A tutorial for SU(2) and spin waves

Shannon Starr

16 June 2014
(Last updated 25 June 2014)

1 Overview

These are tutorial notes that are intended to accompany Bruno Nachtergaele's lectures on an introduction to quantum spin systems for the NSF-CBMS Regional

The first part is intended to be a review of quantum spins and especially SU(2) as introduced, for example, in a first course on quantum mechanics. Many students will recognize all the facts in the first part from an undergraduate physics course:

- The Lie group SU(2).
- Representations of SU(2).
- Tensor products of representations.

The best reference for this material is probably any upper division (or introductory graduate) textbook on quantum mechanics, such as

> **[1]** *Quantum Mechanics: Two Volumes Bound as One,* Albert Messiah. Dover Publications, Inc., Mineola, NY, 1999. (Reprint of 1958 edition, John Wiley & Sons.)

There are also excellent mathematical references such as

> **[2]** *Representations of Finite and Compact Groups,* Barry Simon. American Mathematical Society, Providence, RI, 1995.

In the second part of these notes we will try to introduce the topic of

- Spin waves in the quantum Heisenberg ferromagnet.

A reference for spin waves at the same level as the one for quantum spins would be, for example, Kittel's, "Introduction to Solid State Physics," or Ashcroft and Mermin's, "Solid State Physics." But in the United States it is less common for mathematics graduate students to have seen statistical mechanics or condensed matter physics than a basic course in quantum mechanics. Therefore, we will just try to develop the subject from scratch. Our goal is to get an intuitive idea, at least, of the topic which would allow for an appreciation of the following important recent paper:

[**3**] Michele Correggi, Alessandro Giuliani and Robert Seiringer. Validity of the spin-wave approximation for the free energy of the Heisenberg ferromagnet. *Preprint, 2013.* `http://arxiv.org/abs/1312.7873`

2 The Lie Group SU(2) and the quaternions

The Lie group SU(2) is the symmetry group of the quantum spin. For example, the electron carries a spin with magnitude 1/2 in the physics language. Other particles have spin, such as the photon which has spin 1. But many of the origins of quantum spins in solid state physics are due to the spin of the electron.

Compton suggested that the electron would have an intrinsic spin beyond its orbital angular momentum, and later Goudsmit and Uhlenbeck proved that the electon spin is described by a representation of SU(2) with total spin 1/2, in the physics language. An excellent reference for some of the history of quantum spin systems is

[**4**] *The Theory of Magnetism Made Simple. An Introduction to Physical Concepts and to Some Useful Mathematical Methods.* Daniel C. Mattis. World Scientific Publ. Co. Inc, Singapore, 2006.

In particular, the website for the textbook is here

`http://www.worldscientific.com/worldscibooks/10.1142/5372`

On that website the first two chapters are available for download for free. They describe the history of magnetism and the history of the discovery of the intrinsic spin of an electron.

Of course there is also a treatment in Messiah, [**1**], in Chapter XIII, "Angular Momentum in Quantum Mechanics," specifically in part IV, "Spin."

Let us simply accept that SU(2) is a Lie group of interest, for the moment. Let $M_2(\mathbb{C})$ denote the set of all 2×2 matrices

$$A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}, \qquad \alpha, \beta, \gamma, \delta \in \mathbb{C}.$$

Then SU(2) is defined as the special unitary group of 2×2 matrices

$$\mathrm{SU}(2) = \{U \in M_2(\mathbb{C}) : \det(U) = 1 \text{ and } U^*U = \mathbb{1}\}.$$

For the time being, we will denote the 2×2 identity matrix as $\mathbb{1}$.

The condition that $U^*U = \mathbb{1}$ clearly makes the matrix U unitary. (This is the definition of unitarity, since that implies that for any complex [column] vector $x \in \mathbb{C}$, we have for $y = Ux$ that $\|y\|^2 = y^*y = x^*U^*Ux = \|x\|^2$.) The words "special" in SU(2) is due to the condition that $\det(U) = 1$. This is the right condition to make, for example because the identity matrix has this condition. (If one *understands* SU(2) then it is a small step to do the same for U(2): if U is a matrix such that $U^*U = I$ and $\det(U) = e^{i\theta}$ then $e^{-i\theta/2}U$ is in SU(2).)

We wish to relate the group SU(2) to the quaternions. The quaternions are defined as a division algebra (also known as a skey field). This is an algebraic object with all the conditions for a field, such as an Abelian group structure associated to + and a product $\cdot$ making the set into a ring, which has the property that all non-zero elements have multiplicative inverses. In particular, one of the conditions to be a ring is associativity of multiplication. (The next most interesting algebra may be the octonians, which lack even the associative property.) But $\cdot$ itself is not Abelian. So this lacks that property for the usual definition of a field.

The quaternions are well known. They can be viewed as an algebra which is a four dimensional real vector space, spanned by 1, i, j and k where 1 is the multiplicative identity, $i^2 = j^2 = k^2 = -1$, and $ij = k$, $jk = i$, $ki = j$. These conditions also imply that $ji = -k$, $kj = -i$ and $ik = -j$. For example $(ji)k = (ji)(ij) = j(i^2)j = -j^2 = 1$ by associativity. So (ji) is the multiplicative inverse of k which is evidently $-k$.

The matrices which we will choose to represent the quaternions are $\mathbb{1}$ along with

$$\mathbb{i} = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}, \quad \mathbb{j} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad \mathbb{k} = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}.$$

Another textbook for representation theory one may note is

[5] *Naive Lie Theory,* John Stillwell. Springer Verlag, Berlin, 2008.

Stillwell's book is written to accompany capstone course for undergraduates. It is not intended to be a complete introduction to the theory of Lie groups or Lie algebras. But it does present some parts of the theory nicely, such as the geometry of SU(2), which is what we are describing, now. (Stillwell also gives a nice proof of the Baker-Campbell-Hausdorff formula in the context of exponentiating Lie algebras to get Lie groups.)

Exercise 1. *Check that for these matrices* $\mathbb{i}^2 = \mathbb{j}^2 = \mathbb{k}^2 = \mathbb{i}\mathbb{j}\mathbb{k} = -\mathbb{1}$. *Explain why this is sufficient to prove that* $\mathbb{i}\mathbb{j} = \mathbb{k}$, $\mathbb{j}\mathbb{k} = \mathbb{i}$, $\mathbb{k}\mathbb{i} = \mathbb{j}$, *and* $\mathbb{j}\mathbb{i} = -\mathbb{k}$, $\mathbb{k}\mathbb{j} = -\mathbb{i}$, $\mathbb{i}\mathbb{k} = -\mathbb{j}$.

Now we want to relate this to SU(2). Suppose that we have a matrix

$$U = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}, \qquad \alpha, \beta, \gamma, \delta \in \mathbb{C},$$

satisfying $\det(U) = 1$ and $U^* U = \mathbb{1}$.

Exercise 2. *Check that these conditions are equivalent to*

$$|\alpha|^2 + |\gamma|^2 = 1, \tag{1}$$

$$|\beta|^2 + |\delta|^2 = 1, \tag{2}$$

$$\alpha\overline{\beta} = -\gamma\overline{\delta}, \tag{3}$$

$$\alpha\delta - \beta\gamma = 1. \tag{4}$$

Exercise 3. *Start with (3) and multiply both sides by β. Then use (4) and then (2) to conclude that $\alpha = \overline{\delta}$. Then do a similar calculation to prove that $\beta = -\overline{\gamma}$.*

Because of Exercise 3, we know that if U is in SU(2) then

$$U = x\mathbb{i} + y\mathbb{j} + z\mathbb{k} + t\mathbb{1}\,.$$

Moreover, from (1) or (2) then $x^2 + y^2 + z^2 + t^2 = 1$. So we have deduced that

$$\mathrm{SU}(2) = \{x\mathbb{i} + y\mathbb{j} + z\mathbb{k} + t\mathbb{1} \,:\, x^2 + y^2 + z^2 + t^2 = 1\}\,.$$

These are the unit quaternions. Moreover, as a manifold this is equivalent to the three-dimensional sphere $\mathbb{S}^3$ sitting inside $\mathbb{R}^4$ (because the quaternions are $\mathbb{R}^4$ as a vector space).

2.1 Orbital angular momentum, SO(3)

One of the discoveries of QM is that the electron has an intrinsic spin, in addition to its orbital angular momentum. It is useful to start by considering angular momentum.

Consider the Schrödinger equation for a particle in a central potential in $3d$:

$$i\hbar \frac{\partial}{\partial t}\Psi(\boldsymbol{x},t) = \left(-\frac{\hbar^2}{2m}\Delta + V(\boldsymbol{x})\right)\Psi(\boldsymbol{x},t)\,, \tag{5}$$

This PDE needs to be supplemented by an initial condition, such as $\Psi(\boldsymbol{x},0) = \Psi_0(x)$. But let us just consider (5) without initial condition, for now. Let us also define the Schrödinger operator H on $L^2(\mathbb{R}^3)$,

$$H\Psi(\boldsymbol{x}) = \left(-\frac{\hbar^2}{2m}\Delta + V(\boldsymbol{x})\right)\Psi(\boldsymbol{x})\,.$$

Of course, this operator is not really defined on all of $L^2(\mathbb{R}^3)$. But there is a dense core of smooth Ψ with good decay properties for which $H\Psi$ is well-defined. See Gunter Stolz's equation on the spectral theorem and the Schrödinger equation for a more precise explanation. For now, since we are only interested in this equation as motivation, let us assume that Ψ is sufficiently nice to be in the dense domain of H.

Moreover, let us assume that $V : \mathbb{R}^3 \to \mathbb{R}$ may actually be written as $V(\boldsymbol{x}) = v(|\boldsymbol{x}|)$ for some "nice" function $v : [0,\infty) \to \mathbb{R}$. This is what is known as a central potential. This class of Schrödinger operators is considered in Messiah [**1**] in Chapter IX, "Solution of the Schrödinger equation by separation of variables: central potential." The idea is to use spherical harmonics to solve H. (Note that some choices of central potentials are amenable to solution by other methods, for example harmonic potentials. These can be solved using the ladder operators of the harmonic oscillator, in three dimensions. This actually leads to some tools to study spherical harmonics.)

We will not delve into spherical harmonics, here. Again, our explanation is that we just want a motivation at this point. But it is a detail that a student would want to return to, the first time one studies QM.

Let SO(3) denote the group of special orthogonal matrices on $\mathbb{R}^3$. So, letting $M_3(\mathbb{R})$ denote the 3×3 real matrices,

$$\mathrm{SO}(3) \;=\; \{R \in M_3(\mathbb{R})\; :\; R^T R = I\,,\ \det(R) = 1\}\,,$$

where now we are writing I for the identity matrix (to preserve the symbol $\mathbb{1}$ for the 2×2 identity matrix on $\mathbb{C}^2$).

2.1.1 Digression into the form of matrices R in SO(3)

Every matrix $R \in \mathrm{SO}(3)$ has the following form. Choose an angle $\theta \in [0, \pi]$. If $\theta = 0$ then just take R_0 to be the identity matrix I.

If $0 < \theta < \pi$ then choose a unit vector $\boldsymbol{x} \in \mathbb{R}^3$. I.e., $\boldsymbol{x}$ is in $\mathbb{S}^2$. Using the right-hand-rule, let $R_{\boldsymbol{x},\theta}$ be the rotation about $\boldsymbol{x}$ by θ. More precisely, let $\boldsymbol{y}_1$ and $\boldsymbol{y}_1$ be any vectors in $\mathbb{R}^3$ such that $\boldsymbol{x}$, $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are ortho-normal and $\boldsymbol{x} \wedge \boldsymbol{y}_1 = \boldsymbol{y}_2$, where $\wedge$ is the cross-product. Then, in the basis $\{\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2\}$, $R_{\boldsymbol{x},\theta}$ has the form

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}\,.$$

There is one more case. But first let us note that the above definition does not depend on the choice of the pair $(\boldsymbol{y}_1, \boldsymbol{y}_2)$. It just depends on the "right-hand-rule" being satisfied for the triple $(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2)$.

