Appendix G: Rotation Matrices and Related Theorems **1**
G.1 Generators and finite rotation matrices 1
G.2 About the general rotation matrix 3
G.3 The Baker-Campbell-Hausdorff and Sandwich Formulas 6
G.4 Two more theorems for the rotation matrix toolbox 9
G.5 Generalizations of the Rotation Group 11

This has been installed, do not edit here.

Appendix G: Rotation Matrices and Related Theorems

Here we present still more information on rotation matrices and demonstrate some typical manipulations done with such matrices.

G.1 Generators and finite rotation matrices

The general active rotation matrix for a rotation by angle θ about some **n** *unit vector* axis is given by,

$$R_{\hat{\mathbf{n}}}(\varphi) = \exp(-i\,\theta\,\mathbf{n}\bullet\mathbf{J})\ , \qquad \text{(G.1.1)}$$

where the J_k are 3x3 matrices known as the rotation **generator** matrices,

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \qquad J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad . \qquad \text{(G.1.2)}$$

The numbers in these three matrices can be summarized in this single statement,

$(J_a)_{bc} = -\,i\,\varepsilon_{abc}$ // for example, $(J_1)_{23} = -\,i\,\varepsilon_{123} = -i$

or

$$(iJ_a)_{bc} = \varepsilon_{abc} \qquad \text{(G.1.3)}$$

where the ε_{abc} permutation tensor is described in (1.5.3).

Defining the commutator of two square matrices X,Y as [X, Y] = XY-YX, one can show that the above three matrices satisfy this "commutation relation",

$$[J_i, J_j] = i\varepsilon_{ijk}J_k\ . \qquad \text{// } [J_1,J_2] = iJ_3 \text{ and cyclic} \qquad \text{(G.1.4)}$$

Proof: These easily demonstrated facts will be used in the proof (implied sum on repeated indices) ,

$$\varepsilon_{abc}\varepsilon_{ABc} = \delta_{aA}\delta_{bB} - \delta_{aB}\delta_{bA} \qquad \delta_{ab} \equiv \delta_{a,b} \qquad \varepsilon_{abc} = \varepsilon_{bca} = \varepsilon_{cab} \qquad \text{(G.1.5)}$$

$$LHS_{ac} = (J_i)_{ab}(J_j)_{bc} - (J_j)_{ab}(J_i)_{bc} = (-i)^2[\varepsilon_{iab}\varepsilon_{jbc} - \varepsilon_{jab}\varepsilon_{ibc}] = -[\varepsilon_{iab}\varepsilon_{cjb} - \varepsilon_{jab}\varepsilon_{cib}]$$

$$= -[(\delta_{ic}\delta_{aj} - \delta_{ij}\delta_{ac}) - (\delta_{jc}\delta_{ai} - \delta_{ji}\delta_{ac})] = \delta_{jc}\delta_{ai} - \delta_{ic}\delta_{aj} \quad .$$

$$RHS_{ac} = i\varepsilon_{ijk}(J_k)_{ac} = i\varepsilon_{ijk}[-i\varepsilon_{kac}] = \varepsilon_{ijk}\varepsilon_{kac} = \varepsilon_{ijk}\varepsilon_{ack} = \delta_{ia}\delta_{jc} - \delta_{ja}\delta_{ic} \, . \qquad \text{QED}$$

The expression shown in (G.1.1) involves the exponentiation of a square matrix to produce a new square matrix of the same dimension. This notion of exponentiating a matrix is straightforward as we demonstrate with a simple example:

$$R_x(\theta) = \exp(-i\,\theta\,\hat{\mathbf{x}} \bullet \mathbf{J}) = \exp(-i\,\theta J_1) \equiv \Sigma_{n=0}^{\infty} (-i\theta)^n (J_1)^n / n!$$

$$= 1 + \Sigma_{2,4,6..} (-i\theta)^n (J_1)^n / n! + \Sigma_{1,3,5..} (-i\theta)^n (J_1)^n / n!$$

It is easy to show that $(J_1)^n = J_1$ for odd n, while $(J_1)^n = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ for even n > 0. Thus,

$$R_x(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Sigma_{2,4,6..} (-i\theta)^n/n! + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \Sigma_{1,3,5..} (-i\theta)^n/n! \, .$$

But

$$\Sigma_{2,4,6..} (-i\theta)^n/n! = -\theta^2/2! + \theta^4/4! + ... = (\cos\theta - 1)$$
$$\Sigma_{1,3,5..} (-i\varphi)^n/n! = (-i\theta) + (-i\theta)^3/3! + ... = (-i)\,[\,\theta - \theta^3/3! + ..] = -i\sin\theta \qquad \text{(G.1.6)}$$

and therefore

$$R_x(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} (\cos\theta - 1) + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} (-i\sin\theta)$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & 0 \\ 0 & 0 & \cos\theta \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\sin\theta \\ 0 & \sin\theta & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}. \qquad \text{(G.1.7)}$$

The three axis-aligned rotations are found in this manner to be

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \quad R_y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \quad R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \, . \qquad \text{(G.1.8)}$$

The general rotation (G.1.1) applied to a vector **v** produces a **forward** right-hand-rule rotation of vector **v** by an angle θ about an arbitrary $\hat{\mathbf{n}}$ axis to create **v**' = R**v**. We call this an **active rotation**. For example, if

one applies $R_z(\theta)$ shown in (G.1.8) to the unit vector $\hat{\mathbf{x}} = (1,0,0)$ one gets $(\cos\theta,\sin\theta,0)$ which for small θ is a vector in the first quadrant of the x,y plane.

