Useful Notes for the Rotation Group

I. The Definition using 3×3 Matrices

In physical 3-dimensional (vector) space, E^3, **rotations** are transformations that preserve lengths of vectors while changing their directions. Using the symbol R for both the transformation on the vectors $\vec{x}$ and the matrix that acts on the column vectors x used to present the components of $\vec{x}$, we have

$$\forall \vec{x}, \vec{y} \in E^3, \quad (R\vec{x}) \cdot (R\vec{y}) = \vec{x} \cdot \vec{y} \quad \Longrightarrow \quad (Rx)^T(Ry) = x^T y \quad \Longleftrightarrow \quad R^T R = I \; , \tag{1.1}$$

where $I \equiv I_3$ is the 3×3 identity matrix. Rotations should depend **continuously** on the values of some choice of parameters that specify which rotation one is considering. Since Eq. (1.1) implies that the determinant of R has value ± 1, the set of all rotations must come in two separate parts. Examples from each part are the identity, I_3, and its negative, $-I_3$, often called *parity*.

The action of two rotations on a vector, one after the other, surely is also a rotation, since the length is still unchanged. Therefore **the set of all rotations forms a group**, often called by the symbols **O(3)**, where the **O** stands for the word "orthogonal." The subset with determinant $+1$ is also a group, since the product of two $+1$'s is also $+1$; therefore, it constitutes a *subgroup* of all rotations, called **SO(3)**, where the S stands for the word "special." It is also usual to refer to the elements of this subgroup as *proper rotations*; this language means that one excludes those with negative determinant.

From geometric pictures of rotations a choice of parameters can be found by saying that we "rotate through some angle θ, about some axis, $\hat{n}$," giving rise to the notation $R(\hat{n}; \theta)$. (If desired we can specify the unit vector, $\hat{n}$ by giving its spherical coordinates, say (ϑ, φ). This is a particular "choice of parameters" for rotations that is familiar. A good "visualization" of the entire set of (3-dimensional) rotations is then obtained by taking each rotation as specified by a point somewhere in the (3-dimensional) *ball* of radius π. We use the radial coordinate of a point in this ball as the angle and the direction to that point as the direction for the axis in question. Rotations with an angle greater than π are accomplished by rotating around the opposite axis with a corresponding angle less than π, i.e., we use the identity $R(\hat{n};\theta) = R(-\hat{n}; 2\pi - \theta)$ to describe them. This identity, however, does cause the surface of the ball to be a "little bit strange" since $R(\hat{n};\pi) = R(-\hat{n};\pi)$. This requires that we treat every pair of antipodal points on the surface of this ball as only one point; the mathematical language is that we *identify* those two points as simply one point, on a curved space (or manifold).

Using elementary geometry in the plane, and the usual Cartesian coordinates specified by a basis set $\{\hat{x}, \hat{y}, \hat{z}\}$, then it is easy to see the matrix form of the following simple rotations:

$$
\begin{aligned}
R(\hat{z};\theta) &\Longrightarrow \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} , \\
R(\hat{y};\varphi) &\Longrightarrow \begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix} , \\
R(\hat{x};\zeta) &\Longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\zeta & -\sin\zeta \\ 0 & \sin\zeta & \cos\zeta \end{pmatrix} .
\end{aligned}
\tag{1.2}
$$

Before looking at more general cases, it is useful to write out the lowest order expansions, for very small angles, for each of these three cases:

$$
\begin{aligned}
R(\hat{z};\theta) &\Longrightarrow I_3 + \theta \begin{pmatrix} 0 & -1 & 0 \\ +1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + O^2(\theta) \equiv I_3 + \theta\, \mathfrak{J}_z + O^2(\theta) , \\
R(\hat{y};\varphi) &\Longrightarrow I_3 + \varphi \begin{pmatrix} 0 & 0 & +1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} + O^2(\varphi) \equiv I_3 + \varphi\, \mathfrak{J}_y + O^2(\varphi) , \\
R(\hat{x};\zeta) &\Longrightarrow I_3 + \zeta \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & +1 & 0 \end{pmatrix} + O^2(\zeta) \equiv I_3 + \zeta\, \mathfrak{J}_x + O^2(\zeta) ,
\end{aligned}
\tag{1.3}
$$

where the "big-O" symbols indicate the existence of additional terms of second-order or higher in the various angles. The equations define the three *infinitesimal generators* of these three (basic) rotations, namely $\{\mathfrak{J}_x, \mathfrak{J}_y, \mathfrak{J}_z\}$. Following quantum mechanical procedure, it is also customary to introduce associated Hermitean (angular momentum) matrices, $\{J_x, J_y, J_z\}$:

$$\begin{aligned} \mathfrak{J}_i &\equiv -i\, J_i \ , \\ \text{or} \quad \vec{\mathfrak{J}} &\equiv -i\,\vec{J} = -i\,(J_x\hat{x} + J_y\hat{y} + J_z\hat{z}) \ , \end{aligned} \tag{1.4}$$

$$(\mathfrak{J}_i)^j{}_k \quad \Longrightarrow \quad -\epsilon_{ijk} \ , \quad \text{or} \quad (J_i)^j{}_k \quad \Longrightarrow \quad -i\epsilon_{ijk} \ , \tag{1.5}$$

where the ϵ_{ijk} are the usual Levi-Civita symbols, which are completely skew-symmetric, and normalized so that they generate determinants of square matrices. As a useful aside, because of the relationship of determinants and cross products of 3-dimensional vectors, it is straightforward to see that the matrices $\mathfrak{J}_i$ generate the usual cross product, in the following way:

$$\left(\vec{v}\cdot\vec{\mathfrak{J}}\right)\vec{r} = v^a(\mathfrak{J}_a)^b{}_c r^c(\hat{e}_b) = -v^a\epsilon_{abc}r^c(\hat{e}^b) = \epsilon_{bac}(\hat{e}^b)v^a r^c = \vec{v}\times\vec{r}\ , \tag{1.6}$$

where the $\{\hat{e}_b | b = 1,2,3\}$ constitute a Cartesian basis for the (3-dimensional) space being studied.