Exercise 4. *(a) Suppose that $\boldsymbol{x}$, $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are three ortho-normal vectors in $\mathbb{R}^3$. Make the matrix $R \in M_3(\mathbb{R})$ such that $R = [\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2]$. Check that $R^T R = I$ and explain why this implies that* $\det(R)$ *is either* 1 *or* -1*. We will consider $(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2)$ to satisfy the "right-hand-rule" when the determinant is* 1 *(instead of* -1*).*
(b) For any $\theta \in \mathbb{R}$ consider the matrix $A_\theta \in M_2(\mathbb{R})$ given by

$$A_\theta \;=\; \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}\,.$$

For any $\theta_1 \in \mathbb{R}$, let v_{θ_1} be the vector in $\mathbb{R}^2$

$$v_{\theta_1} \;=\; \begin{bmatrix} \cos(\theta_1) \\ \sin(\theta_1) \end{bmatrix}\,.$$

Prove that the only normalized vector $w \in \mathbb{R}^2$ such that $\det(v_{\theta_1}, w) = 1$ *is $w = v_{\theta_2}$ where $\theta_2 = \theta_1$ (modulo 2π).*
(c) Prove that in the basis $\{v_{\theta_1}, v_{\theta_2}\}$, with $\theta_2 = \theta_1 + (\pi/2)$, the matrix for A_θ is the same as in the canonical basis (which is v_0, $v_{\pi/2}$). (This is an elementary property of SO(2)*.)*

The last case of θ, other than the first case 0 and the second case $(0, \pi)$, is $\theta = \pi$. If $\theta = \pi$ then choose an axis. This means choose an antipodal pair of unit vectors $\{\boldsymbol{x}, -\boldsymbol{x}\}$ where $\boldsymbol{x}$ (and hence $-\boldsymbol{x}$)

is in $\mathbb{S}^2$. Note that this means we are really choosing a point in the real projective space $\mathbb{RP}^2$. Let $R_{\pm\boldsymbol{x},\pi}$ be the rotation by π about either vector, since rotation by π is the same as rotation by $-\pi$.

This shows that SO(3) is three-dimensional as a manifold because, at least for the middle case we see that SO(3) contains an open subset homeomorphic to $(0,\pi) \times \mathbb{S}^2$.

But the first and third cases also show that SO(3) has a somewhat complicated topology. Note that $R_{\boldsymbol{x},0} = I$ for any $\boldsymbol{x}$. Therefore, $\{I\}$ should be glued to $\{R_{\boldsymbol{x},\theta} \,:\, \theta \in (0,\pi)\,,\ \boldsymbol{x} \in \mathbb{S}^2\}$ at $\theta = 0$. These two sets together give us essentially the open three-dimensional ball, where one may think of θ as the radius for $R_{\boldsymbol{x},\theta}$ and $\boldsymbol{x}$ as the point on $\mathbb{S}^2$ giving the direction. Then I is viewed as the origin.

But the last part is complicated. Since $R_{\boldsymbol{x},\pi} = R_{-\boldsymbol{x},\pi}$, this means that $\{R_{\pm\boldsymbol{x},\pi} \,:\, \{\pm\boldsymbol{x}\} \in \mathbb{RP}^2\}$ is equivalent to $\mathbb{RP}^2$, itself. Therefore, instead of, sayin, gluing the 2-sphere onto the boundary of the open 2-ball, instead we glue $\mathbb{RP}^2$. This is well-known to give the cellular decomposition of $\mathbb{RP}^3$.

One can also view $\mathbb{RP}^3$ as the quotient of $\mathbb{S}^3$, the 3-sphere sitting in real 4-space, under the identification of antipodal points. This is how we will see SO(3) descending from SU(2).

(As usual, Wikipedia is a good place to look up classical facts such as the topology of SO(3) or the cell decomposition of RP^3, as appears above.)

We will have some exercises at the end of these notes. In these exercises we will lead the reader to conclude that every matrix $R \in \mathrm{SO}(3)$ really does have the form described above.

2.1.2 Back to the Schrödinger operator

For each $R \in \mathrm{SO}(3)$, define an operator U_R on $L^2(\mathbb{R}^3)$ given by

$$U_R\Psi(\boldsymbol{x}) = \Psi(R\boldsymbol{x})\,.$$

It is easy to see that U_R is unitary since the Lebesgue measure on $\mathbb{R}^3$ is rotationally invariant, so U_R preserves the L^2-norm of Ψ. Because Δ and V are rotationally invariant, it happens that $HU_R = U_RH$.

Exercise 5. *Use the chain rule to check that*

$$\Delta(\Psi(R\boldsymbol{x}))\ =\ (\Delta\Psi)(R\boldsymbol{x})\,,$$

for orthogonal matrices R (i.e., just using $R^TR = I$ not the determinant condition).

If we had two diagonalizable matrices on a finite dimensional space A and B and if we had $AB = BA$ then that would mean that we could simultaneously diagonalize A and B.

The analogue of this for H and the set of all operators U_R, for $R \in \mathrm{SO}(3)$ is that if we have two subspaces $V, W \subset L^2(\mathbb{R}^3)$ such that V and W are sub-representations of SO(3) which are inequivalent then HV is orthogonal to W (and by self-adjointness of H this also means HW is orthogonal to V) assuming that W and V are both subspaces of $\mathrm{dom}(H)$.

The analytical details here are involved. We again refer to Gunter Stolz's lecture on self-adjointness and Schrödinger operators. In Messiah [**1**], Chapter IX, the point is to consider subspaces V spanned functions which satisfy separation of variables $\Psi(\boldsymbol{x}) = F(\|\boldsymbol{x}\|)Y(\boldsymbol{x}/\|\boldsymbol{x}\|)$, and where, moreover, the functions Y are spherical harmonics. These lead to the subrepresentations of SO(3) on $L^2(\mathbb{R}^3)$ such as V and W that are inequivalent.

We will stop here, we will not proceed to study the spherical harmonics or the representations of SO(3) acting on $L^2(\mathbb{R}^3)$.

But the motivation should be clearer, now. If we did determine the subrepresentations of SO(3), then these would be invariant spaces for H. This would aid in calculating the spectrum of H. (At an intuitive level, it would reduce the 3d Schrödinger operator to a 1d problem since F is just a function of $\|\boldsymbol{x}\|$ which is in $[0,\infty)$.)

2.2 The Lie algebra for SO(3) on $L^2(\mathbb{R}^3)$

Let us consider a one-parameter family of elements $R_t \in \mathrm{SO}(3)$ for each $t \in \mathbb{R}$. Let us take

$$R_t = \begin{bmatrix} \cos(t) & -\sin(t) & 0 \\ \sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

These are rotations about z by angle t.

Now define the differential operator $\mathcal{L}_z$ on $L^2(\mathbb{R}^3)$ by

$$\mathcal{L}_z \Psi(\boldsymbol{x}) = \frac{1}{i}\frac{d}{dt} U_{R_t}\Psi(\boldsymbol{x})\Big|_{t=0} = \frac{1}{i}\frac{d}{dt}\Psi(R_t\boldsymbol{x})\Big|_{t=0}.$$

Exercise 6. *Use the chain rule to check that* $\mathcal{L}_z = \frac{1}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$.

The reason for including the factor $1/i$ has to do with the fact that U_{R_t} is unitary for each t. In some heuristic sense, we will have $U_{R_t} = \exp(it\mathcal{L}_z)$ for these rotations about the z axis. Physicists and mathematical physicists like to write unitary operators as exponentials of i times a self-adjoint operator. See Gunter Stolz's introduction to spectral theory (the next tutorial) especially Stone's theorem. Also, see Messiah's definition in **[1]**, Chapter IX, Part 1, "Expression of the Hamiltonian in Spherical Polar Coordinates."

Of course in doing this, it is understood that i is available. This is okay for physicists and mathematical physicists, because the appropriate context of QM is for operators on complex Hilbert spaces. But in Lie group theory, especially when studying real Lie groups, one may equally well merely exponentiate an anti-Hermitian matrix.

Similarly to Exercise 6, one may see that if one replaces the z direction by the x direction or the y direction then what results is

$$\mathcal{L}_x = \frac{1}{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) \quad \text{and} \quad \mathcal{L}_y = \frac{1}{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right).$$

An important fact to notice, which is true of vector fields in general (first order differential operators) is that the product of two is not another vector field, since it is second order. But the commutator is again first order. Here the commutator is $[A,B] = AB - BA$.

Exercise 7. *Check that, thinking of x, y and z as the operators of multiplication by these factors, the following commutation relations hold:*

$$\left[\frac{\partial}{\partial x}, x\right] = 1\,, \quad \left[\frac{\partial}{\partial x}, y\right] = 0\,, \quad \left[\frac{\partial}{\partial x}, z\right] = 0\,,$$

and with similar formulas for $\partial/\partial y$ and $\partial/\partial z$. Using this, conclude that

$$[\mathcal{L}_x, \mathcal{L}_y] = i\mathcal{L}_z\,.$$

Formulas similar to the conclusion of Exercise 7 also hold,

$$[\mathcal{L}_y, \mathcal{L}_z] = i\mathcal{L}_x \quad \text{and} \quad [\mathcal{L}_z, \mathcal{L}_x] = i\mathcal{L}_y\,.$$

The complexified version of the Lie algebra for SO(3) consists of three generators L_x, L_y and L_z (i.e., these are the basis elements for the complex vector space which comprises the complex Lie algebra) along with the relations: $[L_x, L_y] = iL_z$, $[L_y, L_z] = iL_x$ and $[L_z, L_x] = iL_y$.

2.3 The Lie algebra for SU(2) and the relation to SO(3)

Recall that SU(2) may be expressed as the set of all matrices of the form

$$U = x\mathbb{i} + y\mathbb{j} + z\mathbb{k} + t\mathbb{1}\,, \quad (x, y, z, t) \in \mathbb{S}^3\,.$$

There is a 2-to-1 homomorphism of SU(2) onto SO(3), also called a double covering.

To present this, first let us denote the set of all quaternions as

$$\mathbb{H} = \{x\mathbb{i} + y\mathbb{j} + z\mathbb{k} + t\mathbb{1} \,: x, y, z, t \in \mathbb{R}\}\,.$$

Also, let us denote the set of "pure imaginary" quaternions

$$\mathbb{H}_i = \{x\mathbb{i} + y\mathbb{j} + z\mathbb{k} \,: x, y, z, \in \mathbb{R}\}\,.$$

This is equivalent to $\mathbb{R}^3$. Given $U \in \mathrm{SU}(2)$ we define a mapping $\Phi_U : \mathbb{H} \to \mathbb{H}$ as $\Phi_U(A) = UAU^*$ for each $A \in \mathbb{H}$. It is obvious that $\Phi_U(\mathbb{1}) = \mathbb{1}$.

Exercise 8. *Use unitarity to prove that since $\{t\mathbb{1} \,: t \in \mathbb{R}\}$ is invariant for Φ_U, so is $\mathbb{H}_i$.*

Because of the result in Exercise 8, we may define $\phi_U : \mathbb{H}_i \to \mathbb{H}_i$ as the restriction of Φ_U to $\mathbb{H}_i$. This is a linear mapping, and by multiplicativity of the norm, and unitarity of U, one may see that ϕ_U is an isometry. Since SU(2) is connected and since $\phi_{\mathbb{1}} = I$, it follows that the determinant of ϕ_U is 1 for all U in SU(2), as opposed to the alternative -1.

Therefore $U \mapsto \phi_U$ is a homomorphism of SU(2) into SO(3). To see that it is onto, note the following. Given any $\boldsymbol{x} \in \mathbb{R}^3$, let $Q_{\boldsymbol{x}}$ denote the pure imaginary vector

$$Q_{\boldsymbol{x}} = x\mathbb{i} + y\mathbb{j} + z\mathbb{k} .$$

Then if $\boldsymbol{x}$ and $\boldsymbol{x}'$ are two vectors in $\mathbb{R}^3$, we have

$$Q_{\boldsymbol{x}} Q_{\boldsymbol{x}'} = Q_{\boldsymbol{x}\wedge\boldsymbol{x}'} - (\boldsymbol{x}\cdot\boldsymbol{x}')\mathbb{1} . \tag{6}$$

Exercise 9. *Prove (6).*

Now, suppose we choose any unit vector $\boldsymbol{x}$ in $\mathbb{R}^3$. Then $Q_{\boldsymbol{x}}$ is both a unit quaternion and a pure imaginary quaternion. For each $\theta \in \mathbb{R}$ define $U_{\boldsymbol{x},\theta} \in \mathrm{SU}(2)$ as

$$U_{\boldsymbol{x},\theta} = \sin(\theta) Q_{\boldsymbol{x}} + \cos(\theta)\mathbb{1} .$$

Let $\boldsymbol{x}'$ be any unit vector in $\mathbb{R}^3$ orthogonal to $\boldsymbol{x}$. Then, letting $\boldsymbol{x}'' = \boldsymbol{x}\wedge\boldsymbol{x}'$, we have $Q_{\boldsymbol{x}}Q_{\boldsymbol{x}'} = Q_{\boldsymbol{x}''}$, and cyclic permutations of this relation hold, as well. Note that $U^*_{\boldsymbol{x},\theta} = U_{\boldsymbol{x},-\theta}$. It follows that

$$\phi_{U_{\boldsymbol{x},\theta}}(Q_{\boldsymbol{x}}) = Q_{\boldsymbol{x}} ,$$

by commutativity.