Some authors define rotation matrices which rotate vectors **backwards** according to the right hand rule, with the connection to our matrices then being $\theta \leftrightarrow -\theta$. The motivation for doing this is the fact that, whereas $\mathbf{v'} = R\mathbf{v}$ in the active view where the vector moves and the axes stay put, one can instead do $\mathbf{e'}_n = R^{-1}\mathbf{e}_n$ and have the vector stay put and the axes are back-rotated, which is the **passive** viewpoint. The definition of the rotation matrices is just a convention and (G.1.8) shows our definitions.

One says that the matrices J_i "generate" the finite rotations when they are exponentiated.

An alternative approach: First, one can show that $R_i(d\theta) \approx 1 - id\theta J_i$ describes a 3x3 matrix which, when applied to a vector, causes that vector to rotate by the small amount $d\theta$ about the i axis (an "infinitesimal rotation"). Setting $d\theta = \theta/n$, one shows that $\lim_{n\to\infty} (1-i[\theta/n]J_i)^n = \exp(-i\theta J_i)$. This last limit is analogous to $\lim_{n\to\infty}(1+x/n)^n = e^x$ for a scalar value x.

G.2 About the general rotation matrix

An explicit expression for the general rotation matrix

We present this subsection as a Reader Exercise with a set of steps. The result is stated in (G.2.8).

We are interested in the following general rotation matrix:

$$R_n(\theta) = \exp(-i\,\theta\,\mathbf{n}\bullet\mathbf{J}) \equiv \Sigma_{n=0}^{\infty} (-i\theta)^n (\mathbf{n}\bullet\mathbf{J})^n/n! \quad . \quad \mathbf{n} = \hat{\mathbf{n}} = \text{a unit vector} \qquad \text{(G.2.1)}$$

(a) using (G.1.3) and (G.1.5), show that $(\mathbf{n}\bullet\mathbf{J})^2 = T$ where $T_{ab} = (\delta_{ab} - n_a n_b)$. (G.2.2)

(b) show that

$$(\mathbf{n}\bullet\mathbf{J})^n = (\mathbf{n}\bullet\mathbf{J})\,T \qquad \text{for } n = 3,5,7,....$$
$$(\mathbf{n}\bullet\mathbf{J})^n = T \qquad \text{for } n = 2,4,6... \qquad \text{(G.2.3)}$$

(c) show that $\exp(-i\theta\mathbf{n}\bullet\mathbf{J}) = \cos(\theta\mathbf{n}\bullet\mathbf{J}) - i\sin(\theta\mathbf{n}\bullet\mathbf{J})$ where each term is defined by its series

(d) show that

$$\cos(\theta\mathbf{n}\bullet\mathbf{J}) = 1 + T(\cos\theta - 1)$$

$$\sin(\theta\mathbf{n}\bullet\mathbf{J}) = (\theta\mathbf{n}\bullet\mathbf{J}) + (\mathbf{n}\bullet\mathbf{J})T\,(\sin\theta - \theta) \qquad \text{(G.2.4)}$$

(e) show therefore that

$$\exp(-i\theta\mathbf{n}\bullet\mathbf{J}) = 1 + T\,(\cos\theta - 1) - i\,(\mathbf{n}\bullet\mathbf{J})\{\,\theta + T\,(\sin\theta - \theta)\,\} \qquad \text{(G.2.5)}$$

(f) show that the terms linear in θ all cancel leaving this simpler result

$$\exp(-i\theta \mathbf{n}\bullet\mathbf{J}) = 1 + (\cos\theta - 1)\, T + \sin\theta[\, -i(\mathbf{n}\bullet\mathbf{J})] \qquad \text{(G.2.6)}$$

Hint: $(\mathbf{n}\bullet\mathbf{J})_{ab} n_b n_c = n_i(-i\varepsilon_{iab}) n_b n_c = -i n_c [\, \varepsilon_{iab} n_i n_b] = -i n_c [0] = 0$ since [antisym x sym].

(g) show that

$$T = \begin{pmatrix} 1-n_1^2 & -n_1 n_2 & -n_1 n_3 \\ -n_1 n_2 & 1-n_2^2 & -n_2 n_3 \\ -n_1 n_3 & -n_2 n_3 & 1-n_3^2 \end{pmatrix} \qquad -i(\mathbf{n}\bullet\mathbf{J}) = \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix} \qquad \text{(G.2.7)}$$

giving this final result,

$$R_{\mathbf{n}}(\theta) = \exp(-i\theta \mathbf{n}\bullet\mathbf{J}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (\cos\theta - 1) \begin{pmatrix} 1-n_1^2 & -n_1 n_2 & -n_1 n_3 \\ -n_1 n_2 & 1-n_2^2 & -n_2 n_3 \\ -n_1 n_3 & -n_2 n_3 & 1-n_3^2 \end{pmatrix} + \sin\theta \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix} . \qquad \text{(G.2.8)}$$

Note that the first two terms are symmetric matrices, while the third is antisymmetric.

(h) verify the special cases shown in (G.1.8) above. These cases have **n** = (1,0,0), (0,1,0, (0,0,1).

Finding **n** and θ

Problem: One is handed a 9-element rotation matrix R and one wants to find **n**,θ so R = exp(-iθ**n**•**J**) .

Solution: We are unaware of a single magic formula that solves this problem, so we use brute force.