One returns to the complete rotations by using the exponential function; for example

$$R(\hat{z};\theta) = e^{\theta\mathfrak{J}_z} = e^{-i\theta J_z} \ , \tag{1.7}$$

where the exponential function is defined by its infinite power series.

II. A General Parametrization for Arbitrary Axes

To consider general rotation matrices, rather than simply those which rotate about the 3 basis vectors, we again consider them in terms of their generators by looking at small angles. Since the identity is a rotation by exactly zero angle, we can "measure" how much an arbitrary rotation, R deviates from the identity in terms of its logarithm, Q, by first writing

$$R = e^Q \ , \qquad Q \equiv \log[I_3 + (R - I_3)] = \sum_{s=1}^{\infty}\frac{(-1)^{s+1}}{s}(R - I_3)^s \ , \tag{2.1}$$

which will converge for R "sufficiently near" the identity, which is the same as saying that the elements of $R - I_3$ should be "sufficiently small." We may then take the defining requirement, Eq. (1.1), that R be a rotation, and find its counterpart for the logarithm Q, as follows:

$$e^0 = I_3 = R^T R = \left(e^Q\right)^T e^Q = e^{Q^T} e^Q = e^{Q^T + Q}$$
$$\Longrightarrow \quad 0 = Q^T + Q \; , \tag{2.2}$$

which is the same as saying that Q must be a skew-symmetric matrix $\Longleftrightarrow$ R is an orthogonal matrix, i.e., a rotation, near the identity. As it turns out, all the series defining the logarithm happens to converge for all **proper** rotations. This is not true for general Lie groups where the range of convergence is often smaller than the entire, connected group.

In general, in n dimensional space, a skew-symmetric matrix has $\frac{1}{2}n(n-1)$ degrees of freedom. In our 3-dimensional space, this tells us that there are exactly 3 independent degrees of freedom to such matrices, so that it is also true that

a. there are 3 degrees of freedom to the group of all (3-dimensional) proper rotations.

b. The 3 skew-symmetric matrices, $\mathfrak{J}_i$, already presented in Eq. (1.3), clearly constitute a choice of basis for the vector space of all Q's; i.e. any skew-symmetric 3×3 matrix is a linear combination of those 3 matrices, say $a\mathfrak{J}_x + b\mathfrak{J}_y + c\mathfrak{J}_z \equiv \vec{A} \cdot \vec{\mathfrak{J}}$. The 3 scalars $\{a, b, c\}$ may then be thought of as the components of a vector $\vec{A}$, which is a choice of parametrization for all proper rotations. However, it is more convenient to separate out the direction and magnitude of $\vec{A} \equiv \zeta\hat{n}$, so that $Q = \zeta\hat{n} \cdot \vec{\mathfrak{J}}$, and we will see that $\hat{n}$ is an eigenvector of the rotation it describes, i.e., it is the axis about which the rotation is being described:

$$R(\hat{n};\zeta) = e^{\zeta\,\hat{n}\cdot\vec{\mathfrak{J}}} = \cos\theta \; I_3 + (1 - \cos\theta)\hat{n}\hat{n}^T + \sin\theta \; \hat{n} \cdot \vec{\mathfrak{J}} \; , \tag{2.3a}$$

$$\text{or} \qquad R(\hat{n};\zeta)\vec{V} = \cos\theta \, \vec{V} + (1 - \cos\theta)(\hat{n} \cdot \vec{V})\hat{n} + \sin\theta \, \hat{n} \times \vec{V} \; . \tag{2.3b}$$

The exponential has been evaluated by an iteration procedure rather than by an explicit calculation of the matrices. This is found by beginning with the vector cross product in

Eq. (1.6), which allows us to calculate the second, and third powers:

$$\begin{aligned} \left(\vec{v}\cdot\vec{\mathfrak{J}}\right)^2 \vec{r} \; &= \; \vec{v}\times(\vec{v}\times\vec{r}) = \vec{v}(\vec{v}\cdot\vec{r}) - v^2\vec{r} = \left[\vec{v}\vec{v}^T - v^2 I_3\right]\vec{r} \, , \\ \Rightarrow \left(\vec{v}\cdot\vec{\mathfrak{J}}\right)^3 \vec{r} = \vec{v}\times\left(v^2\vec{r}\right) &= v^3(\hat{v}\times\vec{r}) \qquad \Longrightarrow \qquad \left(\hat{n}\cdot\vec{\mathfrak{J}}\right)^3 = -\left(\hat{n}\cdot\vec{\mathfrak{J}}\right) \, , \end{aligned} \tag{2.4}$$

III. Action of the Group on its Algebra

Let $R_0 \equiv e^{\theta\hat{n}\cdot\vec{\mathfrak{J}}}$ be a particular rotation, and R some arbitrary rotation. Since these are matrices, acting as linear operators on a single vector space, the action of R on R_0 is determined by a similarity transformation:

$$R^{-1}R_0R = R^{-1}\left\{I_3 + \theta\hat{n}\cdot\vec{\mathfrak{J}} + O^2(\theta)\right\}R = I_3 + \theta\{R^{-1}\left(\hat{n}\cdot\vec{\mathfrak{J}}\right)R\} + O^2(\theta) \, , \tag{3.1}$$

which shows the relation of this action to the action on the Lie algebra, i.e., the $\mathfrak{J}$'s, again via a similarity transformation.