Exercise 10. *Check that*

$$U_{\boldsymbol{x},\theta} Q_{\boldsymbol{x}'} = \sin(\theta) Q_{\boldsymbol{x}'} + \cos(\theta) Q_{\boldsymbol{x}''} .$$

Then prove that

$$\phi_{U_{\boldsymbol{x},\theta}}(Q_{\boldsymbol{x}'}) = [\cos^2(\theta) - \sin^2(\theta)] Q_{\boldsymbol{x}'} + 2\sin(\theta)\cos(\theta) Q_{\boldsymbol{x}''} .$$

Finally, explain what needs to be done, and do it, in order to prove

$$\phi_{U_{\boldsymbol{x},\theta}}(Q_{\boldsymbol{y}}) = Q_{\boldsymbol{y}'} , \quad \text{for } \boldsymbol{y}' = R_{\boldsymbol{x},2\theta}\boldsymbol{y} ,$$

for each $\boldsymbol{y} \in \mathbb{R}^3$.

This proves that $U \mapsto \phi_U$ is onto because by Subsection 2.1.1, we know that every element of SO(3) is of the form $R_{\boldsymbol{x},\theta}$ for some $\boldsymbol{x} \in \mathbb{R}^3$ and $\theta \in [0,\pi]$. (If $\theta = 0$ then $\boldsymbol{x}$ is irrelevant, and if $\theta = \pi$ then we may choose between $\boldsymbol{x}$ and $-\boldsymbol{x}$.) We may obtain this by taking $U_{\boldsymbol{x},\theta/2}$. Actually every U has the form $U_{\boldsymbol{x},\theta}$. We see that ϕ_U is the identity if and only if U commutes with all $Q \in \mathbb{H}_i$. But then U commutes with all $Q \in \mathbb{H}$, because everything commutes with $\mathbb{1}$. The center of $\mathbb{H}$ is $\{+1,-1\}$. Therefore, the kernel of $U \mapsto \phi_U$ is $\{+1,-1\}$. Since this is a homomorphism, this proves that the mapping is 2-to-1.

At this point it is obvious that SO(3) is the real projective space $\mathbb{RP}^3$ since it is the quotient of $\mathrm{SU}(2) \cong \mathbb{S}^3$ by the identification of antipodes $U \sim -U$. This is the definition of real projective space.

The important implication of this will be that there are roughly half as many representations of SO(3) and there are of SU(2). In the physics language this corresponds to the fact that irreducible representations of SU(2) are also representations of SO(3) if and only if the "spin" is integer, as opposed to half-odd-integer. We will explain the irreps next, and then this will be clear.

3 Irreducible representations of SU(2)

A finite dimensional representation of SU(2) is a (complex) vector space V, along with operators S^x, S^y and S^z on V satisfying the commutation relations

$$[S^x, S^y] = iS^z\,, \quad [S^y, S^z] = iS^x \text{ and } [S^x, S^x] = iS^y\,.$$

(We will only consider complex representations.) The representation is said to be unitary if V has a Hilbert space structure and S^x, S^y and S^z are all Hermitian operators on V.

If $W \subseteq V$ is a subspace which is invariant for S^x, S^y and S^z, i.e., each of these three operators maps W back into itself, then W is called a sub-representation.

Exercise 11. *Suppose that V is a unitary representation of* SU(2) *and that $W \subseteq V$ is a subrepresentation. Since W is a subrepresentation, $S^x w$, $S^y w$ and $S^z w$ are all in W for each $w \in W$. Let $W^\perp$ be the orthogonal complement of W. Using this, prove that*

$$\langle u, S^x v\rangle = \langle u, S^y v\rangle = \langle u, S^z v\rangle = 0\,,$$

for each $u \in W$ and $v \in W^\perp$. Conclude that $W^\perp$ is also a subrepresentation.

Note that Exercise 11 is similar to Exercise 8. The idea of taking orthogonal complements for unitary representations is often useful.

A representation is called irreducible if there are no non-trivial subrepresentations. I.e., a representation V is irreducible if whenever $W \subseteq V$ is a sub-representation this requires W equals V or $\{0\}$.

An important fact from the general theory of representations of finite and compact groups is "complete reducibility." Let us list some of the elements quickly, here.

- Firstly, if V is any finite-dimensional representation, then the vector space may be equipped with an inner-product such that V becomes unitary.
- Then if W is a sub-representation then so is the orthogonal complement $W^\perp$, by Exercise 11.
- Because of this, if V has a non-trivial sub-representation, W such that W is neither 0 nor V, then $V = W \oplus W^\perp$, where $W^\perp$ is also non-trivial, and both are sub-representations.
- Finally, this implies, by an inductive argument that every finite dimensional representation V may be written as $V = W_1 \oplus \cdots \oplus W_n$, for some $n \leq \dim(V)$, where $W_1, \dots, W_n$ are all irreducible representations.

A great reference for this type of thing is [2].

If V and W are both representations of SU(2), then $T : V \to W$ is an intertwiner if T is a linear mapping and T respects the structure of the representations. Here this means that for the three spin matrices on V, call them S_V^x, S_V^y and S_V^z, and for the three spin matrices on W, call them S_W^x, S_W^y and S_w^z, one has the intertwining relations

$$TS_V^x = S_W^x T\,, \quad TS_V^y = S_W^y T \text{ and } TS_V^z = S_W^z T\,.$$

Another important result from representation theory is Schur's lemma, which implies that if V and W are each irreducible, then T may only be the zero-mapping or an isomorphism, the latter only if V and W are equivalent representations (meaning that there is an intertwiner between them which is an isomorphism). Again, [2] is a good reference.

3.1 Highest weight vectors

We will now characterize all finite dimensional irreducible representations of SU(2). We will take the perspective of most QM textbooks, and *deduce* what they are. The key is to look for highest weight vectors.

Suppose that V is a finite-dimensional, unitary representation of SU(2). So $V \cong \mathbb{C}^n$ for some n, and there are three matrices S^x, S^y and S^z, satisfying the commutation relations

$$[S^x, S^y] = iS^z\,, \quad [S^y, S^z] = iS^x \text{ and } [S^z, S^x] = iS^y\,.$$

Moreover, because the representation is unitary, S^x, S^y and S^z are all Hermitian: $(S^x)^* = S^x$, $(S^y)^* = S^y$ and $(S^z)^* = S^z$. In particular, this means that they are all diagonalizable and the eigenvalues are all real. Let j be the largest eigenvalue of S^z, and let Ψ_j be a normalized vector such that

$$S^z \Psi_j = j\Psi_j\,.$$

We are going to call Ψ_j a "highest weight vector."

The key ingredients to be introduced now are the spin-raising and spin-lowering operators, S^+ and S^-. These are defined as $S^+ = S^x + iS^y$ and $S^- = S^x - iS^y$, for short.

Exercise 12. *Prove that*
(a) $(S^+)^* = S^-$ *(and hence* $(S^-)^* = S^+$*);*
(b) $S^x = (S^+ + S^-)/2$ *and* $S^y = (S^+ - S^-)/(2i)$*;*
(c) $(S^x)^2 + (S^y)^2 = (S^+S^- + S^-S^+)/2.$

The Lie group SU(2) is three-dimensional, meaning that we need 3 generators for the Lie algebra. These are S^x, S^y and S^z. But now that we have a unitary representation, we may instead consider our three main operators as S^z, S^+ and S^-. (One may think of this at the level of the Lie algebra, by taking the complexified Lie algebra, where one tensors with $\mathbb{C}$ over $\mathbb{R}$, in order to allow ourselves to take complex linear combinations of the three generators to get $S^\pm = S^x \pm iS^y$.) So let us calculate the commutation relations.

$$\begin{aligned} [S^z, S^{\pm}] &= [S^z, S^x \pm i S^y] \\ &= [S^z, S^x] \pm i [S^z, S^y] \\ &= [S^z, S^x] - (\pm i)[S^y, S^z] \\ &= i S^y - (\pm i) i S^x \\ &= i S^y \pm S^x \\ &= \pm S^{\pm}\,. \end{aligned}$$

Exercise 13. *Prove that* $[S^+, S^-] = 2S^z$.

Because of this, in particular, note that

$$S^z S^+ \Psi_j \;=\; (S^+ S^z + [S^z, S^+])\Psi_j \;=\; S^+ S^z \Psi_j + S^+ \Psi_j \;=\; (j+1) S^+ \Psi_j\,.$$

But note that, by our assumption, j was the largest eigenvalue of S^z. If $S^+\Psi_j$ were any vector other than 0, then that would imply that $S^+\Psi_j$ were an eigenvector for S^z for the eigenvalue $j+1$, contrary to our assumption. Therefore, we have given a contradiction proof that $S^+\Psi_j = 0$. This helps to explain why Ψ_j is called a "highest weight" vector.

If Φ is a general eigenvector of S^z with $S^z\Phi = m\Phi$, then we may call m the "weight." (The name may be better motivated by the more general theory of representations of semi-simple Lie groups, in particular, where the representation theory of SU(2) [or SL(2)] fits in as a main tool, finding representations of this small group inside representations of larger groups and piecing the representations of SU(2) together.) Then S^+ will raise the weight because

$$S^z \Phi \;=\; m\Phi \;\Rightarrow\; S^z(S^+\Phi) \;=\; (m+1) S^+ \Phi\,, \tag{7}$$

and S^- will lower the weight because

$$S^z \Phi \;=\; m\Phi \;\Rightarrow\; S^z(S^-\Phi) \;=\; (m-1) S^- \Phi\,. \tag{8}$$

But if Ψ_j already has the highest weight possible, then this must mean that $S^+\Psi_j$ annihilates Ψ_j.

Next, consider the set of vectors $\{\Psi_j, S^-\Psi_j, (S^-)^2\Psi_j, \dots\}$. These vectors cannot all be linearly independent, because V is only finite dimensional. On the other hand, by (8) we would know that all of these vectors would be eigenvectors for S^z with different eigenvalues $\{j, j-1, j-2, \dots\}$, unless some of them are zero. If none of them were zero, then that would necessitate that they would all be linearly independent, because eigenvectors of a matrix with distinct eigenvalues are always linearly indpendent. Therefore, there is some $k \in \{1, 2, \dots\}$ such that $(S^-)^\ell \Psi_j = 0$ for all $\ell \geq k$. We will take k to be the smallest possible integer. Then it will turn out that $\{\Psi_j, S^-\Psi_j, \dots, (S^-)^{k-1}\Psi_j\}$ span a k-dimensional sub-representation of V.