One can break R into the sum of two pieces, one symmetric and one antisymmetric by constructing

$$S_{ab} \equiv (R_{ab}+R_{ba})/2 \qquad A_{ab} \equiv (R_{ab}-R_{ba})/2 \qquad R_{ab} = S_{ab} + A_{ab} \qquad R = S + A\ . \qquad \text{(G.2.9)}$$

Looking at (G.2.8) one then has,

$$S = \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{22} & S_{23} \\ S_{31} & S_{23} & S_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (\cos\theta - 1) \begin{pmatrix} 1-n_1^2 & -n_1 n_2 & -n_1 n_3 \\ -n_1 n_2 & 1-n_2^2 & -n_2 n_3 \\ -n_1 n_3 & -n_2 n_3 & 1-n_3^2 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & A_{12} & A_{13} \\ -A_{12} & 0 & A_{23} \\ -A_{13} & -A_{23} & 0 \end{pmatrix} = \sin\theta \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix} \qquad \text{(G.2.10)}$$

which can be written as these nine equations

$$A_{12} = -n_3\sin\theta \qquad S_{11} = 1 + (\cos\theta-1)(1-n_1^2) \qquad S_{12} = (\cos\theta-1)(-n_1n_2)$$
$$A_{13} = n_2\sin\theta \qquad S_{22} = 1 + (\cos\theta-1)(1-n_2^2) \qquad S_{13} = (\cos\theta-1)(-n_1n_3)$$
$$A_{23} = -n_1\sin\theta \qquad S_{33} = 1 + (\cos\theta-1)(1-n_3^2) \qquad S_{23} = (\cos\theta-1)(-n_2n_3) \ . \qquad \text{(G.2.11)}$$

• If $\theta = 0$, then the original matrix must be R = 1 and then we have no work to do.

• If $\theta = \pi$, the entire A matrix vanishes according to (G.2.10) leaving only S. The equations above are then

$$S_{11} = 1 + (-2)(1-n_1^2) \qquad S_{12} = 2n_1n_2$$
$$S_{22} = 1 + (-2)(1-n_2^2) \qquad S_{13} = 2n_1n_3$$
$$S_{33} = 1 + (-2)(1-n_3^2) \qquad S_{23} = 2n_2n_3 \ .$$

Since $n_1^2 + n_2^2 + n_3^2 = 1$, not all three n_i can vanish. The left column of equations says

$$n_i^2 = 1 - (1-S_{ii})/2 \qquad i = 1,2,3 \ .$$

Inspect the three n_i^2 and select one n_r^2 which is non-zero. Then select the plus sign to get,

$$n_r = +\sqrt{1 - (1-S_{rr})/2} \ .$$

To be specific, assume $n_r = n_1$. Then from the right column, $n_2 = S_{12}/(2n_1)$ and $n_2 = S_{13}/(2n_1)$ and we are done. So at least one of the following solutions must be viable:

$$\theta = \pi \qquad n_1 = \sqrt{1 - (1-S_{11})/2} \qquad n_2 = S_{12}/(2n_1) \qquad n_3 = S_{13}/(2n_1)$$
$$\theta = \pi \qquad n_2 = \sqrt{1 - (1-S_{22})/2} \qquad n_1 = S_{12}/(2n_2) \qquad n_3 = S_{23}/(2n_2)$$
$$\theta = \pi \qquad n_3 = \sqrt{1 - (1-S_{33})/2} \qquad n_1 = S_{13}/(2n_3) \qquad n_2 = S_{23}/(2n_3) \qquad \text{(G.2.12)}$$

• If $\theta \neq 0$ or π, then note from (G.2.10) that

$$A_{12}^2 + A_{13}^2 + A_{23}^2 = \sin^2\theta$$
$$S_{11} + S_{22} + S_{33} = 1 + 2\cos\theta \ . \qquad \text{(G.2.13)}$$

Using the sign ambiguity in the direction of **n**, we can assume that $0 < \theta < \pi$ so $\sin\theta > 0$. Then from (G.2.13) and the left column of (G.2.11) we have this solution,

$$\sin\theta = \sqrt{A_{12}^2 + A_{13}^2 + A_{23}^2}$$
$$\cos\theta = [\, S_{11} + S_{22} + S_{33} - 1]/2$$
$$n_1 = -A_{23}/\sin\theta$$
$$n_2 = A_{13}/\sin\theta$$
$$n_3 = -A_{12}/\sin\theta \qquad \text{(G.2.14)}$$

and thus a viable θ and **n** have been found such that $R = \exp(-i\theta\mathbf{n}\bullet\mathbf{J})$.

Example: Let $R = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$ so then $S = \begin{pmatrix} \cos\alpha & 0 & 0 \\ 0 & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $A = \begin{pmatrix} 0 & -\sin\alpha & 0 \\ \sin\alpha & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

$\sin\theta = \sqrt{A_{12}^2 + A_{13}^2 + A_{23}^2} = \sqrt{\sin\alpha^2 + 0 + 0} = \sin\alpha$

$\cos\theta = [S_{11} + S_{22} + S_{33} - 1]/2 = [\cos\alpha + \cos\alpha + 1 - 1]/2 = \cos\alpha \quad \Rightarrow \theta = \alpha$

$n_1 = -A_{23}/\sin\theta = 0$

$n_2 = A_{13}/\sin\theta = 0$

$n_3 = -A_{12}/\sin\theta = \sin\alpha/\sin\theta = 1 \quad \Rightarrow \quad \mathbf{n} = (0,0,1)$

Therefore $R = \exp(-i\alpha J_3) = R_z(\alpha)$, as verified in (G.1.8).