There are various ways to calculate this result; I will sketch two distinct proofs. The first uses particular properties of our 3×3 matrices, while the second uses only the properties of the **commutators** of the Lie algebra elements.

Proof No. 1

This method will use explicit properties of the 3×3 matrices, but properties that give a little extra "insight" into the structure of the rotations. We begin with an explicit statement that the determinant of a rotation is just $+1$:

$$\epsilon_{ijk}R^i{}_aR^j{}_bR^k{}_c = |R|\epsilon_{abc} = \epsilon_{abc} \, , \tag{3.2}$$

$$\Longrightarrow \qquad \epsilon_{ijm}R^i{}_aR^j{}_b = \epsilon_{abc}(R^{-1})^c{}_m \, , \tag{3.3}$$

where the second line follows by multiplying both sides of the equation by $(R^{-1})^c{}_m$. Now, using Eq. (1.5), which relates the matrix elements of $(\mathfrak{J}_i)^j{}_k$ and ϵ_{ijk}, we first relate the action of the rotation R on the generator $\mathfrak{J}^a$:

$$(R^{-1}\mathfrak{J}^mR)^b{}_a = -(R^{-1})^b{}_j\epsilon^{mj}{}_iR^i{}_a = -\epsilon_{mji}R_j{}^bR^i{}_a = \epsilon^m_{ij}R^{jb}R^i{}_a = \epsilon_{ij}{}^mR^i{}_aR^{jb} \, . \tag{3.4}$$

The final result is the same as the left-hand side of Eq. (3.3) except that the m and b indices have been raised. Therefore, we may equate our desired similarity transformation of $\boldsymbol{\mathcal{J}}^a$ to the right-hand side of Eq. (3.3) with those indices raised and continue:

$$\epsilon_a{}^b{}_c(R^{-1})^{cm} = \epsilon_a{}^b{}_f cR^{mc} = -\epsilon_c{}^b{}_a R^{mc} = -\epsilon^{cb}_a R^m{}_c = (\boldsymbol{\mathcal{J}}^c)^b{}_a R^m{}_c = R^m{}_c(\boldsymbol{\mathcal{J}}^c)^b{}_a \ . \tag{3.5}$$

The entire equality, from Eq. (3.3) through Eq. (3.5), can now be summarized in a very important way, where we use a matrix notation for the "component" notation just used above:

$$R^{-1}\,\vec{\boldsymbol{\mathcal{J}}}\,R \;=\; R\vec{\boldsymbol{\mathcal{J}}}\,, \tag{3.6}$$

$$\text{or} \qquad R^{-1}\left(\vec{V}\cdot\vec{\boldsymbol{\mathcal{J}}}\right)R \;=\; \vec{V}\cdot\left(R\vec{\boldsymbol{\mathcal{J}}}\right) = \left(R^{-1}\vec{V}\right)\cdot\vec{\boldsymbol{\mathcal{J}}}\,. \tag{3.7}$$

A very useful way to think about Eq. (3.6) is engendered by remembering that the quantity $\vec{\boldsymbol{\mathcal{J}}}$ is both a matrix and a (3-dimensional) vector. On the left-hand side of the equation we have the rotation R acting on $\vec{\boldsymbol{\mathcal{J}}}$ as if it were a matrix—via a similarity transformation—while on the right-hand side the rotation acts as if $\vec{\boldsymbol{\mathcal{J}}}$ were a vector. The equation says that these two distinct sorts of actions give the same result: that it is indeed legitimate to say that $\vec{\boldsymbol{\mathcal{J}}}$ is *a matrix-valued vector.*

We can now proceed further, by taking R also as infinitesimally near the identity; inserting into our equation the exponential form for the rotation, namely $R = e^{\theta\hat{n}\cdot\vec{\boldsymbol{\mathcal{J}}}} = I_3 + \theta\hat{n}\cdot\vec{\boldsymbol{\mathcal{J}}} + O^2(\theta)$, will allow us to calculate the commutator of two of the angular-momentum generators. The left-hand side of Eq. (3.7) can then be written out as

$$\begin{aligned}
&[I_3 - \theta\hat{n}\cdot\vec{\boldsymbol{\mathcal{J}}} + O^2(\theta)]\left(\vec{V}\cdot\vec{\boldsymbol{\mathcal{J}}}\right)[I_3 + \theta\hat{n}\cdot\vec{\boldsymbol{\mathcal{J}}} + O^2(\theta)] \\
&\qquad = \left(\vec{V}\cdot\vec{\boldsymbol{\mathcal{J}}}\right) - \theta\left(\hat{n}\cdot\vec{\boldsymbol{\mathcal{J}}}\right)\left(\vec{V}\cdot\vec{\boldsymbol{\mathcal{J}}}\right) + \theta\left(\vec{V}\cdot\vec{\boldsymbol{\mathcal{J}}}\right)\left(\hat{n}\cdot\vec{\boldsymbol{\mathcal{J}}}\right) + O^2(\theta) \\
&\qquad = \left(\vec{V}\cdot\vec{\boldsymbol{\mathcal{J}}}\right) + \theta\left[\left(\vec{V}\cdot\vec{\boldsymbol{\mathcal{J}}}\right),\left(\hat{n}\cdot\vec{\boldsymbol{\mathcal{J}}}\right)\right] + O^2(\theta)\,,
\end{aligned} \tag{3.8}$$

where the brackets in the last line are the matrix commutator of the two matrices involved. On the other hand, we evaluate the right-hand side of Eq. (3.7) by the use of Eq. (2.3b):

$$\begin{aligned}
\left(R^{-1}\vec{V}\right)\cdot\vec{\boldsymbol{\mathcal{J}}} &= \left\{\cos(-\theta)\,\vec{V} + [1-\cos(-\theta)](\hat{n}\cdot\vec{V}) + \sin(-\theta)\,\hat{n}\times\vec{V}\right\}\cdot\vec{\boldsymbol{\mathcal{J}}} \\
&= \vec{V}\cdot\vec{\boldsymbol{\mathcal{J}}} - \theta\,\hat{n}\times\vec{V}\cdot\vec{\boldsymbol{\mathcal{J}}} + O^2(\theta)\ .
\end{aligned} \tag{3.9}$$