Exercise 14. *For $m \in \{j, j-1, \dots, j-k+1\}$, let $\Psi_m = (S^-)^{j-m}\Psi_j / \|(S^-)^{j-m}\Psi_j\|$. Prove that*

$$S^-\Psi_m \;=\; \sqrt{j(j+1)-m(m-1)}\,\Psi_{m-1}\,,$$

interpreting $0\Psi_{j-k}$ as just 0 (since Ψ_{j-k} does not really exist, but the coefficient is zero anyway). Do this by induction, as follows.
(a) Starting with the fact that $S^+\Psi_j = 0$ and $S^z\Psi_j = j\Psi^j$, use the commutation relations $S^+S^- - S^-S^+ = 2S^z$ to conclude that

$$\langle S^-\Psi_j, S^-\Psi_j\rangle - \langle S^+\Psi_j, S^+\Psi_j\rangle \;=\; \langle \Psi_j, 2S^z\Psi_j\rangle \;=\; 2j\|\Psi_j\|^2 \;=\; 2j\,,$$

using the assumption that Ψ_j was normalized. Explain why this implies that $j \geq 0$ and also why this implies the initial step $S^-\Psi_j = \sqrt{2j}\,\Psi_{j-1}$.
(b) Suppose that you know that $S^-\Psi_m = \sqrt{j(j+1)-m(m-1)}\,\Psi_{m-1}$ for some m. Use unitarity to show that $S^+\Psi_{m-1}$ is orthogonal to $(\mathrm{span}(\Psi_m))^\perp$. Hence, $S^+\Psi_{m-1}$ equals $\lambda\Psi_m$ for some $\lambda \in \mathbb{C}$.
(c) Prove that if $S^-\Psi_m = \sqrt{j(j+1)-m(m-1)}\,\Psi_{m-1}$ for some m, then

$$S^+\Psi_{m-1} \;=\; \sqrt{j(j+1)-m(m-1)}\,\Psi_m.$$

(d) Suppose that for some m you know $S^+\Psi_m = \sqrt{j(j+1)-m(m+1)}\,\Psi_{m+1}$. You already know $S^z\Psi_m = m\Psi_m$. Using this, prove that $S^-\Psi_m = \sqrt{j(j+1)-m(m-1)}\,\Psi_{m-1}$. This is the induction step.

Because of Exercise 14, we see in particular that $S^-\Psi_{-j} = \sqrt{j(j+1)-(-j)(-j-1)}\,\Psi_{-j-1}$, which is $0\Psi_{-j-1}$, which is interpreted as just 0 (because there is no actual vector Ψ_{-j-1}, but this does not really matter because the coefficient is 0 anyway). So we see that $k \leq 2j+1$ since unless Ψ_{-j} also does not exist we have $\Psi_{-j} = (S^-)^{2j}\Psi_j/C$ for some constant C, so this means $(S^-)^{2j+1}\Psi_j = 0$.

But in fact, k is not less than $2j+1$. To see this, note that for $m \in \{j, j-1, \dots, -j+1\}$ we have

$$S^-\Psi_m \;=\; \sqrt{j(j+1)-m(m-1)}\,\Psi_{m-1}\,,$$

and this is only zero if $\sqrt{j(j+1)-m(m-1)} = 0$. But that requires $j(j+1)-m(m-1) = 0$, and by solving the quadratic equation that requires $m = -j$ or $2j+1$. But as m must be less than j, and as $j \geq 0$ by part (a) above, this means $m = -j$ is the only possible value such that $S^-\Psi_m = 0$ among the weights $\{j, j-1, \dots, -j\}$. So really it is true that $k = 2j+1$. This is the dimension of the representation spanned by $\{\Psi_j, \Psi_{j-1}, \dots, \Psi_{-j}\}$.

Let us quickly note why this is a sub-representation. For this to be a sub-representation we require this subspace to be mapped back into itself under each of the three matrices S^x, S^y and S^z. We already know $S^z\Psi_m = m\Psi_m$ for each $m \in \{j, j-1, \dots, -j\}$. We also know

$$S^+\Psi_m \;=\; \sqrt{j(j+1)-m(m+1)}\,\Psi_{m+1} \quad\text{and}\quad S^-\Psi_m \;=\; \sqrt{j(j+1)-m(m-1)}\,\Psi_{m-1}\,,$$

with the interpretation that $0\Psi_{j+1}$ and $0\Psi_{-j-1}$ are both zero. But that means that this subspace is also mapped into itself by S^+ and S^-. Moreover, by Exercise 12, we know that S^x and S^y are linear combinations of S^+ and S^-. Therefore, this is an invariant subspace also for S^x and S^y.

Therefore, this is a sub-representation. It is not zero because it always contains the non-zero vector Ψ_j. So, since we assumed that V is irreducible, it must be the entire sub-representation. Note that we call two representations "equivalent" if there is an intertwiner between them which is an isomorphism. We classify irreducible representations up to equivalence.

3.2 Conclusion

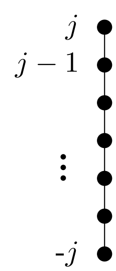
Therefore, what we have just deduced is that the irreducible representations may be listed as $\mathcal{D}^{(j)}$ for $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$, (so that $k = 2j+1$ is an allowable dimension $1, 2, 3, \dots$), where $\mathcal{D}^{(j)}$ may be written with an orthonormal basis $\Psi_j, \Psi_{j-1}, \dots, \Psi_{-j}$ satisfying

$$\begin{aligned} S^z\Psi_m &= m\Psi_m\,, \quad \text{for } m \in \{j, j-1, \dots, -j\}, \\ S^+\Psi_m &= \sqrt{j(j+1) - m(m+1)}\,\Psi_{m+1}\,, \quad \text{for } m \in \{j-1, j-2, \dots, -j\} \text{ and } S^+\Psi_j = 0, \\ S^-\Psi_m &= \sqrt{j(j+1) - m(m-1)}\,\Psi_{m-1}\,, \quad \text{for } m \in \{j,, j-1, \dots, -j+1\} \text{ and } S^-\Psi_{-j} = 0, \end{aligned}$$

and $S^x = (S^+ + S^-)/2$ and $S^y = (S^+ - S^-)/(2i)$.

4 Tensor products of representations

A good application of the use of highest weight vectors for representations of SU(2) is the decomposition of the tensor product of two irreducible representations into a direct sum of irreducible representations. A good picture to think of for the representation $\mathcal{D}^{(j)}$ is a vertical chain



representing the possible eigenvalues m of S^z in $\mathcal{D}^{(j)}$. Of course, there is also an eigenvector associated to each m-value, Ψ_m. We also need to discuss how one forms a representation of SU(2) on the tensor product of two representations.

4.1 Brief review of tensor products

Recall that if V and W are two vector spaces then $V \otimes W$ is another vector space. More precisely, if $\dim(V) = n$ and $\dim(W) = m$ and if $V = \text{span}\{v_1, \dots, v_n\}$ and $W = \text{span}\{w_1, \dots, w_m\}$ for some vectors $v_1, \dots, v_n \in V$ and $w_1, \dots, w_m \in W$, then we may define $V \otimes W$ as the span of nm linearly independent basis vectors

$$\{v_i \otimes w_j \, : \, i \in \{1, \ldots, n\}\,,\ j \in \{1, \ldots, m\}\}\,.$$

At this level the notation $\otimes$ for vectors is just a notation: we define mn new vectors called $v_1 \otimes w_1, \ldots, v_1 \otimes w_m, \ldots, v_n \otimes w_1, \ldots, v_n \otimes w_m\}$ (in a vector space called $V \otimes W$).

But then, once these are defined, we may actually define a new operation with algebraic meaning $\otimes : V \times W \to V \otimes W$, wherein, for

$$v \,=\, \sum_{j=1}^{n} c_j v_j \quad \text{and} \quad w \,=\, \sum_{k=1}^{m} d_k w_k\,,$$

for $c_1, \ldots, c_n, d_1, \ldots, d_m \in \mathbb{C}$, we have the formula

$$v \otimes w \,=\, \sum_{j,k=1}^{n} (c_j d_k) v_j \otimes w_k\,.$$

Moreover, it is seen that for $v, v' \in V$, $w, w' \in W$ and $c, d \in \mathbb{C}$,

$$(v + cv') \otimes (w + dw') \,=\, v \otimes w + c(v' \otimes w) + d(v \otimes w') + cd(v' \otimes w')\,,$$

which expresses bilinearity of $\otimes$. There is a universal property for the tensor product. For any other bilinear mapping $h : V \times W \to Z$, for some vector space Z, there exists a unique mapping $h' : V \otimes W \to Z$ such that $h(v, w) = h'(v \otimes w)$. See Wikipedia for more details (from which the notation of the last sentence was adopted):

`http://en.wikipedia.org/wiki/Tensor_product`

Note that one particular case is if we have linear operators $A : V \to V$ and $B : W \to W$. Then we may define $A \otimes B$ to be an operator from $V \otimes W$ to itself, defined as

$$(A \otimes B)(e_i \otimes f_j) \,=\, (Ae_i) \otimes (Bf_j)\,,$$

where on the left-hand-side the $\otimes$ is mainly notation for expressing the basis vector $e_i \otimes f_j$, and on the right hand side it is the bilinear operation we described before (since Ae_i is a general vector in V and Bf_j is a general vector in W, not necessarily basis vectors).

One can also check that $(A, B) \mapsto A \otimes B$ is also bilinear. In fact it is the tensor product of the vector spaces $\mathrm{End}(V)$ and $\mathrm{End}(W)$, where the endomorphism algebra of a vector space is the set of all linear mappings of that space back into itself. Moreover, it is easy to see that with this definition, writing $\mathbb{1}_V$ and $\mathbb{1}_W$ for the identities on A and B, respectively, we have

$$A \otimes B \,=\, (A \otimes \mathbb{1}_W)(\mathbb{1}_V \otimes B) \,=\, (\mathbb{1}_V \otimes B)(A \otimes \mathbb{1}_W)\,.$$

There is frequently an identification of $\mathrm{End}(V)$ with a subalgebra of $\mathrm{End}(V \otimes W)$ obtained by mapping A to $A \otimes \mathbb{1}_W$, and similarly an identification of $\mathrm{End}(W)$ with another subalgebra of $\mathrm{End}(V \otimes W)$ obtained by mapping B to $\mathbb{1}_V \otimes B$. Then the formula above implies that the two subalgebras commute with each other. (The commutator of the two subalgebras is zero.) This will be used frequently in the theory of quantum spin systems. It will be explained in other lectures more methodically.

4.2 The tensor product of two finite dimensional representations

Now we may finally define the representation structure (also known as module structure) on $V \otimes W$ in the case that V and W are each SU(2) representations. Suppose that the three spin operators on V are denotes S_V^x, S_V^y and S_V^z, and suppose a similar notation for W. Then we may define three spin matrices on $V \otimes W$ which we denote $S_{V\otimes W}^x$, $S_{V\otimes W}^y$ and $S_{V\otimes W}^z$, defined as

$$\begin{aligned} S_{V\otimes W}^x &= S_V^x \otimes \mathbb{1}_W + \mathbb{1}_V \otimes S_W^x\,, \\ S_{V\otimes W}^y &= S_V^y \otimes \mathbb{1}_W + \mathbb{1}_V \otimes S_W^y\,, \\ S_{V\otimes W}^z &= S_V^z \otimes \mathbb{1}_W + \mathbb{1}_V \otimes S_W^z\,. \end{aligned}$$

It remains to check the commutation relations, but note that since the subalgebra $\{A \otimes \mathbb{1}_W \,:\, A \in \mathrm{End}(V)\}$ commutes with the subalgebra $\{\mathbb{1}_V \otimes B \,:\, B \in \mathrm{End}(B)\}$, we have, for instance

$$\begin{aligned} [S_{V\otimes W}^x, S_{V\otimes W}^y] &= [S_V^x \otimes \mathbb{1}_W + \mathbb{1}_V \otimes S_W^x, S_V^y \otimes \mathbb{1}_W + \mathbb{1}_V \otimes S_W^y] \\ &= [S_V^x, S_V^y] \otimes \mathbb{1}_W + \mathbb{1}_V \otimes [S_W^x, S_W^y] \\ &= iS_V^z \otimes \mathbb{1}_W + \mathbb{1}_V \otimes (iS_W^z) \\ &= iS_{V\otimes W}^z\,. \end{aligned}$$

So, for example, this example of a commutation relation works out. It is easy to see by similar arguments that the other commutations work out, as well.

You can also motivate this choice by considering Schŕodinger operators with SO(3) symmetry. Let us do this very quickly.

4.3 Motivation by analogy with SO(3)

A typical two-body Hamiltonian, say for two electrons in $\mathbb{R}^3$ may be written as follows. Say $\boldsymbol{x}_1 = (x_1, y_1, z_1)$ and $\boldsymbol{x}_2 = (x_2, y_2, z_2)$. Then the Hamiltonian is an unbounded operator on $L^2(\mathbb{R}^3 \times \mathbb{R}^3)$,

$$H\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{k=1}^{2} \Big(-\frac{1}{2}\nabla_{\boldsymbol{x}_k}^2 + V(\boldsymbol{x}_k)\Big)\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) + W(\boldsymbol{x}_1 - \boldsymbol{x}_2)\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2)\,,$$

where $V : \mathbb{R}^3 \to \mathbb{R}$ is called a "central potential" if it is rotationally invariant, and W is an isotropic interaction if it is rotationally invariant.