G.3 The Baker-Campbell-Hausdorff and Sandwich Formulas

Statement and proof of the Baker-Campbell-Hausdorff (BCH) formula

We now quote a fascinating fact known as the Baker-Campbell-Hausdorff formula involving two square matrices A and B, where commutator $[X,Y] \equiv XY-YX$,

$$e^{-A}Be^{A} = B + [B,A]/1! + [[B,A],A]/2! + [[[B,A],A],A]/3! + \ldots\ldots \qquad (G.3.1)$$

Outline of BCH proof: LHS = RHS

(a) Show that LHS $= \Sigma_{n=0}^{\infty}\Sigma_{m=0}^{\infty} (-A)^n BA^m / (n!m!)$.

(b) Set $k = n+m$ to rewrite as LHS $= \Sigma_{k=0}^{\infty}\Sigma_{n=0}^{k} (-A)^n BA^{k-n} / (n![k-n]!)$.

(c) Rewrite again as LHS $= \Sigma_{k=0}^{\infty} T_k/k!$ where $T_k \equiv \Sigma_{n=0}^{k} \binom{k}{n} (-A)^n BA^{k-n}$.

(d) Define $C_0 = B$, $C_1 = [B,A]$, $C_2 = [[B,A],A]$, etc. so that RHS $= \Sigma_{k=0} C_k/k!$.
Note that $[C_k,A] = C_{k+1}$.

The proof LHS = RHS is complete if one can show that $T_k = C_k$.

(e) Show $T_k = C_k$ by induction: show $T_0 = C_0$ and $T_k = C_k \Rightarrow T_{k+1} = C_{k+1}$. QED

Statement and proof of the Sandwich Formula

Now *using* this BCH formula, along with the commutation relation (G.1.4) that $[J_i, J_j] = i\varepsilon_{ijk}J_k$, one can show that

$$\exp(-i\theta \mathbf{n}\bullet\mathbf{J})\ \mathbf{J}\ \exp(+i\theta \mathbf{n}\bullet\mathbf{J}) = \cos\theta\ \mathbf{J} + \sin\theta\ \mathbf{J} \times \mathbf{n} + (1 - \cos\theta)\ \mathbf{n}(\mathbf{n}\bullet\mathbf{J}) . \qquad (G.3.2)$$

This "vector of matrices" notation is just a shorthand for the following equations for k = 1,2,3 :

$$\exp(-i\theta \mathbf{n}\bullet\mathbf{J})\ J_k \exp(+i\theta \mathbf{n}\bullet\mathbf{J}) = \cos\theta J_k + \sin\theta[\mathbf{J} \times \mathbf{n}]_k + (1-\cos\theta)\, n_k (\mathbf{n}\bullet\mathbf{J})$$

$$= \cos\theta\, J_k + \sin\theta\, \varepsilon_{kms} n_s J_m + (1-\cos\theta)\, n_k\ (\mathbf{n}\bullet\mathbf{J})\ . \qquad \text{(G.3.3)}$$

This is the "sandwich formula" since J_k on the left is sandwiched between two rotations.

Outline of Sandwich proof: LHS = RHS

(a) We will use the BCH formula with $A = i\theta\ \mathbf{n}\bullet\mathbf{J}$ and $B = J_k$. First, define C_k as in (d) above.

(b) show that $C_0 = J_k$, $C_1 = -\theta n_i \varepsilon_{kij} J_j$, $C_2 = \theta^2 (n_k n_i J_i - J_k)$, $C_3 = -\theta^2 C_1$, $C_4 = -\theta^2 C_2$

(c) deduce (or use induction to show) that in general,

$$C_n = -(-1)^{n/2}\ \theta^n\ (\ n_k\ n_i\ J_i\ - J_k\ ) \qquad n = 2,4,6....$$

$$C_n = -(-1)^{(n-1)/2}\ \theta^n\ n_i\ \varepsilon_{kij}\ J_j \qquad n = 1,3,,5...$$

At this point we have from the BCH formula,

$$\exp(-i\theta \mathbf{n}\bullet\mathbf{J})\ J_k \exp(+i\theta \mathbf{n}\bullet\mathbf{J}) = \Sigma_{n=0}^{\infty} C_n/n! \quad \text{with } C_n \text{ as in (c) above}$$

(d) Show that $\Sigma_{n=0}^{\infty} C_n/n! =$

$$J_k - n_i\ \varepsilon_{kij}\ J_j\ \Sigma_{n=1,3,5..}\ (-1)^{(n-1)/2}\ \theta^n/n! - (\ n_k\ n_i\ J_i\ - J_k\ )\ \Sigma_{n=2,4,6..}\ (-1)^{n/2}\ \theta^n/n!$$

$$= J_k - n_i\ \varepsilon_{kij}\ J_j\ [\theta - \theta^3/3!\ + ...] - (\ n_k\ n_i\ J_i\ - J_k\ )\ [\ -\theta^2/2! + \theta^4/4! + ...\ ]$$

$$= J_k - n_i\ \varepsilon_{kij}\ J_j\ \sin\theta - (\ n_k\ n_i\ J_i\ - J_k\ )(\cos\theta - 1)$$

$$= J_k\cos\theta + \varepsilon_{kji}\ J_j\ n_i\ \sin\theta + n_k\ (\mathbf{n}\bullet\mathbf{J})(1-\cos\theta)\ . \qquad \text{QED}$$

Comments

A vector **v** (rank-1 tensor) transforms ("rotates") according to $\mathbf{v}' = R\mathbf{v}$ (Active View) .
For a matrix M (rank-2 tensor) the corresponding transformation is $M' = RMR^{-1}$, and this is what one sees on the left side of (G.3.3) where $M = J_k$ and $R = \exp(-i\theta\ \hat{\mathbf{n}} \bullet \mathbf{J})$. In the expression RJ_kR^{-1} the rotation generator J_k is "sandwiched" between the two rotations.