Comparing the final results, then, of Eqs. (3.8) and (3.9), cancelling the zeroth-order term in θ, dividing by θ, and taking the limit as $\theta \to 0$, we have the desired result for the commutator of two infinitesimal generators, i.e., two elements of the Lie algebra:

$$\left[\left(\vec{V} \cdot \vec{\mathfrak{J}}\right), \left(\hat{n} \cdot \vec{\mathfrak{J}}\right)\right] = + \left(\vec{V} \times \hat{n}\right) \cdot \vec{\mathfrak{J}} \, . \tag{3.10}$$

Taking the vector $\vec{V}$ simply as $\hat{e}_i$ and $\hat{n}$ as $\hat{e}_j$, we can also pick out the indicial version of this relationship, and using the form of the cross product that involves the Levi-Civita symbol:

$$[\mathfrak{J}_i, \mathfrak{J}_j] = \epsilon_{ij}{}^k \mathfrak{J}_k \, . \tag{3.11}$$

Proof No. 2

A more general method to determine the action of the rotations themselves on their Lie algebra does exist. Its value is that it depends only on the properties of the commutators of the Lie algebra; these commutators have the same form for every representation of the generators, i.e., every triplet of matrices, of any dimension, that satisfies Eqs. (3.11). We have of course not yet discussed representations in other dimensions, which we will do somewhat later on in these notes. Nonetheless, it seems to be of value to give this much more general proof of the result in Eqs. (3.6-7), since we will be able to use it to considerable advantage later on.

To begin this discussion, then, we must have some fairly general lemmas concerning the behavior of exponential functions of matrices. I first list some simple properties, where we take A, B, and C, below to be arbitrary square matrices, all three of the same size and type:

$$(e^A)^T = e^{A^T} \, , \quad (e^A)^\dagger = e^{A^\dagger} \, , \quad (e^A)^{-1} = e^{-A} \, ,$$

$$\det(e^A) = e^{\mathrm{tr}\ A} \, ,$$

$$A\, e^B \, A^{-1} = e^{ABA^{-1}} \, ,$$

$$e^A \, e^B \neq e^{A+B} \, , \qquad \text{when} \ \ [A, B] \neq 0 \, .$$

The last line above is not extremely useful, since we are only told what the product is not. On the other hand, the Baker-Campbell-Hausdorff theorem, in principle, answers the question as to what it is:

$$e^{sA}\, e^{tB} \equiv e^{H(sA,tB)} \ ,$$

$$\text{where} \quad H(sA,tB) = sA + tB + \tfrac{1}{2}st[A,B] + \tfrac{1}{12}\left(s^2t[A,[A,B]] + st^2[B,[B,A]]\right)$$
$$+\tfrac{1}{48}s^2t^2\left([A,[B,[B,A]]] - [B,[A,[A,B]]]\right) + O^5(t,s) \ . \tag{3.12}$$

Unfortunately, while I can indeed write down the fifth-order terms, the sixth-order terms, etc., there is no general form that gives the n-th order terms without having first calculated the $(n-1)$-st order terms; nonetheless this is in fact the "best one can do." However, there is a property that can be shown concerning $H(sA,tB)$, namely, that it only involves commutators made from A and B, as one sees in the example given. Since A and B are matrices, they of course have just the ordinary matrix product available to them; nonetheless, this theorem tells us that the desired quantity, $H(sA,tB)$, involves only this skew-symmetric, commutator product, i.e., $[A,B] \equiv AB - BA$.

The similarity transformation describes a different sort of a commutator, namely $R^{-1}\,B\,R$, which describes the action of the group element R on the matrix B. Because of the value of studying the group elements in terms of their logarithms, i.e., their infinitesimal generators, this commutator involves the properties of exponentials of matrices, i.e., this (matrix) similarity transformation could be thought of as trying to determine a simple formulation for the quantity $e^{tA}\,B\,e^{-tA}$. As it turns out this is a much simpler thing to do than to answer the question in the B-C-H theorem; however, because of it, the answer will only involve commutators, and of course repeated commutators, of A and B.

We may define the problem as follows, and will then give a derivation of the result:

$$f(t) \equiv e^{tA}Be^{-tA} \implies (d/dt)f(t) = Af(t) - f(t)A \equiv [A, f(t)] \equiv (\mathrm{ad}\,A)f(t) \ , \tag{3.13}$$

where this is the definition of the operation referred to as (ad A). It is a mapping, from a group of matrices, say, to operators on those matrices, called the *adjoint* operator; in this instance it is just the commutator, but we will see it become more useful than that might directly imply. The derivative with respect to t gives us a simple differential equation for the desired quantity, $f(t)$, which will be a very convenient way to calculate its value. As $f(t)$ involves exponentials of t, it must surely be analytic for sufficiently small values of t, so that we can expand it in a power series about the origin and then construct a power-series expansion for the solution of the differential equation:

$$f(t) \equiv \sum_{n=0}^{\infty} \frac{t^n}{n!} f_n \qquad \Longrightarrow f_{n+1} = [A, f_n] = (\mathrm{ad}\, A) f_n \ , \ \text{and} \ f_0 = f(0) = B \ ,$$

$$\Longrightarrow \quad f_n = (\mathrm{ad}\, A)^n B \quad \Rightarrow f(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} (\mathrm{ad}\, A)^n B \tag{3.14}$$

Between Eqs. (3.13-14), we now have an explicit formultation for the action, via similarity transformation, of a group element on an arbitrary matrix, involving only repeated commutators. (This motivates the abstract definition of a Lie algebra as a vector space of objects that also have (only) a skew-symmetric product. The vector space of all square matrices, with the commutator product, is an obvious way to acquire a particular such algebra.)