Under these conditions, for any $R \in \mathrm{SO}(3)$, we have that, if we define $\Phi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \Psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$, then

$$H\Phi(\boldsymbol{x}_1, \boldsymbol{x}_2) \;=\; (H\Psi)(R\boldsymbol{x}_1, R\boldsymbol{x}_2)\,,$$

where $H\Psi$ is the function obtained by applying H to $\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$, but $(H\Psi)(R\boldsymbol{x}_1, R\boldsymbol{x}_2)$ means that you should apply this function to the point $(R\boldsymbol{x}_1, R\boldsymbol{x}_2)$. This motivates the definition of a unitary operator $U^{\{1,2\}}(R) : L^2(\mathbb{R}^3 \times \mathbb{R}^3) \to L^2(\mathbb{R}^3 \times \mathbb{R}^3)$ defined as

$$U_{\{1,2\}}(R)\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) \;=\; \Psi(R\boldsymbol{x}_1, R\boldsymbol{x}_2)\,.$$

Note that it is essential that you rotate $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ by the same rotation R in order to take advantage of the symmetry of W.

Now one could define the 1-parameter subgroup R_z^t, for $t \in \mathbb{R}$, of rotations by angle t about z given by $R_z^t \boldsymbol{e}_3 = \boldsymbol{e}_3$, $R_z^t \boldsymbol{e}_1 = \cos(t)\boldsymbol{e}_1 + \sin(t)\boldsymbol{e}_2$, $R_z^t \boldsymbol{e}_2 = \cos(t)\boldsymbol{e}_2 - \sin(t)\boldsymbol{e}_1$. (Note that the notation is somewhat consistent because $R_z^t R_z^s = R_z^{t+s}$, although of course unique roots are not well-defined.) Then we may as usual define

$$\begin{aligned}
\mathcal{L}^z_{\{1,2\}}\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2) \;&=\; \frac{1}{i}\,\frac{d}{dt}\,U_{\{1,2\}}(R_z^t)\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2)\bigg|_{t=0} \\
&=\; \frac{1}{i}\left(x_1\,\frac{\partial}{\partial y_1} - y_1\,\frac{\partial}{\partial x_1} + x_2\,\frac{\partial}{\partial y_2} - y_2\,\frac{\partial}{\partial x_2}\right)\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2)\,.
\end{aligned}$$

We may identify $L^2(\mathbb{R}^3 \times \mathbb{R}^3)$ as the tensor product $L^2(\mathbb{R}^3) \otimes L^2(\mathbb{R}^3)$, wherein the identification is, for $\Psi_1, \Psi_2 \in L^2(\mathbb{R}^3)$, we may define

$$(\Psi_1 \otimes \Psi_2)(\boldsymbol{x}_1, \boldsymbol{x}_2) \;=\; \Psi_1(\boldsymbol{x}_1)\Psi_2(\boldsymbol{x}_2)\,.$$

It is easy to see that this is an isomorphism of the vector spaces. Moreover, with this representation, one may view $\mathcal{L}^z_{\{1,2\}}$ as an (unbounded) operator on $L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3)$ where

$$\mathcal{L}^z_{\{1,2\}} \;=\; \mathcal{L}^z \otimes \mathbb{1} + \mathbb{1} \otimes \mathcal{L}^z\,.$$

Or, defining $\mathcal{L}^z_1 = \mathcal{L}^z \otimes \mathbb{1}$ and $\mathcal{L}^z_2 = \mathbb{1} \otimes \mathcal{L}^z$. So $\mathcal{L}^z_{\{1,2\}} = \mathcal{L}^z_1 + \mathcal{L}^z_2$. This is the same as the formalism for the representation of SU(2) on a tensor product of two representations.

4.4 Back to tensor products of finite dimensional representations of SU(2)

Now, let us consider the tensor product of two irreducible representations,

$$V \;=\; \mathcal{D}^{(j)} \otimes \mathcal{D}^{(j')}\,.$$

We will write the spin operators on the tensor product as

$$S^x_1 \;=\; S^x \otimes \mathbb{1}\,, \quad S^x_2 \;=\; \mathbb{1} \otimes S^x \;\text{ and }\; S^x_{\{1,2\}} \;=\; S^x_1 + S^x_2\,,$$

and similarly for y and z. Then $S^x_{\{1,2\}}$, $S^y_{\{1,2\}}$ and $S^z_{\{1,2\}}$ are three spin matrices defining a representation of SU(2) on $\mathcal{D}^{(j)} \otimes \mathcal{D}^{(j')}$. This representation is not irreducible. But it may be written as a direct sum decomposition of irreducible representations. We will prove the following:

Theorem. *For any* $j, j' \in \{0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots\}$, *we have*

$$\mathcal{D}^{(j)} \otimes \mathcal{D}^{(j')} \cong \bigoplus_{J=|j-j'|}^{j+j'} \mathcal{D}^{(J)} ,$$

where the sum on ℓ *is understood to go in steps of* 1.

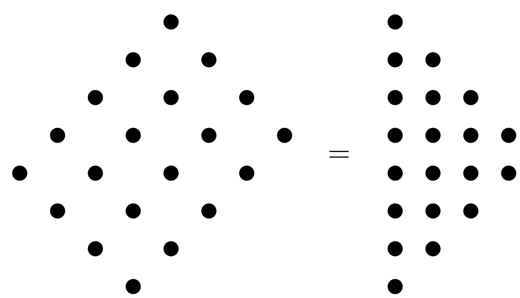
As a warm-up we recommend the following exercise.

Exercise 15. *Prove that*

$$(2j+1)(2j'+1) = \sum_{J=|j-j'|}^{j+j'} (2J+1) ,$$

which shows that the dimensions work out.

In pictures one way to understand this is to consider a picture

$4 \cdot 5 =$  $= 8+6+4+2\,.$

The particular formula could be rewritten as

$$\Big(2 \cdot \frac{3}{2} + 1\Big)(2 \cdot 2 + 1) = \Big(2 \cdot \frac{7}{2} + 1\Big) + \Big(2 \cdot \frac{5}{2} + 1\Big) + \Big(2 \cdot \frac{3}{2} + 1\Big) + \Big(2 \cdot \frac{1}{2} + 1\Big) .$$

In other words, writing d_j for $2j+1$ we have verified in picture the formula

$$d_{3/2} \cdot d_2 = \sum_{J=1/2}^{7/2} d_J ,$$

which is a special case of Exercise 15 since $(3/2) + 2 = 7/2$ and $|(3/2) - 2| = 1/2$.

The idea of the proof of the theorem will be somewhat inductive. We will construct the irreducible representations in $\mathcal{D}^{(j)} \otimes \mathcal{D}^{(j')}$. The key to this will be to identify the highest weight vectors. Let us say that $\psi \in V$ is a highest weight vector of weight m if $S^z\psi = m\psi_z$ and $S^+\psi = 0$. In the present context, where $V = \mathcal{D}^{(j)} \otimes \mathcal{D}^{(j')}$, this means we seek vectors such that $S^z_{\{1,2\}}\psi = m\psi$, and, defining $S^{\pm}_{\{1,2\}} = S^x_{\{1,2\}} \pm iS^y_{\{1,2\}}$, we also require $S^+_{\{1,2\}}\psi = 0$.

Since $S^-_{\{1,2\}}$ is the adjoint of $S^+_{\{1,2\}}$, this means that ψ is a highest weight vector of weight m if $S^z\psi = m\psi$ and if ψ is orthogonal to the range of $S^-_{\{1,2\}}$. More precisely, let us define $V_m = \{\psi \in V :$

$S^z_{\{1,2\}}\psi = m\psi\}$. Then we require that $\psi \in V_m$ and $\psi \perp S^-_{\{1,2\}}\phi$ for every $\phi \in V$. But since $S^-_{\{1,2\}}$ maps V_{m+1} into V_m, this really means that ψ is a highest weight vector of weight m if $S^z_{\{1,2\}}\psi \in V_m$ and ψ is orthogonal to $S^-_{\{1,2\}}\phi$ for every $\phi \in V_{m+1}$. This gives us an inductive method for calculating the highest weight vectors.

Exercise 16. *Let $\psi_{j,m}$, for $m = j, j-1, \dots, -j$ be the vectors in $\mathcal{D}^{(j)}$, which are orthonormal such that $S^z\psi_{j,m} = m\psi_{j,m}$ and $S^{\pm}\psi_{j,m} = \sqrt{j(j+1) - m(m\pm 1)}\,\psi_{j,m\pm 1}$. Make a similar definition for vectors $\psi_{j',m'}$ in $\mathcal{D}^{(j')}$.*
(a) Prove that an orthonormal basis for the subspace V_M is

$$\{\psi_{j,m} \otimes \psi_{j',m'} \,:\, m \in \{j, j-1, \dots, -j\}\,,\ m' \in \{j', j'-1, \dots, -j'\} \text{ and } m + m' = M\}\,.$$

(b) Prove that $S^-_{\{1,2\}}V_M$ is spanned by the vectors

$$\sqrt{j(j+1) - m(m-1)}\,\psi_{j,m-1} \otimes \psi_{j',m'} + \sqrt{j'(j'+1) - m'(m'-1)}\,\psi_{j,m} \otimes \psi_{j',m'-1}\,,$$

for the set of all pairs (m, m') such that $m \in \{j, \dots, -j\}$, $m' \in \{j', \dots, -j'\}$ and $m + m' = M$.
(c) Prove that $V_{M-1} \cap (S^-_{\{1,2\}}V_M)^\perp$ contains a non-zero vector if and only if $\dim(V_{M-1}) > \dim(V_M)$, in which case the subspace is 1-dimensional.
(d) Prove that $\dim(V_M) = j + j' + 1 - M$ as long as (M is in $\{j+j', j+j'-1, \dots, -j-j'\}$ and) $j' + j' - M \le \min\{2j, 2j'\}$ (so that both $\psi_{j,M-j'}$ and $\psi_{j',M-j}$ exist) while $\dim(V_M) = \min\{2j, 2j'\} + 1$ if $0 \le M \le |j - j'|$.

It is common to denote a normalized vector in $V_J \cap (S^-_{\{1,2\}}V_{J+1})^\perp$ by the symbol $\Psi_{J,J}$, for $J \in \{j+j', j+j'-1, \dots, |j-j'|\}$. (Note that in the case that $J = j + j'$ the appropriate space is just $V_{j+j'}$ since $V_{j+j'+1} = \{0\}$.) Then one defined $\Psi_{J,M}$ for $M \in \{J, J-1, \dots, -J\}$ by the formula

$$\Psi_{J,M} \;=\; \frac{(S^-_{\{1,2\}})^{J-M}\Psi_{J,J}}{\|(S^-_{\{1,2\}})^{J-M}\Psi_{J,J}\|}\,.$$

Using the techniques of the last section, one may realize that $\{\Psi_{J,M}\}_{M=-J}^{J}$ forms an orthonormal basis for a sub-representation of V equivalent to $\mathcal{D}^{(J)}$. Then this implies that

$$\mathcal{D}^{(j)} \otimes \mathcal{D}^{(j')} \;\supseteq\; \bigoplus_{J=|j-j'|}^{j+j'} \mathcal{D}^{(J)}\,.$$

But by Exercise 15, we know that the dimension of the direct sum on the right equals the dimension of the tensor product on the left. Therefore, there is not just inclusion, there is identity. This prove the theorem.

Before moving on to the next section, let us introduce the Casimir operator.

The total spin operator is $\boldsymbol{S}^2 = (S^x)^2 + (S^y)^2 + (S^z)^2$. This operator is well-defined on any representation, or in the universal enveloping algebra of the Lie algebra for SU(2). It is also called the

Casimir operator. In the universal enveloping algebra for SU(2) it generates the center, which is the defining characteristic of "Casimir operators."