The above sandwich formulas play a major role in Magnetic Resonance Imaging. The connection is that protons in your body have magnetic moments (spins) which can be lined up by a strong magnetic field. When the proton spins are slammed with a certain radio frequency pulse, they do conical rotation

(precession) about the magnetic field axis at the so-called Larmor frequency. After the pulse this proton precession decays away (time T1) and bulk-decoheres (time T2) producing a certain return RF signal which can be analyzed. These return signals are sensitive to the local environment of the protons. The location of a particular response is determined by giving the magnetic field a spatial gradient which affects the Larmor frequency. In this manner, an image can be formed. The sandwich formulas are not applied directly to individual spin angular momenta **J**, but to the average spin (polarization) density in the object being scanned. The analysis is quite complicated since it must take into account thermal and statistical effects which are managed with the use of the density matrix formalism. See the excellent text of Levitt for all the details.

Special cases of the sandwich formula

We shall have our own purposes for the sandwich formulas in Appendix H. For rotations about the i axis we set

$$n_s = \delta_{s,i} \quad \Rightarrow \quad \mathbf{n} \bullet \mathbf{J} = J_i \qquad \text{and} \qquad \varepsilon_{kms} J_m n_s = \varepsilon_{kmi} J_m = \varepsilon_{ikm} J_m \ . \tag{G.3.4}$$

Then from the sandwich formula (G.3.3),

$$\exp(-i\theta \mathbf{n}\bullet\mathbf{J})\ J_k \exp(+i\theta \mathbf{n}\bullet\mathbf{J}) = \cos\theta\, J_k + \sin\theta\, \varepsilon_{kms} n_s J_m + (1 - \cos\theta)\, n_k (\mathbf{n}\bullet \mathbf{J}), \tag{G.3.3}$$

we find that (there is no implied sum on i in the rightmost term and $\varepsilon_{kmi} = \varepsilon_{ikm}$)

$$\exp(-i\theta J_i)\ J_k \exp(+i\theta J_i) = \cos\theta\, J_k + \sin\theta \varepsilon_{ikm} J_m + (1 - \cos\theta)\, \delta_{k,i}\, J_i$$

so then

$$R_i(\theta)\, J_k\, R_i(-\theta) = \cos\theta\, J_k + \sin\theta\, \varepsilon_{ikm}\, J_m + (1 - \cos\theta)\, \delta_{k,i}\, J_i \ . \tag{G.3.5}$$

In the case i = k we know that the left side is just J_k since everything then commutes. This is verified on the right since $\varepsilon_{iim} = 0$ and the other two terms then add up to J_k. So although obvious, we state:

$$R_k(\theta)\, J_k\, R_k(-\theta) = J_k \ . \qquad i = k \tag{G.3.6}$$

In the case $i \neq k$ the third term in (G.3.5) does not contribute and we have

$$R_i(\theta)\, J_k\, R_i(-\theta) = \cos\theta\, J_k + \varepsilon_{ikm} \sin\theta\, J_m \qquad i \neq k$$

We shall now construct a table of all the cases for which $i \neq k$: (G.3.7)

$$\begin{aligned}
R_1(\theta)\, J_2 R_1(-\theta) &= \cos\theta\, J_2 + \varepsilon_{12m} \sin\theta\, J_m = \cos\theta\, J_2 + \sin\theta\, J_3 \\
R_1(\theta)\, J_3 R_1(-\theta) &= \cos\theta\, J_3 + \varepsilon_{13m} \sin\theta\, J_m = \cos\theta\, J_3 - \sin\theta\, J_2 \\
\\
R_2(\theta)\, J_1 R_2(-\theta) &= \cos\theta\, J_1 + \varepsilon_{21m} \sin\theta\, J_m = \cos\theta\, J_1 - \sin\theta\, J_3 \\
R_2(\theta)\, J_3 R_2(-\theta) &= \cos\theta\, J_3 + \varepsilon_{23m} \sin\theta\, J_m = \cos\theta\, J_3 + \sin\theta\, J_1
\end{aligned}$$

$$R_3(\theta)\, J_1 R_3(-\theta) = \cos\theta\, J_1 + \varepsilon_{31m} \sin\theta\, J_m = \cos\theta J_1 + \sin\theta\, J_2$$
$$R_3(\theta)\, J_2 R_3(-\theta) = \cos\theta\, J_2 + \varepsilon_{32m} \sin\theta\, J_m = \cos\theta J_2 - \sin\theta\, J_1 \ . \qquad \text{(G.3.8)}$$

G.4 Two more theorems for the rotation matrix toolbox

Theorem 1: $R\mathbf{J}R^{-1} = R^{-1}\mathbf{J}$ (G.4.1)

This is another vector/matrix notation theorem which makes a claim about rotating a vector of matrices. The above ambiguous notation is a shorthand for the following,

$$RJ_iR^{-1} = [R^{-1}\mathbf{J}]_i = R^{-1}{}_{ij}J_j \qquad i = 1,2,3\ . \qquad \text{(G.4.2)}$$

The object on the left is a product of three 3x3 matrices, while the right side is a linear combination of 3x3 matrices, so at least the theorem's claim is dimensionally reasonable. More generally J_i might be an abstract "operator" and R a rotation which acts on that operator, see Section G.5.