Taking as fundamental the commutator product for our generators $\{\mathfrak{J}_i\}_1^3$, as given by Eqs. (3.11), or equivalently by Eq. (3.10), we may now calculate our desired result for the rotation group:

$$R(\hat{n};\theta)(\hat{\imath} \cdot \vec{\mathfrak{J}}) R^{-1}(\hat{n};\theta) = e^{\theta \hat{n} \cdot \vec{\mathfrak{J}}} \ (\hat{\imath} \cdot \vec{\mathfrak{J}}) e^{-\theta \hat{n} \cdot \vec{\mathfrak{J}}} \quad = \sum_{n=0}^{\infty} \frac{\theta^n}{n!} (\mathrm{ad}\, \hat{n} \cdot \vec{\mathfrak{J}})^n \, (\hat{\imath} \cdot \vec{\mathfrak{J}}) \ . \tag{3.15}$$

Using explicitly the cross product relationship for the commutator, Eq. (3.10), and the procedure for multiple products already used earlier, we first see that the infinite series allows an iteration procedure:

$$(\mathrm{ad}\, \hat{n} \cdot \vec{\mathfrak{J}})^2 \, (\hat{\imath} \cdot \vec{\mathfrak{J}}) = [\hat{n} \cdot \vec{\mathfrak{J}}, \Big(\hat{n} \times \hat{\imath}\Big) \cdot \vec{\mathfrak{J}}] = \Big(\hat{n} \times (\hat{n} \times \hat{\imath})\Big) \cdot \vec{\mathfrak{J}} \ ,$$

$$(\mathrm{ad}\, \hat{n} \cdot \vec{\mathfrak{J}})^3 \, (\hat{\imath} \cdot \vec{\mathfrak{J}}) = \hat{n} \times \Big((\hat{n} \cdot \hat{\imath})\hat{n} - 1\hat{\imath}\Big) \cdot \vec{\mathfrak{J}} = -\Big(\hat{n} \times \hat{\imath}\Big) \cdot \vec{\mathfrak{J}} = -(\mathrm{ad}\, \hat{n} \cdot \vec{\mathfrak{J}})^1 \, (\hat{\imath} \cdot \vec{\mathfrak{J}}) \ .$$

This then allows the infinite series to be explicitly summed, giving

$$R(\hat{n};\theta)(\hat{\imath}\cdot\vec{\mathbf{J}})R^{-1}(\hat{n};\theta) = \hat{\imath}\cdot\vec{\mathbf{J}}\,\cos\theta + (\hat{n})\cdot\hat{\imath}(\hat{n}\cdot\vec{\mathbf{J}})(1-\cos\theta) + \hat{n}\times\hat{\imath}\cdot\vec{\mathbf{J}}\sin\theta\;. \tag{3.16}$$

Comparing this result with the general, 3-dimensional expression for the action of a rotation on a vector, as given in Eq. (2.3b), we see that this result is identical with that already given via Proof No. 1, in Eq. (3.7). However, since it has been shown for arbitrary matrices, with only the requirement that they satisfy the commutation relations given in Eqs. (3.11), we may then use this result later on, for any such triplet of matrices. Therefore, we will soon proceed to looking at more general such triplets.

As a concluding thought here, however, we will use this result to consider the action of the group on itself, i.e., to consider $R_1^{-1}R_2R_1$, for any two rotations R_1 and R_2, where, for later convenience, we have switched the order of R_1 and R_1^{-1} from our equations above Referring back to our listing of properties of the exponentials of matrices, we can re-write this concept in terms of the appropriate generators,:

$$\begin{aligned} R^{-1}(\hat{n};\theta)\,R(\hat{\imath};\zeta)\,R(\hat{n};\theta) =& e^{-\theta\hat{n}\cdot\vec{\mathbf{J}}}\;e^{\zeta\hat{\imath}\cdot\vec{\mathbf{J}}}\;e^{+\theta\hat{n}\cdot\vec{\mathbf{J}}} \;=\; e^{\zeta[R^{-1}(\hat{n};\theta)(\hat{\imath}\cdot\vec{\mathbf{J}}\;)R(\hat{n};\theta)]} \\ &= e^{\zeta[R(\hat{n};\theta)\hat{\imath}]\cdot\vec{\mathbf{J}}} \;= R[R(\hat{n};\theta)\hat{\imath};\zeta]\;, \end{aligned} \tag{3.17}$$

which, in English tells us that the action of R_1 on R_2 is to maintain the angle appropriate to R_2 but to change its axis, by using R_1 to rotate it! (A very reasonable and pleasant physical happening.)