Exercise 17. *(a) Using the formulas* $S^{\pm} = (S^x \pm iS^y)/2$*, prove that*

$$\boldsymbol{S}^2 \,=\, (S^z)^2 + \frac{1}{2}\, S^+ S^- + \frac{1}{2}\, S^- S^+ \,.$$

(b) Using the fact that $S^+ S^z = (S^z - 1)S^+$ *and* $S^- S^z = (S^z + 1)S^-$*, prove that* $\boldsymbol{S}^2$ *commutes with* S^z*.*
(c) Using the symmetry $(S^x, S^y, S^z) \mapsto (S^y, S^z, S^x)$*, which preserves the commutation relations, argue that* $\boldsymbol{S}^2$ *commutes with* S^x *and* S^y *as well.*
(d) In $\mathcal{D}^{(j)}$*, with the orthonormal basis* $\psi_j, \psi_{j-1}, \dots, \psi_{-j}$*, argue that*

$$\boldsymbol{S}^2 \psi_j \,=\, j^2 \psi_j + S^+ S^- \psi_j \,=\, j(j+1)\psi_j \,.$$

(e) Using the commutativity of $\boldsymbol{S}^2$ *with* S^x *and* S^y*, argue that for every vector* $\psi \in \mathcal{D}^{(j)}$*, one has* $\boldsymbol{S}^2\psi = j(j+1)\psi$*.*
(f) Recall that the determination that $S^+S^-\psi_j = 2j\psi_j$ *comes from the fact that* $S^+\psi_j = 0$ *and* $S^z\psi_j = j\psi_j$ *– in other words that* ψ_j *is a highest weight vector of weight* j *– and the commutation relation* $S^+S- = 2S^z - S^-S^+$*. Conclude that in* $\mathcal{D}^{(j)} \otimes \mathcal{D}^{(j')}$*, it is also true that* $\boldsymbol{S}^2_{\{1,2\}}\Psi_{J,J} = J(J+1)\Psi_{J,J}$*, where* $\boldsymbol{S}^2_{\{1,2\}} = (S^x_{\{1,2\}})^2 + (S^y_{\{1,2\}})^2 + (S^z_{\{1,2\}})^2$*.*

From Exercise 17, part e, we have the reasoning for called $\mathcal{D}^{(j)}$ the "spin-j irreducible representation of SU(2)." The fact that $\boldsymbol{S}^2\psi = j(j+1)\psi$ instead of equalling $j^2\psi$ is typically understood in terms of the non-commutativity of S^x, S^y and S^z. Moreover, one may use this to define a version of the Heisenberg uncertainty principle for SU(2) (in much the same way that the commutation relations of the position and momentum operators in the Heisenberg-Weyl algebra lead to the usual uncertainty principle):

$$\langle (S^x - \langle S^x\rangle)^2\rangle\langle (S^y - \langle S^y\rangle)^2\rangle \,\geq\, \frac{1}{4}\langle S^z\rangle^2\,,$$

as well as the full family of inequalities one obtains by rotating x, y and z in the formula above by an arbitary element of SO(3):

$$\langle (\boldsymbol{x}' \cdot \boldsymbol{S} - \langle \boldsymbol{x}' \cdot \boldsymbol{S}\rangle)^2\rangle\langle (\boldsymbol{x}'' \cdot \boldsymbol{S} - \langle \boldsymbol{x}'' \cdot \boldsymbol{S}\rangle)^2\rangle \,\geq\, \frac{1}{4}\langle \boldsymbol{x} \cdot \boldsymbol{S}\rangle^2\,,$$

whenever $\boldsymbol{x}$, $\boldsymbol{x}'$ and $\boldsymbol{x}''$ are three vectors in S^2 which are pairwise orthogonal. The vectors which maximize the left-hand-side of one of these inequalities are called "SU(2) coherent states" in analogy with the canonical coherent states of the Heisenberg-Weyl algebra which maximize the usual uncertainty. It turns out that they are the same as states which maximize $\langle \boldsymbol{x} \cdot \boldsymbol{S}\rangle$ for some $\boldsymbol{x} \in \boldsymbol{S}^2$. Another description is to start with the highest weight vector $\psi_j \in \mathcal{D}^{(j)}$ and then rotate it all over the sphere by acting by SO(3) rotations lifted to SU(2). We know that for j not an integer, this is not a well-defined procedure because SU(2) is a double-cover of SO(3). The vector obtained $\psi_{\boldsymbol{x}}$ will depend on the path taken to go from $\boldsymbol{e}_3$ to $\boldsymbol{x}$. But the dependence will only show up in the phase factor. Therefore, it will cancel out when one defines the actual state $|\psi_{\boldsymbol{x}}\rangle\langle\psi_{\boldsymbol{x}}|$. A standard reference for this type of thing is:

[6] *Generalized Coherent States and Their Applications,* A. Perelomov. Springer-Verlag, Berlin, 1986.

From Exercise 17, part f, and the argument for part e, it follows that $\boldsymbol{S}^2_{\{1,2\}}\Psi_{J,M} = J(J+1)\Psi_{J,M}$ while $S^z_{\{1,2\}}\Psi_{J,M} = M\Psi_{J,M}$, for every $J \in \{j+j',\dots,|j-j'|\}$ and every $M \in \{J,\dots,-J\}$.

There are a couple of additional exercises that one can consider doing.

(1) One is to calculate the Clebsch-Gordan coefficients. These are the coefficients $cJ,M(m,m')$ such that

$$\Psi_{J,M} = \sum_{(m,m')\in\mathcal{A}_{j,j'}(M)} c_{J,M}(m,m')\psi_{j,m}\otimes\psi_{j',m'}\,,$$

where

$$\mathcal{A}_{j,j'}(M) = \{(m,m') \in \{j,\dots,-j\}\times\{j',\dots,-j'\} \,:\, m+m' = M\}\,.$$

In principle we do have a straightforward algorithm to calculate these since, at least for $M = J$ we know $\sqrt{j(j+1)-m(m-1)}\,c_{J,J}(m-1,m') + \sqrt{j'(j'+1)-m'(m'-1)}\,c_{J,J}(m,m'-1) = 0$ for each $(m,m') \in \mathcal{A}_{j,j'}(J+1)$. Moreover, we know that $\Psi_{J,M} = (S^-_{\{1,2\}})^{J-M}\Psi_{J,J}$ divided by a positive normalization constant. But one may be interested in the precise combinatorics of these coefficients. In fact there is an even more interesting combinatorial question related to this which is the elucidation of the basis of all highest weight vectors in a large tensor product. This was done by Lieb and Temperley for tensor products of spins-1/2, when they constructed Hulthen's bracket basis. Indeed their definition, in the same paper which gave rise to the title "Temperley-Lieb algebra" (by quantum algebraists following much after their original paper), generalizes to quantum groups. (There is a family of q-combinatorial numbers called quantum Clebsch-Gordan coefficients.) This was also generalized to tensor products of higher spin irreducible representations in this nice reference:

[7] Igor B. Frenkel and Mikhail G. Khovanov. Canonical bases in tensor products and graphical calculus for $U_q(\mathfrak{sl}_2)$. *Duke Math. J.* **87**, no. 3 (1997), 409–480.

(2) The Clebsch-Gordan coefficients alternate for highest weight vectors in the following sense. Re-define a new basis for the second tensor factor

$$\tilde{\psi}_{j',m'} = (-1)^{j'-m'}\psi_{j',m'}\,.$$

This amounts to rotating by $e^{i\pi S^z_2}$ modulo a global constant (the same one for all basis vectors). Then, if one defined $\tilde{c}_{J,M}(m,m')$ such that

$$\Psi_{J,M} = \sum_{(m,m')\in\mathcal{A}_{j,j'}(M)} \tilde{c}_{J,M}(m,m')\psi_{j,m}\otimes\tilde{\psi}_{j',m'}\,,$$

one sees that all the coefficients $\tilde{c}_{J,J}(m,m')$ may be chosen positive. Moreover, the inductive relations for the highest weight vector become

$$\sqrt{j(j+1)-m(m-1)}\,\tilde{c}_{J,J}(m-1,m') = \sqrt{j'(j'+1)-m'(m'-1)}\,\tilde{c}_{J,J}(m,m'-1)\,.$$

This may seem like a trivial observation. But one can observe that if one conjugates by $\exp(i\pi S_2^z)$ then the Casimir operator becomes an operator satisfying the Perron-Frobenius theorem. More precisely, we generally have the formula

$$\boldsymbol{S}_{\{1,2\}}^2 \;=\; \boldsymbol{S}_1^2 + \boldsymbol{S}_2^2 + 2\boldsymbol{S}_1 \cdot \boldsymbol{S}_2\,,$$

where $\boldsymbol{S}_1 \cdot \boldsymbol{S}_2$ is known as "Heisenberg's exchange" operator

$$\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 \;=\; S_1^x S_2^x + S_1^y S_2^y + S_1^z S_2^z \;=\; S_1^z S_2^z + \frac{1}{2}\left(S_1^+ S_2^- + S_1^- S_2^+\right)\,.$$

Conjugating by $\exp(i\pi S_2^z)$ will not change either $\boldsymbol{S}_1^2$ or $\boldsymbol{S}_2^2$ because they both commute with S_2^z (one because it is in the algebra localized at the other tensor factor, and the other because it is the Casimir operator, which commutes with every spin operator in the appropriate algebra). But it does alter the Heisenberg exchange operator. In fact

$$e^{-i\pi S_2^z}\boldsymbol{S}_{\{1,2\}}^2 e^{i\pi S_2^z} \;=\; \boldsymbol{S}_1^2 + \boldsymbol{S}_2^2 + 2S_1^z S_2^z - S_1^+ S_2^- - S_1^- S_2^+\,.$$

In the basis $\psi_{j,m} \otimes \psi_{j',m'}$ (which is sometimes known as the "Ising" basis for historical reasons, since Lenz and Ising postulated their discrete classical model even before quantum mechanics gave any sort of a posteriori justification as a classical analogue of Heisenberg's model) the first three operators are all diagonal, while the last two have non-positive off-diagonal entries. This means, by the Perron-Frobenius theorem, that there is a unique eigenvector with lowest eigenvalue in each subspace spanned by Ising basis vectors which may be connected to each other by the off-diagonal hopping matrices. These subspaces are the V_M's for $M \in \{j+j',\dots,-j-j'\}$, and the vectors are the highest weight vectors with smallest possible total spin $J = M$, as long as M is in $\{j+j',\dots,|j-j'|\}$. This may seem like much ado about nothing, since there are other ways of calculating these vectors. But in the right hands, the ideas implicit in this simple fact were turned into one of the few general results for quantum spin systems. In their landmark paper, Lieb and Mattis (following a previous idea of Marshall related to "good signs") proved that for general Heisenberg antiferromagnets on bipartite lattices that the ground states would have lowest possible spin, and that the energies would be related to the spin in the direct order relation, at least up until a maximum spin which may be less than the maximal possible total spin for ferrimagnets, but which is easily calculable a priori without any need to diagonalize the model. See their reference

[**8**] E. H. Lieb and D. C. Mattis. Ordering Energy Levels of Interacting Spin Systems, *J. Math. Phys.* **3**, 749–751 (1962).

5 Spin wave approximation

To be completed. Proper notes on this topic will probably take the same amount of space as everything else, combined.

6 Additional exercises

6.1 Diagonalizability of unitary matrices

The next exercise is to explain why each matrix $R \in \mathrm{SO}(3)$ must have at least one eigenvalue 1, and must have the other two eigenvalues equal to complex conjugates of the form $e^{i\theta}$ and $e^{-i\theta}$ for $\theta \in [0, \pi]$.

Exercise A. *Suppose you have a matrix $R \in \mathrm{SO}(3)$. Hence, suppose that R is in $M_2(\mathbb{R})$ with* $\det(R) = 1$ *and $R^T R = I$. Now, instead, view R as a matrix in $M_2(\mathbb{C})$ which just happens to have real entries. Assume that R is diagonalizable. In other words, assume that there are three numbers $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ and three linearly vectors $\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3 \in \mathbb{C}^3$ such that $R\boldsymbol{z}_k = \lambda_k \boldsymbol{z}_k$ for $k = 1, 2, 3$.*
(a) Prove that $\|R\boldsymbol{z}\|^2 = \|\boldsymbol{z}\|^2$ for each $\boldsymbol{z} \in \mathbb{C}^3$.
(b) Prove that $|\lambda_k| = 1$ for each $k = 1, 2, 3$.
(c) Prove that $\lambda_1\lambda_2\lambda_3 = 1$.
(d) Prove that $\{\lambda_1, \lambda_2, \lambda_3\} = \{\overline{\lambda}_1, \overline{\lambda}_2, \overline{\lambda}_3\}$. In other words, from what is given, find three linearly independent vectors $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3 \in \mathbb{C}^3$ such that $R\boldsymbol{w}_k = \overline{\lambda}_k \boldsymbol{w}_k$ for $k = 1, 2, 3$.
(e) Prove that $1 \in \{\lambda_1, \lambda_2, \lambda_3\}$, and if $\lambda_1 = 1$ (which may be assumed without loss of generality) then $\lambda_3 = \overline{\lambda}_2$.

