function of j.

The representation on $V_1 \otimes V_2$ is generated by the operators $S^3 = S_1^3 + S_2^3$ and $S^\pm = S_1^\pm + S_2^\pm$. In general this will be a direct sum of irreps given by spin j matrices for j taking on some values. Note that the Casimir operator for the tensor product is

$$|\mathbf{S}|^2 = |\mathbf{S}_1 + \mathbf{S}_2|^2 = |\mathbf{S}_1|^2 + |\mathbf{S}_2|^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2 \,,$$

where

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = S_1^3 S_2^3 + \frac{1}{2}(S_1^+ S_2^- + S_1^- S_2^+) \,.$$

Since $|\mathbf{S}_1|^2$ and $|\mathbf{S}_2|^2$ are essentially constants times the identity matrix, $(|\mathbf{S}_1|^2, |\mathbf{S}_2|^2, |\mathbf{S}|^2, S^3)$ is a commuting family of operators. We prefer to keep the decorations $|\mathbf{S}_1|^2$,$|\mathbf{S}_2|^2$ since it makes clear what the dimensions of the two irreps are that we are tensoring. Suppose $\psi(j_1, j_2, j, m)$ is a simultaneous eigenstate with eigenvalues $(j_1^2 + j_1, j_2^2 + j_2, j^2 + j, m)$. Then

$$\psi(j_1, j_2, j, m) = \sum_{\substack{m_1, m_2 \\ m_1 + m_2 = m}} M(j_1, m_1, j_2, m_2; j_1, j_2, j, m)\psi_1(j_1, m_1) \otimes \psi_2(j_2, m_2) \,.$$

It is an abuse of notation to label the eigenstates $\psi(j_1, j_2, j, m)$ before proving that for each quadruple of eigenvalues there is at most one eigenstate. But we will systematically prove this fact regardless of the label for the eigenstates, so the abuse is not important.

We consider $\psi_1(j_1, j_1) \otimes \psi_2(j_2, j_2)$. This is a simulatneous eigevector of $|\mathbf{S}_1|^2$, $|\mathbf{S}_2|^2$ and S^3, with eigenvalues $j_1(j_1 + 1)$, $j_2(j_2 + 1)$ and $j_1 + j_2$. But also, since $S_\alpha^+ \psi_1(j_1, m_1) \otimes \psi_2(j_2, m_2) = 0$ for $\alpha = 1, 2$, we see that it is an eigenvector of $\mathbf{S}_1 \cdot \mathbf{S}_2$ with eigenvalue $j_1 j_2$. This means it is an eigenvector of $|\mathbf{J}|^2$, with eigenvalue $(j_1 + j_2)(j_1 + j_2 + 1)$. In other words,

$$\psi_1(j_1, m_2) \otimes \psi_2(j_2, m_2) = \psi(j_1, j_2, j_1 + j_2, j_1 + j_2) \,.$$

So there is at least one copy of the irrep $\mathcal{D}^{j_1+j_2}$, generated by the spin j_1+j_2 matrices. This is the only state $\psi(j_1,j_2,j,j_1+j_2)$ for any j. Thus every irrep in the direct sum decomposition of $V_1\otimes V_2$ has spin at mose j_1+j_2, and in fact there is only one copy of that irrep. (If there were any other irrep with spin at least j_1+j_2 it would contain an eigenstate of J^3 with eigenvalue j_1+j_2, orthogonal to the one we just determined.)

Suppose now that we have proved there is a unique copy of the irrep $\mathcal{D}^{(j)}$ in $\mathcal{D}^{(j_1)}\otimes\mathcal{D}^{(j_2)}$ for $j=j_1+j_2,j_1+j_2-1,\ldots,j'$, where $j'>|j_1-j_2|$. The eigenspace of J^3 with eigenvalue $j'-1$ has dimension $j_1+j_2-j'+2$, while the eigenstates $\{\psi(j_1,j_2,j,j'-1)\in\mathcal{D}^{(j)}:j=j_1+j_2,j_1+j_2-1,\ldots,j'\}$ only account for a $(j_1+j_2-j'+1)$-dimensional subspace. Taking the unique vector orthogonal to all of these yields a state $\psi(j_1,j_2,j,j'-1)$ with $j<j'$. (Because the orthogonal complement of an invariant subspace is invariant.) But since the third component of spin for this state is $j'-1$, it must be that $j=j'-1$. So there is at least one copy of $\mathcal{D}^{(j'-1)}$. Since any other copy of $\mathcal{D}^{(j'-1)}$ would give an additional orthogonal state in the eigenspace of J^3 with eigenvalue $j'-1$, there is a unique copy of $\mathcal{D}^{(j'-1)}$.

Thus we have proved that there is a unique copy of $\mathcal{D}^{(j)}$ in the tensor product $\mathcal{D}^{(j_1)}\otimes\mathcal{D}^{(j_2)}$ for each $j=j_1+j_2,j_1+j_2-1,\ldots,|j_1-j_2|$. To see that this is actually the entire list of irreps, note that the dimensions match. The dimension of $\mathcal{D}^{(j_1)}\otimes\mathcal{D}^{(j_2)}$ is $(2j_1+1)(2j_2+1)$, and

$$\dim(\mathcal{D}^{(j_1+j_2)}\oplus\mathcal{D}^{(j_1+j_2-1)}\oplus\cdots\oplus\mathcal{D}^{(|j_1-j_2|)})=\sum_{j=|j_1-j_2|}^{j_1+j_2}(2j+1)=(2j_1+1)(2j_2+1)\,.$$

We have thus solved the problem of stating which irreps appear in the direct sum decomposition of $V_1\otimes V_2$. We have not said what the matrix $M(j_1,m_1,j_2,m_2;j_1,j_2,j,m)$ is, which connects them. Foregoing the analysis, the result is zero unless $m=m_1+m_2$ and

$$\begin{aligned}&M(j_1,j_2,m_1+m_2,j;j_1,m_2,j_2,m_2)\\&=\left[\frac{(2j+1)(j_1+j_2-j)!(j_1-j_2+j)!(-j_1+j_2+j)!}{(j_1+j_2+j+1)!}\right]^{1/2}\\&\times[(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!(j+m)!(j-m)!]^{1/2}\\&\times\sum_z(-1)^z\frac{1}{z!(j_1+j_2-j-z)!(j_1-m_1-z)!(j_2+m_2-z)!(j-j_2+m_1+z)!(j-j_1-m_2+z)!}\end{aligned}$$

Details of this calculation, as well as more symmetric forms of the vector-coupling coefficients can be found in Edmonds.