IV. Representations of the Rotation Group A representation of a group is a (continuous) mapping that sends each element of the group into a continuous linear operator that acts on some vector space, **and** which preserves the group product. An irreducible representation is one which does not leave invariant any proper subspace of the vector space on which it acts. The rotation group, as an example of a compact group, has some nice properties for its irreducible representations:

a. They are finite-dimensional.

b. The matrices in question are all unitary. In particular, one can recall, from quantum mechanics classes, perhaps, that there is exactly one irreducible representation for each finite dimension, n. It is customary to label these representations not by n, but by the value of j, where $n \equiv 2j+1$; the value of j is often called the "spin" of the representation.

c. Infinite-dimensional representations, created from infinite sums or integrals of irreducible ones, are composed on unitary operators.

We will denote an arbitrary (unitary) representation of the rotation group, by U, so that the linear operators which are its values can be denoted by $U(R)$, or $U[R(\hat{n};\theta)]$, or, perhaps, as a sometimes-shorthand, by just $U(\hat{n};\theta)$. A representation of the group can be used to determine, uniquely, a representation of the Lie algebra of that group, since the Lie algebra are just the generators of the group near the identity. Denoting the generators by $\overline{\mathfrak{J}}_i \equiv -i\,\overline{J}_i$, we can write, for any choice of axis, $\hat{n}$,

$$U[R(\hat{n};\theta)] = e^{-i\theta\hat{n}\cdot\overline{\vec{J}}}\ . \tag{4.1}$$

$$\hat{n}\cdot\overline{\vec{J}} = \lim_{\theta\to 0}\frac{i}{\theta}\Big\{U[R(\hat{n};\theta)] - I\Big\} = \frac{d}{d\theta}U[R(\hat{n};\theta)]\bigg|_{\theta=0}\ . \tag{4.2}$$

However, it is quite important to point out that a given representation of a Lie algebra does **not** uniquely determine the group that it generates. For instance, integer values of j do indeed generate representations of the rotation group; however, half-odd-integer values of j generate representations in which the product rule is preserved **only** to within a plus or a minus sign! In fact, the half-odd-integer values of j generate representations of a different group, namely **SU(2)**. The relation between the two is that **SU(2)** is *the universal covering group* for **SO(3)**.

For an irreducible representation, a standard convention for choosing a basis for the n-dimensional vector space is given by the following set of $n = 2j+1$ vectors:

$$\Big\{|jm\rangle \mid m = -j, -j+1, -j+2,\ \ldots\ , j-2, j-1, j\,\Big\}\ . \tag{4.3}$$

Denoting the (Hermitean) operators on this space that represent the generators, i.e., the basis of the Lie algebra for **SO(3)**, by $J_i^{(j)}$, these vectors are mutual eigenvectors of the matrices $J_z^{(j)}$

and the Casimir operator, $(\vec{J}^{(j)})^2$, of course therefore presenting those operators as diagonal matrices:

$$(\vec{J}^{(j)})^2|jm\rangle = j(j+1)|jm\rangle\ , \qquad J_z^{(j)}|jm\rangle = m|jm\rangle\ . \tag{4.3}$$

As well, one also learns there the linear combinations, $J_\pm^{(j)}$, of J_x and J_y, can be thought of as "raising" and "lowering" operators in the sense that transform any given basis vector into one with a value of m that is either $+1$ higher, or lower:

$$\begin{gathered} J_\pm \ \equiv J_x \pm iJ_y\ , \\ J_+^{(j)}|jm\rangle = \sqrt{(j-m)(j+1+m)}|jm+1\rangle\ , \quad J_-^{(j)}|jm\rangle = \sqrt{(j+m)(j+1-m)}|jm-1\rangle\ . \end{gathered} \tag{4.4}$$

The matrices for the irreducible representations are conventionally denoted by the symbol $D^{(j)}[R(\hat{n};\theta)]$—sometimes written simply as $D^{(j)}(\hat{n};\theta)$—and are unitary. Their matrix elements can be given as

$$D^{(j)}(R)^{m'}{}_m \equiv \langle jm'|U(R)|jm\rangle = \langle jm'|e^{-i\theta\hat{n}\cdot\overline{\vec{J}}}|jm\rangle = \left(e^{-i\theta\hat{n}\cdot\vec{J}^{(j)}}\right)^{m'}{}_m\ . \tag{4.5}$$

1. **The case $\mathbf{j=1}$.**

Using the formulae above, we easily find that

$$J_z^{(1)} \Longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}\ , \ J_x^{(1)} \Longrightarrow \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\ , \ J_y^{(1)} \Longrightarrow \frac{i}{\sqrt{2}}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}\ . \tag{4.6}$$

Since these are 3×3 matrices, and there is only one irreducible representation for each dimension, one might wonder why these matrices are distinct from the 3×3 matrices given in Eqs. (1.3) or Eqs. (1.5). The answer is that they are equivalent, via a similarity transformation, to those others. The matrices in §1 are relative to a Cartesian basis set. On the other hand, these are relative to a basis in which J_z is diagonal. Finding the eigenvectors for J_z, and making the standard choices about signs and normalizations, one finds the following relationships:

$$|10\rangle = \hat{z}\ , \quad |11\rangle = \frac{-1}{\sqrt{2}}(\hat{x}+i\hat{y})\ , \quad |1-1\rangle = \frac{+1}{\sqrt{2}}(\hat{x}-i\hat{y})\ . \tag{4.7}$$

2. **The case $\mathbf{j} = \frac{1}{2}$.**

In this case, the vector space is only 2-dimensional, and the representations of the generators are

$$J_z^{(\frac{1}{2})} = \tfrac{1}{2}\begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix} , \quad J_x^{(\frac{1}{2})} = \tfrac{1}{2}\begin{pmatrix} 0 & +1 \\ +1 & 0 \end{pmatrix} , \quad J_y^{(\frac{1}{2})} = \tfrac{1}{2}\begin{pmatrix} 0 & -i \\ +i & 0 \end{pmatrix} , \tag{4.8}$$

where these three matrices are just half the standard three *Pauli matrices*, usually called σ_i, although R. Cahn's book often refers to them as τ_i. By explicit calculation, one finds that

$$D^{(\frac{1}{2})}[R(\hat{n};\theta)] = \cos(\theta/2)\, I_2 - i\hat{n}\cdot\vec{J}^{(\frac{1}{2})}\,\sin(\theta/2)\ . \tag{4.9}$$

It is therefore immediately clear that

$$D^{(\frac{1}{2})}(\hat{n};2\pi) = -I_2\ , \quad \forall \hat{n}\ , \tag{4.10}$$

which makes it clear that this is not a representation of $\mathbf{SO(3)}$.