Proof: Start with the general rotation form given in (G.2.6),

$$R = \exp(-i\theta\mathbf{n}\bullet\mathbf{J}) = 1 + (\cos\theta - 1)\, T + \sin\theta[\, -i(\mathbf{n}\bullet\mathbf{J})] \quad \text{where } T_{ab} = (\delta_{ab} - n_a n_b) \ . \qquad \text{(G.4.3)}$$

Sandwich this rotation around J_i and then use the sandwich formula (G.3.3),

$$RJ_iR^{-1} = \exp(-i\theta\mathbf{n}\bullet\mathbf{J})\ J_i \exp(+i\theta\mathbf{n}\bullet\mathbf{J})$$
$$= \cos\theta\, J_i + \sin\theta\, \varepsilon_{ijk}n_k J_j + (1 - \cos\theta)\, n_i\ (\mathbf{n}\bullet\mathbf{J})$$
$$= \cos\theta\, \delta_{ij}J_j + \sin\theta\, \varepsilon_{ijk}n_k J_j + (1 - \cos\theta)\, n_i n_j J_j$$
$$= [\, \cos\theta\, \delta_{ij} + \sin\theta\, \varepsilon_{ijk}n_k + (1 - \cos\theta)\, n_i n_j\,]\, J_j \ . \qquad \text{(G.4.4)}$$

Our theorem is proved if, comparing (G.4.2) and (G.4.4), we can show that

$$R^{-1}{}_{ij} = \cos\theta\, \delta_{ij} + \sin\theta\, \varepsilon_{ijk}n_k + (1 - \cos\theta)\, n_i n_j \ . \qquad \text{(G.4.5)}$$

Looking back at (G.4.3), one has

$$R^{-1} = \exp(+i\theta\mathbf{n}\bullet\mathbf{J}) = 1 + (\cos\theta - 1)\, T + \sin\theta[\, +i(\mathbf{n}\bullet\mathbf{J})]$$

so

$$R^{-1}{}_{ij} = \delta_{ij} + (\cos\theta - 1)\, T_{ij} + \sin\theta[+n_k\, (iJ_k)_{ij}]$$
$$= \delta_{ij} + (\cos\theta - 1)\ (\delta_{ij} - n_i n_j) + \sin\theta\, n_k\, \varepsilon_{kij} \qquad \text{// (G.1.3) and def of } T_{ij}$$
$$= \delta_{ij}\cos\theta + (1-\cos\theta)\, n_i n_j + \sin\theta\, \varepsilon_{ijk}n_k \qquad \text{(G.4.6)}$$

But this is the same as (G.4.5) so the theorem is proved.

Theorem 2: $R\exp(-i\theta \mathbf{n}\bullet\mathbf{J})\,R^{-1} = \exp(-i\theta \mathbf{n}'\bullet\mathbf{J})$ where $\mathbf{n}' = R\mathbf{n}$ (G.4.7)

or $R\,R_{\mathbf{n}}(\theta)\,R^{-1} = R_{\mathbf{n}'}(\theta)$ where $\mathbf{n}' = R\mathbf{n}$

The matrix R is an arbitrary rotation. On the left we have a product of three 3x3 matrices, while the right side is a 3x3 matrix. One proof of this theorem might be to say "what else could it be? ". We shall provide a more substantial proof below.

Proof: Start by inserting the general rotation form (G.2.6) into the left side of (G.4.7),

$$R\exp(-i\theta \mathbf{n}\bullet\mathbf{J})\,R^{-1} = R\,[\,1 + (\cos\theta - 1)\,T - i\sin\theta\,(\mathbf{n}\bullet\mathbf{J})\,]\,R^{-1}$$

$$= 1 + (\cos\theta - 1)RTR^{-1} - i\sin\theta\; n_k\, RJ_kR^{-1}\,. \qquad \text{(G.4.8)}$$

Now consider

$$[RTR^{-1}]_{ad} = R_{ab}T_{bc}R^{-1}{}_{cd} = R_{ab}(\delta_{bc} - n_bn_c)R^{T}{}_{cd}$$

$$= \delta_{ad} - R_{ab}\,n_bn_cR_{dc} = \delta_{ad} - (\,R_{ab}\,n_b)(R_{dc}n_c) = \delta_{ad} - n'_an'_d \qquad \mathbf{n}' = R\mathbf{n}$$

$$\equiv T'_{ad}\,. \qquad \text{// in other words, } RTR^{-1} = T' \qquad \text{(G.4.9)}$$

From Theorem 1 (G.4.2) we know that

$$RJ_kR^{-1} = R^{-1}{}_{kj}J_j\,. \qquad \textit{(G.4.2)}$$

Inserting this last item and (G.4.9) into (G.4.8) gives

$$R\exp(-i\theta \mathbf{n}\bullet\mathbf{J})\,R^{-1} = 1 + (\cos\theta - 1)T' - i\sin\theta\; n_kR^{-1}{}_{kj}J_j$$

$$= 1 + (\cos\theta - 1)T' - i\sin\theta\, R_{jk}n_k\, J_j = 1 + (\cos\theta - 1)T' - i\sin\theta\,[Rn]_j\, J_j$$

$$= 1 + (\cos\theta - 1)T' - i\sin\theta\, n'_jJ_j = 1 + (\cos\theta - 1)T' - i\sin\theta\,(\mathbf{n}'\bullet\mathbf{J})$$

$$= \exp(-i\theta \mathbf{n}'\bullet\mathbf{J}) \qquad \text{// using (G.2.6) with } \mathbf{n}\rightarrow\mathbf{n}' \qquad \text{(G.4.10)}$$

Thus the theorem is proved.