It is a well-known fact that any matrix $R \in \mathrm{SO}(3)$ is diagonalizable. In case the reader does not already know this fact, the next exercise leads to a proof.

Theorem *(Jordan canonical form.) Suppose that $L : \mathbb{C}^n \to \mathbb{C}^n$ is any linear operator. Then there is some $k \leq n$ and vector subspaces $V_1, \dots, V_k \subseteq \mathbb{C}^n$ such that the linear span of vectors in $V_1, \dots, V_k$ is all of $\mathbb{C}^n$ and such that* $\dim(V_j) = d_j \in \{1, 2, \dots\}$ *with $d_1 + \dots + d_k = n$. Moreover, the subspaces have the following form. For each $j \in \{1, \dots, k\}$, we can find vectors $\boldsymbol{z}_{j,0}, \dots, \boldsymbol{z}_{j,d_j}$ spanning V_j such that $M\boldsymbol{z}_{j,0} = \lambda_j \boldsymbol{z}_{j,0}$ for some $\lambda_j \in \mathbb{C}$ and (if $d_j > 1$) then, for $\ell = 1, \dots, d_j - 1$, we have $M\boldsymbol{z}_{j,\ell} = \lambda_j \boldsymbol{z}_{j,\ell} + \boldsymbol{z}_{j,\ell-1}$.*

The Jordan canonical form is a well-known result from an introductory course on linear algebra. It may be assumed. It is also easy to prove though requiring several steps. A key to the Jordan canonical form is to notice that in the basis $\{v_{1,0}, \dots, v_{1,d_1-1}, \dots, v_{k,0}, \dots, v_{k,d_k-1}\}$, L has a block-diagonal matrix. I.e., each V_j is an invariant subspace for L, meaning $LV_j \subseteq V_j$. Moreover, whenever $d_j > 1$, the form of L restricted to the invariant subspace V_j, in the basis $v_{j,0}, \dots, v_{j,d_j-1}$ is

$$\begin{bmatrix} \lambda_j & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_j & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda_j & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_j & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_j \end{bmatrix}$$

So $L - \lambda_j I$ is nilpotent on V_j. Whenever there is some j with $d_j > 1$ we say that L has a non-trivial Jordan block.

Exercise B. *Again, suppose that* $R \in M_n(\mathbb{C})$ *is a matrix such that all of the matrix entries happen to be in* $\mathbb{R}$ *and such that* $R^T R = I$*. Suppose that* $\boldsymbol{z} \in \mathbb{C}^n \setminus \{0\}$ *and* $\lambda \in \mathbb{C}$ *are such that* $R\boldsymbol{z} = \lambda \boldsymbol{z}$*. Prove that there is no* $\boldsymbol{w} \in \mathbb{C}^n$ *such that* $R\boldsymbol{w} = \lambda \boldsymbol{w} + \boldsymbol{z}$*, by contradiction, by first checking that should such a* $\boldsymbol{w}$ *exist* $\overline{\lambda}\boldsymbol{z}^* R\boldsymbol{w} = \boldsymbol{z}^*\boldsymbol{w}$*.*

Using the previous two exercises, we can prove the result from Section 2.1.1.

Exercise C. *View* $R \in \mathrm{SO}(3)$ *as a matrix in* $M_3(\mathbb{C})$ *which happens to have real matrix entries, such that* $\det(R) = 1$ *and* $R^T R = I$*.*
(a) If $R\boldsymbol{z} = \lambda \boldsymbol{z}$ *for some* $\boldsymbol{z} \in \mathbb{C}^3 \setminus \{0\}$ *and some* $\lambda \in \mathbb{C}$*, then prove that* $R\overline{\boldsymbol{z}} = \overline{\lambda}\overline{\boldsymbol{z}}$*.*
(b) Using Exercise A, prove that there is some $\boldsymbol{x} \in \mathbb{R}^3 \setminus \{0\}$ *such that* $R\boldsymbol{x} = \boldsymbol{x}$*.*
(c) Given $\boldsymbol{x}$ *from part (b), let* $V = \{\boldsymbol{y} \in \mathbb{R}^3 \,:\, \boldsymbol{x}^T\boldsymbol{y} = 0\}$*. Prove that* R *maps* V *into* V*.*
(d) Given that $\lambda_1 = 1$ *there are several other possibilities for* $\lambda_2 = e^{i\theta}$ *and* $\lambda_3 = e^{-i\theta}$*. If they are both 1 then* $R = I$*. If* $\lambda_2 = \lambda_3 = -1$ *then* R *restricted to* V *acts as negative of the identity. In that case* $R = R_{\boldsymbol{x},\pi}$ *which may also be written as* $R_{\boldsymbol{x},-\pi}$*. Otherwise, if* $\theta \in (0,\pi)$ *(which we may assume without loss of generality, so that* $-\theta \in (-\pi, 0)$*) let* $\boldsymbol{z} \in \mathbb{C}^3$ *be an eigenvector corresponding to the eigenvalue* $\lambda_2 = e^{i\theta}$ *for* R*, viewed as a matrix in* $M_3(\mathbb{C})$*. Prove that* $\boldsymbol{z}$ *is neither purely real, nor purely imaginary.*
(e) Continuing with the last case of (d), let $\boldsymbol{y}_1$ *and* $\boldsymbol{y}_2$ *be the real and imaginary parts of* $\sqrt{2}\boldsymbol{z}$*. Check that these two vectors are ortho-normal.*
(f) Starting with

$$R(\boldsymbol{y}_1 + i\boldsymbol{y}_2) \,=\, e^{i\theta}(\boldsymbol{y}_1 + i\boldsymbol{y}_2)$$

conclude that in the basis $\{\boldsymbol{y}_1, \boldsymbol{y}_2\}$ *the matrix for* R *restricted to* V *is*

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

(f) If $\det(\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2) = -1$*, then prove that* $R = R_{-\boldsymbol{x},\theta}$*. Otherwise, prove that* $R = R_{\boldsymbol{x},\theta}$*.*

6.2 Haar measure and existence of an invariant inner-product

The next exercise is to explore the Haar measure on SU(2). The motivation is to show how to construct an invariant inner-product on any finite dimensional representation. Let V be a finite-dimensional, complex representation of SU(2), by which we mean that there are three linear operators on V, called S^x, S^y and S^z satisfying the SU(2) commutation relations.

Let $v_1, \dots, v_n$ be any basis for V, so we are assuming $\dim(V) = n$. Then let us start by constructing a positive-definite sesqui-linear form

$$B(v, w) \,=\, \sum_{k=1}^{n} \overline{c_k} d_k \,, \quad \text{for } v = c_1 v_1 + \dots + c_n v_n \text{ and } w = d_1 v_1 + \dots + d_n v_n,$$
$$\text{for } c_1, \dots, c_n, d_1, \dots, d_n \in \mathbb{C}.$$

Note that we use the physicists convention for the sesqui-linearity. (Most mathematicians do use this convention, perhaps because it matches with Dirac's "bra-ket" notation which is quite useful, at least mnemonically.) The problem with this potential inner-product is that we do not necessarily have equality of $B(v, S^x w)$ and $B(S^x v, w)$, nor the conditions for S^y and S^z to be self-adjoint with respect to B, either. This is not surprising, since B was constructed in an arbitrary manner.

There is a way to fix this, to give an inner-product $\langle \cdot, \cdot \rangle$ for which the representation is unitary. This is to symmetrize B. For each group element g of the actual Lie group SU(2), not its Lie algebra, one conjugates B by the representation U_g induced by the three spin matrices S^x, S^y and S^z. This gives a new positive-definite sequilinear form B_g. Then one averages over the Haar measure on SU(2). The Haar measure is the (unique up to normalization) measure on a compact Lie group which is invariant under left (and right) multiplication. Using this symmetry property of the Haar measure, one can deduce that the new positive-definite sesquilinear form, obtained by averaging over the Haar measure, does lead to unitarity.

In order to go back and forth between the Lie algebra and the Lie group one needs differentiation and exponentiation. These are related to Stone's theorem and the Baker-Campbell-Hausdorff formula that are introduced in some other lectures. But for these finite dimensional representations, of a very concrete Lie group which we have already understood geometrically, we do not necessarily need those techniques. Those techniques are general tools which help even when one does not have explicit formulas for Lie groups and Lie algebras. (Another way to say this is that Stone's theorem and the Baker-Campbell-Hausdorff formula are theoretical tools which are often used [perhaps in conjuction with other tools] to classify an unknown set of objects, such as, for example, classifying all semi-simple Lie groups using representation theory, which is the actual subject of typical graduate courses on representation theory, involving Coxeter groups, Dynkin diagrams, and all that.)

Exercise D. *We have realized* SU(2) *as a submanifold of* $\mathbb{H} = \{x\mathbb{i} + y\mathbb{j} + z\mathbb{k} + t\mathbb{1} \, : \, x, y, z, t \in \mathbb{R}\}$. *There are various ways to see that the norm on* $\mathbb{H}$ *is multiplicative.*
(a) Assume the formula (or prove it) which we have used before:

$$(Q_{\boldsymbol{x}} + t\mathbb{1})(Q_{\boldsymbol{x}'} + t'\mathbb{1}) \, = \, Q_{\boldsymbol{x}\wedge\boldsymbol{x}'+t\boldsymbol{x}+t'\boldsymbol{x}} + (tt' - \boldsymbol{x}\cdot\boldsymbol{x}')\mathbb{1} \, .$$

Prove that this means that the norm is multiplicative $\|x\mathbb{i} + y\mathbb{j} + z\mathbb{k} + t\mathbb{1}\| = x^2 + y^2 + z^2 + t^2$. *(If it helps, use or prove that* $\|\boldsymbol{x} \wedge \boldsymbol{x}'\|^2 + (\boldsymbol{x}\cdot\boldsymbol{x}')^2 = \|\boldsymbol{x}\|^2\|\boldsymbol{x}'\|^2$.*)*
(b) Extend this by proving that the $\mathbb{R}^4$ *inner-product may be realized as the real part of* $(Q_{\boldsymbol{x}}+t\mathbb{1})(Q_{-\boldsymbol{x}'}+t'\mathbb{1})$*, where* $Q_{-\boldsymbol{x}'} + t'\mathbb{1}$ *is the conjugate of* $Q_{\boldsymbol{x}'} + t'\mathbb{1}$ *and the norm-squared of a quaternion is obtained by multiplying with its conjugate (which equals the norm-square time* $\mathbb{1}$*).*
(c) Using this, prove that if $Q_{\boldsymbol{x}} + t\mathbb{1}$ *is a unit quaternion then left multiplication by this quaternion on* $\mathbb{H}$ *leads to an orthogonal transform when viewed as a linear transformation on* $\mathbb{R}^4$.

It is well understood that Lebesgue measure on $\mathbb{R}^4$ is invariant under orthogonal transformations. (There are many ways to see this, some requiring very little writing, but perhaps more background.) Therefore, putting Lebesgue measure on $\mathbb{H}$ and then using the polar decomposition of Lebesgue measure into spherical coordinates gives the uniform measure on SU(2), which is equivalent to $\mathbb{S}^3$. What remains is to understand what measure this gives when we exponentiate an element $\exp(Q_{\boldsymbol{x}})$ since exponentiation will be our main tool for constructing a representation of the Lie group SU(2), given the representation of the Lie algebra through the three spin matrices S^x, S^y and S^z.