The general notion of a *ray representation* is one where a phase is allowed to "creep into" the representation of the product, as follows:

$$\check{U}(R_1)\check{U}(R_2) = e^{i\eta(R_1,R_2)}\check{U}(R_1R_2)\ , \tag{4.11}$$

where η is a real-valued function of the two rotations. Since this "extra" factor has an absolute value of $+1$, in quantum mechanics it will never be measured. Therefore, there is a general theorem, due to Eugene Wigner, that quantum mechanics "allows," or even "wants," ray representations rather than the more usual sort of representations. As well, Wigner showed that these ray representations, of a group G, can always be normalized so that they are true representations of the *universal covering group* of G, denoted $\overline{G}$. In general, as well, one can show that the universal covering group is the unique group that has a homomorphism from $\overline{G}$ to G and is simply connected. For the rotation group, $\mathbf{SO(3)}$, the universal covering group is $\mathbf{SU(2)}$, the group of unitary, 2×2 matrices with determinant $+1$. The homomorphism in question, from $g \in \mathbf{SU(2)}$ to $R(g) \in \mathbf{SO(3)}$, is given by the following equation:

$$[R(g)]^a{}_b = \tfrac{1}{2}\,\mathrm{tr}\,\{g^\dagger\sigma^a g\sigma_b\}\ . \tag{4.12}$$

This equation is obviously (at least) two to one, since if $g \in \mathbf{SU(2)}$ then $-g$ is also a member of $\mathbf{SU(2)}$. But since the expression is quadratic in g, then both g and $-g$ give the same values for the matrix elements of $R(g)$.

A useful visualization of the manifold for $\mathbf{SU(2)}$ can be given in the following way. First, notice that an arbitrary element of $\mathbf{SU(2)}$ can be written in the form

$$g \in \ \mathbf{SU(2)} \quad \Longrightarrow \quad g \longrightarrow \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} , \quad \text{where } \ |a|^2 + |b|^2 = +1 \ . \tag{4.13}$$

Given the constraint on the complex numbers a and b, we have three real degrees of freedom available, as expected, since there is a 2-1 homomorphism onto $\mathbf{SO(3)}$. Therefore, the set $\{\mathrm{Re}\ a, \mathrm{Im}\ a, \mathrm{Re}\ b, \mathrm{Im}\ b\}$, subject to the constraint $(\mathrm{Re}\ a)^2 + (\mathrm{Im}\ a)^2 + (\mathrm{Re}\ b)^2 + (\mathrm{Im}\ b)^2 = +1$, constitutes a set of free real parameters to describe the group that are simply constrained to lie on the surface of the sphere S^3, the 3-dimensional sphere in 4-dimensional flat space.

V. Representations on Function Spaces

In many subject areas, it is common to use vectors whose components are functions to describe physical systems. The simplest case is of course just scalar functions, i.e., a 1-dimensional vector space, but with a functional dependence for the (single) component, say as a function of location $\vec{r}$. In the case of single, scalar functions of location, say $f(\vec{r})$, when the coordinates are rotated, i.e., $\vec{r} \to R(\hat{n};\theta)\vec{r}$, the behavior of the function of $\vec{r}$ will depend on the physical meaning of the function. I will only consider the behavior of wave functions, $\psi(\vec{r})$, or amplitude functions, which are such that the squares of their absolute values are local probability distribution functions. In quantum mechanics, one usually thinks of these functions both in terms of their behavior as functions of location, and also in abstract terms as being vectors in some Hilbert space, which possesses a scalar product and is closed with respect to convergent sequences. Then, the explicit functions of $\vec{r}$ are simply the set of all components of that vector with respect to one particular choice of "basis" in that vector space. There would of course be many other choices of basis, all related, as usual, by a transformation that changes one basis set into another.

A very standard way to describe this dual viewpoint is to use the Dirac *ket* symbol, with the letter indicating the function in question inside, i.e, $|\psi\rangle$, as a symbol for the actual vector in the Hilbert space. The ket $|\vec{r}\rangle$ is then a symbol for a particular element of that set of basis vectors. Dirac used *bra* vectors, $\langle\eta|$, for example, to denote elements of the dual space to the ket vectors. Then the general scalar product of two vectors can be written as $\langle\eta|\psi\rangle$, and the components of $|\psi\rangle$ are given by the appropriate scalar products, i.e., $\psi(\vec{r}) \equiv \langle\vec{r}|\psi\rangle$. As a basis set which is not square-integrable, the orthonormality relations for $\{|\vec{r}\rangle \mid \vec{r} \in \mathbb{R}^3\}$ involve Dirac distributions, usually just called Dirac delta "functions,"

$$\langle\vec{r}|\vec{r}'\rangle = \delta(\vec{r} - \vec{r}') \; . \tag{5.1}$$

Being symmetries, rotations should transform our amplitude distribution functions so that the probabilities for being located within any (infinitesimal) volume element, $d\vec{r}$, at an arbitrary location $\vec{r}$, should be preserved. For a given wave function, $\psi(\vec{r})$, we denote the transformed wave function by $\psi_R(\vec{r})$ and the transformed volume element by $d\vec{r}_R$. This requirement may then be written as

$$|\psi(\vec{r})|^2 d\vec{r} = |\psi_R(\vec{r})|^2 d\vec{r}_R = |\psi_R(\vec{r})|^2 d\vec{r} \; . \tag{5.2}$$