The fact (G.4.9) that $RTR^{-1} = T'$ is in fact just the standard rule for the transformation of a rank-2 tensor under rotations, see (1.1.21) [also (J.24) and (J.25)]. That is to say, if an Observer in Frame S sees T, an Observer in rotated Frame S' will see $T' = RTR^{-1}$. This same statement applies at the higher level of our theorem

$$R\,R_{\mathbf{n}}(\theta)\,R^{-1} = R_{\mathbf{n}'}(\theta) \qquad \mathbf{n}' = R\mathbf{n} \qquad \text{// Theorem 2 restated} \tag{G.4.11}$$

The rotation $R_{\mathbf{n}}(\theta)$ seen by an Observer in Frame S appears as $R_{\mathbf{n}'}(\theta)$ in Frame S' where $\mathbf{e}'_{\mathbf{n}} = R^{-1}\mathbf{e}_{\mathbf{n}}$ as in (1.1.30). Like T, the rotation $R_{\mathbf{n}}(\theta)$ transforms as a rank-2 tensor under rotations.

Applying (G.4.11) to a vector **v**, one finds,

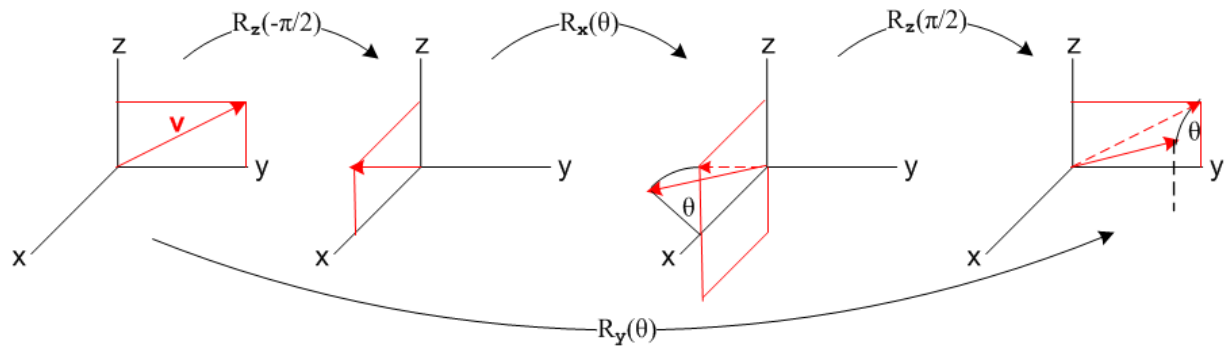
$$R\,R_{\mathbf{n}}(\theta)\,R^{-1}\mathbf{v} = R_{\mathbf{n}'}(\theta)\mathbf{v} \qquad \text{where } \mathbf{n}' = R\mathbf{n} \tag{G.4.12}$$

Instead of rotating a vector **v** by amount θ about axis $\hat{\mathbf{n}}'$, one can first back-rotate the vector **v** by R^{-1}, then rotate by amount θ about axis $\hat{\mathbf{n}}$, the forward rotate the result by R.

Example: $R_z(\pi/2)R_x(\theta)R_z(-\pi/2)$ = ?? Here $\mathbf{n} = \hat{\mathbf{x}}$ and $\mathbf{n}' = R_z(\pi/2)\,\hat{\mathbf{x}} = \hat{\mathbf{y}}$ so we conclude that

$$R_z(\pi/2)R_x(\theta)R_z(-\pi/2) = R_y(\theta)\ . \tag{G.4.13}$$

Here is a graphical interpretation of (G.4.13) applied to a particular vector **v** :



(G.4.14)

G.5 Generalizations of the Rotation Group

Here we consider some generalizations of the ideas presented above in Section G.1 .

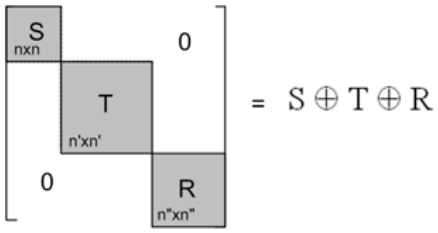
N dimensional representations of the rotation group

The generators J_i shown in (G.1.2) are a special case of a more general idea which starts with the commutation relation

$$[J_i, J_j] = i\varepsilon_{ijk}J_k\ . \qquad \text{// for example, } [J_1,J_2] = iJ_3 \qquad \textit{(G.1.4)} \tag{G.5.1}$$

One first thinks of the J_i as abstract "operators" in some abstract "operator space". One can show that it is possible to find a set of three NxN matrices of any integer dimension N which satisfy (G.5.1). The matrices are not unique, so (G.1.2) for the J_i in three dimensions is not unique, but it is a standard form.

The three NxN generator matrices J_i are said to form an N-dimensional "irreducible representation" of the abstract generators J_i. One can always create new viable generator matrices by taking a "direct sum" of existing viable generator matrices, such as in this block-diagonal-form picture



(G.5.2)

This generator matrix is "reducible" into a direct sum of S, T and R. An "irreducible" representation is one that cannot be reduced in this manner. The same comment applies to rotation matrices.

The commutation relation (G.5.1) is an example of a **Lie Algebra**. Our particular Lie Algebra is called so(3), so we can represent the algebra elements J_i of this algebra by three NxN matrices J_i. It is possible to write down a formula analogous to (G.1.3) $(iJ_a)_{bc} = \varepsilon_{abc}$ which works for any N, but (G.1.3) itself only applies to N = 3. This is so because ε_{abc} has no meaning for $N \neq 3$. But it always has meaning in (G.5.1) because there are only three generators regardless of the value of N.