Exercise E. *(a) Prove that* $\exp(Q_{\boldsymbol{x}}) = \|\boldsymbol{x}\|^{-1}\sin(\|\boldsymbol{x}\|)Q_{\boldsymbol{x}} + \cos(\|\boldsymbol{x}\|)\mathbb{1}$.
(b) Writing $\mathbb{S}^3 = \{(x,y,z,t) \in \mathbb{R}^4 \,:\, x^2+y^2+z^2+t^2 = 1\}$, *and* S_3 *for the three-dimensional surface measure on* $\mathbb{S}^3$, *prove that for any continuous function* $f : [-1,1] \to \mathbb{R}$,

$$\int_{\mathbb{S}^3} f(t)\, dS_3(x,y,z,t) \;=\; 4\pi \int_{-1}^{1} f(t)\sqrt{1-t^2}\, dt\,.$$

(c) Explain why the inner-product $\langle\cdot,\cdot\rangle$ *on* V *should be given, up to a normalization constant, by*

$$\langle v, w\rangle \;=\; \int_{\mathbb{S}^2}\left(\int_0^{\pi} B(e^{i\theta[xS^x+yS^y+zS^z]}v, e^{i\theta[xS^x+yS^y+zS^z]}w)\sin^2(\theta)\,d\theta\right) dS_2(x,y,z)\,,$$

where S_2 *is the two-dimensional surface measure on* $\mathbb{S}^2 = \{(x,y,z)\in\mathbb{R}^3 \,:\, x^2+y^2+z^2=1\}$.

Let us write $\boldsymbol{a} = (a,b,c) \in \mathbb{R}^3$ in order to not re-use the symbol $\boldsymbol{x} = (x,y,z)$, already used. Then the idea is that if we try to calculate $\langle v, e^{i(aS^x+bS^y+cS^z)}w\rangle$, then we have

$$\langle v, e^{i(aS^x+bS^y+cS^z)}w\rangle \;=\; \int_{\mathbb{S}^2}\left(\int_0^{\pi} B(e^{i\theta\boldsymbol{x}\cdot\boldsymbol{S}}v, e^{i\theta\boldsymbol{x}\cdot\boldsymbol{S}}e^{i\boldsymbol{a}\cdot\boldsymbol{S}}w)\sin^2(\theta)\,d\theta\right) dS_2(x,y,z)\,,$$

where we write $\boldsymbol{S}$ for the spin-matrix-vector (S^x,S^y,S^z). Then, using the group structure on SU(2), we can rewrite

$$e^{i\theta\boldsymbol{x}\cdot\boldsymbol{S}}e^{i\boldsymbol{a}\cdot\boldsymbol{S}} \;=\; e^{i\Theta_{\boldsymbol{a}}(\theta,\boldsymbol{x})\boldsymbol{X}_{\boldsymbol{a}}(\theta,\boldsymbol{x})\cdot\boldsymbol{S}}\,.$$

If you want another exercise in long explicit calculations, you can prove that

$$\cos(\Theta) \;=\; \cos(\theta)\cos(\|\boldsymbol{a}\|) - \sin(\theta)\sin(\|\boldsymbol{a}\|)\,\frac{\boldsymbol{x}\cdot\boldsymbol{a}}{\|\boldsymbol{a}\|}\,, \tag{9}$$

$$\sin(\Theta)\boldsymbol{X} \;=\; \sin(\theta)\sin(\|\boldsymbol{a}\|)\frac{\boldsymbol{x}\wedge\boldsymbol{a}}{\|\boldsymbol{a}\|} + \sin(\theta)\cos(\|\boldsymbol{a}\|)\boldsymbol{x} + \cos(\theta)\sin(\|\boldsymbol{a}\|)\,\frac{\boldsymbol{a}}{\|\boldsymbol{a}\|}\,, \tag{10}$$

which is closely related to Exercise D, part (a). But in fact Exercise D should be sufficient without these explicit calculations. Hence, $e^{i\theta\boldsymbol{x}\cdot\boldsymbol{S}} = e^{i\Theta_{\boldsymbol{a}}(\theta,\boldsymbol{x})\boldsymbol{X}_{\boldsymbol{a}}(\theta,\boldsymbol{x})\cdot\boldsymbol{S}}e^{-i\boldsymbol{a}\cdot\boldsymbol{S}}$. Therefore, we have

$$\begin{aligned}\langle v, e^{i\boldsymbol{a}\cdot\boldsymbol{S}}w\rangle \;&=\; \int_{\mathbb{S}^2}\left(\int_0^{\pi} B(e^{i\theta\boldsymbol{x}\cdot\boldsymbol{S}}v, e^{i\theta\boldsymbol{x}\cdot\boldsymbol{S}}e^{i\boldsymbol{a}\cdot\boldsymbol{S}}w)\sin^2(\theta)\,d\theta\right) dS_2(x,y,z)\\ &=\; \int_{\mathbb{S}^2}\left(\int_0^{\pi} B(e^{i\Theta_{\boldsymbol{a}}(\theta,\boldsymbol{x})\boldsymbol{X}_{\boldsymbol{a}}(\theta,\boldsymbol{x})\cdot\boldsymbol{S}}e^{-i\boldsymbol{a}\cdot\boldsymbol{S}}v, e^{i\Theta_{\boldsymbol{a}}(\theta,\boldsymbol{x})\boldsymbol{X}_{\boldsymbol{a}}(\theta,\boldsymbol{x})\cdot\boldsymbol{S}}w)\sin^2(\theta)\,d\theta\right) dS_2(x,y,z)\,.\end{aligned}$$

But then, using the group structure, for (Lebesgue) almost every choice of $(\boldsymbol{x}',\theta') \in \mathbb{S}^2\times(0,\pi)$ there exists a unique $(\boldsymbol{x},\theta)$ in $\mathbb{S}^2\times(0,\pi)$ such that

$$e^{i\Theta_{\boldsymbol{a}}(\theta,\boldsymbol{x})\boldsymbol{X}_{\boldsymbol{a}}(\theta,\boldsymbol{x})\cdot\boldsymbol{S}} \;=\; e^{i\theta'\boldsymbol{x}'\cdot\boldsymbol{S}}\,.$$

(The only exceptions have to do with non-uniqueness at $\theta = 0$ and $\theta = \pi$ which are measure-zero sets).

And the measure satisfies

$$\sin^2(\theta)\, d\theta\, dS_2(\boldsymbol{x}) \;=\; \sin^2(\theta')\, d\theta'\, dS_2(\boldsymbol{x}')\,.$$

Again, we could in principle do a direct calculation using (9) and (10) and try to prove the equivalence of the measures. But this should be unnecessary. We know that multiplication by $e^{i\boldsymbol{a}\cdot\boldsymbol{S}}$ is an orthogonal transformation of $\mathbb{H}$. (Technically, we previously argued that left-multiplication is, but right-multiplication is as well.) And we know that Lebesgue measure on $\mathbb{R}^4$ is invariant under an orthogonal transformation (essentially because the determinant-squared is 1). And we know that the orthogonal transformation leaves the unit sphere invariant.

Therefore, if we calculated the spherical measure correctly in Exercise E, then the equivalence of the measures should be automatic from that. It is perhaps only obscured by our choice of coordinates. So then we would get

$$\langle v, e^{i\boldsymbol{a}\cdot\boldsymbol{S}} w\rangle \;=\; \int_{\mathbb{S}^2} \left(\int_0^\pi B(e^{i\theta'\boldsymbol{x}'\cdot\boldsymbol{S}} e^{-i\boldsymbol{a}\cdot\boldsymbol{S}} v, e^{i\theta'\boldsymbol{x}'\cdot\boldsymbol{S}} w) \sin^2(\theta')\, d\theta'\right) dS_2(\boldsymbol{x}') \;=\; \langle e^{-i\boldsymbol{a}\cdot\boldsymbol{S}} v, w\rangle\,.$$

Differentiating with respect to $\boldsymbol{a}$ and using sesqui-linearity shows that S^x, S^y and S^z are each Hermitian with respect to $\langle\cdot,\cdot\rangle$. That was the desired result.

Exercise F. *The introduction of coordinates $(\boldsymbol{x},\theta) \in \mathbb{S}^2 \times (0,\pi)$ is reasonable. But proving invariance of Haar measure in these coordinates is difficult. This exercise is just to prove to ourselves that it can be done. It is just an exercise of calculus skills.*

Let us suppose $\boldsymbol{a} = t\boldsymbol{e}_3$ and let us write $\boldsymbol{x} = (\cos(\phi)\sqrt{1-z^2}, \sin(\phi)\sqrt{1-z^2}, z)$ for $(z,\phi) \in [-1,1] \times (0,2\pi)$. Then we get a parametrization of $\mathbb{S}^3$ which is $(z,\phi,\theta) \in [-1,1]\times(0,2\pi)\times(0,\pi)$, with associated measure

$$\sin^2(\theta)\, d\theta\, d\phi\, dz\,.$$

Then we have from (9) and (10)

$$\begin{aligned}
\cos(\Theta) \;&=\; \cos(\theta)\cos(t) - \sin(\theta)\sin(t)z\,,\\
\sin(\Theta)\boldsymbol{X} \;&=\; \sin(\theta)\sqrt{1-z^2}\,(\cos(t)\boldsymbol{r}_\phi + \sin(t)\boldsymbol{r}_{\phi+\frac{1}{2}\pi}) + (\cos(\theta)\sin(t) + \sin(\theta)\cos(t)z)\boldsymbol{e}_3\\
&=\; \sin(\theta)\sqrt{1-z^2}\,\boldsymbol{r}_{\phi+t} + (\cos(\theta)\sin(t) + \sin(\theta)\cos(t)z)\boldsymbol{e}_3\,,
\end{aligned}$$

where we write $\boldsymbol{r}_\phi = \cos(\phi)\boldsymbol{e}_1 + \sin(\phi)\boldsymbol{e}_2$.
(a) Use the formula $\|\sin(\Theta)\boldsymbol{X}\|^2 = \sin^2(\theta)(1-z^2) + \big(\cos(\theta)\sin(t) + \sin(\theta)\cos(t)z\big)^2$ and the formula $\cos^2(\Theta) = \big(\cos(\theta)\sin(t) + \sin(\theta)\cos(t)z\big)^2$ to prove that $\boldsymbol{X}$ is in $\mathbb{S}^2$.
(b) Write $\boldsymbol{X} = (\cos(\Phi)\sqrt{1-Z^2}, \sin(\Phi)\sqrt{1-Z^2}, Z)$. Prove that $\Phi = \phi + t$, and

$$\sin(\Theta)Z \;=\; \cos(\theta)\sin(t) + \sin(\theta)\cos(t)z\,.$$

(c) Check that

$$
\begin{aligned}
-\sin(\Theta)\, d\Theta \;&=\; \big(-\sin(\theta)\cos(t) - \cos(\theta)\sin(t)z\big)\, d\theta - \sin(\theta)\sin(t)\, dz\,,\\
\cos(\Theta)Z\, d\Theta + \sin(\Theta)\, dZ \;&=\; \big(-\sin(\theta)\sin(t) + \cos(\theta)\cos(t)z\big)\, d\theta + \sin(\theta)\cos(t)\, dz\,.
\end{aligned}
$$

Using the fact that $d\theta\, d\phi\, dz$ really means the anti-symmetric wedge product $d\theta \wedge d\phi \wedge dz$, prove that

$$
\big(-\sin(\Theta)\, d\Theta\big) \wedge \big(\sin(\Theta)\, dZ\big) \;=\; \begin{vmatrix} -\sin(\theta)\cos(t) - \cos(\theta)\sin(t)z & -\sin(\theta)\sin(t) \\ -\sin(\theta)\sin(t) + \cos(\theta)\cos(t)z & \sin(\theta)\cos(t) \end{vmatrix}\, d\theta \wedge dz\,.
$$

(d) Finally conclude that $\sin^2(\Theta)\, d\Theta \wedge dZ = \sin^2(\theta)\, d\theta \wedge dz$*, while* $d\Phi = d\phi$*.*

So this does indeed mean that the measure $\sin^2(\theta)\, d\theta\, d\phi\, dz$ is invariant, since we get $\sin^2(\Theta)\, d\Theta\, d\Phi\, dZ$, under transformation. Note that we did not lose any generality by assuming the particular form for $\boldsymbol{a}$ since we know that we may rotate any point in $\mathbb{S}^2$ to $\boldsymbol{e}_3$. But it is much harder to prove in coordinates, as we have, although the exercise may be good for our calculus muscles.