The last equality follows because the Jacobian of the transformation is just the determinant of $R = +1$, which then leads to the requirement

$$|\psi_R(R\vec{r})|^2 = |\psi(\vec{r})|^2 \quad \Longleftrightarrow \quad |\psi_R(\vec{r})|^2 = |\psi(R^{-1}\vec{r})|^2 \; . \tag{5.3}$$

To be sure one understands the algebra, it is illustrative to think of the simple case where, for example, our function $\psi(\vec{r})$ is very strongly peaked at a particular place, say $\vec{r}_0$, then the rotated function should be very strongly peaked at the rotated place, i.e., $R\vec{r}_0$.

Requiring that the transformation of these functions be just the identity, i.e., no transformation, when the rotation is the identity, and also insisting that the transformation be a continuous function of the parameters describing the rotation, it is straightforward to prove—Weyl—that the requirements above are uniquely satisfied—for these **scalar** wave functions—by simply removing the square and the absolute values from both sides of the equation. In terms of the abstract Hilbert-space notation, we let $U(R)$ be a representation of the rotations via unitary operators on the Hilbert space, i.e.,

$$|\psi_R\rangle \equiv U(R)|\psi\rangle \equiv e^{-i\theta\hat{a}\cdot\overline{\vec{J}}}|\psi\rangle \underset{|\vec{r}\rangle}{\Longrightarrow} \quad \psi_R(\vec{r}) = \psi(R^{-1}\vec{r}) = \psi\{\vec{r} - \theta\hat{a}\times\vec{r} + O^2(\theta)\} \ . \tag{5.4}$$

Taking the derivative with respect to θ, evaluating at $\theta = 0$, and using Taylor's theorem to approximate the behavior of the function, we find the associated representation for the generator of this rotation:

$$\overline{\vec{J}} = -i\,\vec{r}\times\nabla_{\vec{r}} \ . \tag{5.5}$$

When written out in terms of spherical coordinates, these three operators involve only θ and φ, excluding r; therefore, it is common to think of them in terms of just the direction of some location, i.e., the direction $\hat{r}$, that has those values of θ and φ.

We also know that a maximal commuting set of operators associated with this triplet may be chosen to be the pair $\{\overline{J_z}, \overline{\vec{J}}^2\}$. The eigenvalue problem for this pair is solved by the spherical harmonics, defined as functions of $\hat{r}$, i.e., functions on the sphere of a constant radius, by

$$\overline{J_z}\,Y_{\ell m}(\hat{r}) = m\,Y_{\ell m}(\hat{r}) \ , \qquad \overline{\vec{J}}^2\,Y_{\ell m}(\hat{r}) = \ell(\ell+1)\,Y_{\ell m}(\hat{r}) \ , \tag{5.6}$$

which satisfy the following two orthonormality and closure conditions:

$$\int d\hat{r}\,Y^*_{\ell' m'}(\hat{r})Y_{\ell m}(\hat{r}) = \delta_{\ell'\ell}\delta_{m'm} \quad ; \qquad \sum_{\ell m} Y^*_{\ell m}(\hat{r}')Y_{\ell m}(\hat{r}) = \delta(\hat{r}-\hat{r}') \ . \tag{5.7}$$

Because of these relations they can be taken as a basis for any (sufficiently-well-behaved) functions defined on the sphere, S^2, so that this decomposition may be written out, either in the component notation for functions of $\vec{r}$ or in the abstract Hilbert space notation:

$$\begin{aligned} \Upsilon(\vec{r}) &= \sum_{\ell m} f^{\ell m}(r)\,Y_{\ell m}(\hat{r}) \quad ; \qquad f^{\ell m}(r) = \int d\hat{r}\,Y_{\ell m}(\hat{r})^*\Upsilon(\vec{r}) \ , \\ |\Upsilon\rangle &= \sum_{\ell m} |f^{\ell m}\rangle \otimes |\ell m\rangle \quad ; \qquad |f^{\ell m}\rangle = \langle \ell m|\Upsilon\rangle = \int d\hat{r}\,\langle \ell m|\hat{r}\rangle\langle\hat{r}|\Upsilon\rangle \ , \end{aligned} \tag{5.8}$$

where we have used the notations

$$|r\rangle \otimes |\hat{r}\rangle \equiv |r\,\hat{r}\rangle = |\vec{r}\rangle = |r\,\theta\,\varphi\rangle\ , \quad \langle\hat{r}|\ell m\rangle = Y_{\ell m}(\hat{r})\ , \quad \langle r|f^{\ell m}\rangle = f^{\ell m}(r)\ . \tag{5.9}$$

As the spherical harmonics, for a fixed value of ℓ constitute elements of a vector space of dimension $2\ell + 1$, we can think of them as the $\vec{r}$-representation of our vector space on which acts an irreducible representation of the rotation group for fixed ℓ. To proceed further with that approach, we would define a new basis for our scalar functions,

$$|r\ell m\rangle \equiv \int d\hat{r}\; Y_{\ell m}(\hat{r})|\vec{r}\rangle\ . \tag{5.10}$$

More might be put here sometime, but not yet, as of 25 September, 2001.