For general N, the object $R_i(\theta) = \exp(-i\theta \mathbf{J}_i)$ is an NxN matrix which represents the action of a rotation of an N-vector in a Euclidean space E^N. The set of such rotation matrices forms an "irreducible representation" of the rotation group SO(3) in N dimensions. The integer N is usually written N = 2j+1 where j = 0, 1/2, 1, 3/2 ... and this j then serves as a label for a given matrix representation.

The value of j in N = 2j+1 is associated with "angular momentum" or "spin". In the case N=2 (having j=1/2) the generator matrices are the 2x2 "Pauli matrices". In this case the 2x2 matrices $\exp(-i\theta J_i)$ describe the rotations of spin-1/2 particles such as electrons or protons. Here are the details for N=2 :

$$J_x = (1/2)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad J_y = (1/2)\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad J_z = (1/2)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$R_x = \begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \qquad R_y = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \qquad R_z = \begin{pmatrix} \exp(-i\frac{\theta}{2}) & 0 \\ 0 & \exp(+i\frac{\theta}{2}) \end{pmatrix}$$

(G.5.3)

The "vectors" for spin-1/2 particles have two components $\begin{pmatrix} u \\ d \end{pmatrix}$. The special case $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is called "spin up" and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is "spin down".

For the Lie Algebra so(3) one can show that $J^2 \equiv J_1^2 + J_2^2 + J_3^2$ and any particular J_i commute with each other, so $[J^2, J_i] = 0$. J^2 is called a Casimir operator of this Algebra, and fancier Lie Algebras can have several such Casimirs.

A Differential operator representation of the rotation group

It is also possible to "represent" the three rotation generators J_i by three differential operators in spherical coordinates θ and φ. These operators satisfy $[J_i, J_j] = i\varepsilon_{ijk}J_k$ and in this context they are usually called L_i but we shall stick with J_i. In this case, one can compute the differential operator J^2 and one can ponder differential equations which take the form $J^2\, f_{jm}(\theta,\varphi) = j(j+1)\, f_{jm}(\theta,\varphi)$ and $J_z\, f_{jm}(\theta,\varphi) = m\, f_{jm}(\theta,\varphi)$. [The facts that the eigenvalue of J^2 is j(j+1) and not j^2, and that m runs from -j to j, derive from the structure of the Lie Algebra.] The solutions $f_{jm}(\theta,\varphi)$ are able to have well-defined eigenvalues j(j+1) and m *because* $[J^2,J_3] = 0$. If we instead had $[J^2,J_3] \neq 0$, then $[J^2,J_3]\, f_{jm}(\theta,\varphi) = j(j+1)m - mj(j+1) = 0$ is a contradiction and the two eigenfunction equations could not exist. The solutions $f_{jm}(\theta,\varphi)$ of these equations are called the **spherical harmonics** and are usually written $Y_{jm}(\theta,\varphi)$. Just for the record, here is what the differential operators look like, where C = cos and S = sin (for example, $S_\varphi = \sin\varphi$ and $\partial_\varphi = \partial/\partial\varphi$) :

$$
\begin{aligned}
J_1 &= i\,[\, S_\varphi\, \partial_\theta + \cot\theta\, C_\varphi\, \partial_\varphi] & J_\pm &= e^{\pm i\varphi}\, [\, \pm\partial_\theta + i\, \cot\theta\, \partial_\varphi\,] \\
J_2 &= -i\,[\, C_\varphi\, \partial_\theta - \cot\theta\, S_\varphi\, \partial_\varphi] & J^2 &= -[\, \partial_\theta^2 + \cot\theta\, \partial_\theta + (1/S_\theta)^2\, \partial_\varphi^2] \\
J_3 &= -i\, \partial_\varphi & J^2 &= -[\, (1/S_\theta)\, \partial_\theta\, [\, S_\theta\, \partial_\theta\,] + (1/S_\theta)^2\, \partial_\varphi^2\,] \qquad \text{(G.5.4)}
\end{aligned}
$$

Reader Exercise: Verify that the three operators on the left satisfy the Lie Algebra $[J_i, J_j] = i\varepsilon_{ijk}J_k$.

For the hydrogen atom with a spinless electron, there are three mutually commuting quantities H, J^2 and J_3 where H is the Hamiltonian. This means that the solution eigenfunctions can have well defined E, j and m values and these eigenfunctions are those painful "orbitals" appearing in chemistry books. When the Hamiltonian commutes with some other operator like J^2, that operator is called a "symmetry". Solution functions then bear a label for each such symmetry, such as j for J^2.

The Lie Algebra so(3) is isomorphic (one-to-one related) to another Lie Algebra called su(2).
The Lie Group SO(3) is isomorphic to another Lie Group called SU(2).

In the above we discuss only the Lie Group SO(3) with its three generators J_i. There are many other Lie Groups which have physics applications.

Some Other Lie Groups of Interest

The group SO(n) is the **orthogonal group** in n dimensions and it has n(n-1)/2 generators.

The group SO(3,1) is the **Lorentz Group** which has 6 generators J_i and K_i which generate 3 rotations and 3 "boosts" (velocity transformations). The Lie Algebra is this,

$$
\begin{aligned}
[\, J_i, J_j] &= +i\, \varepsilon_{ijk}J_k \\
[\, J_i, K_j] &= +i\, \varepsilon_{ijk}\, K_k \\
[\, K_i, K_j] &= -i\, \varepsilon_{ijk}J_k\ . \qquad \text{(G.5.5)}
\end{aligned}
$$

There are two Casimirs : $J^2 - K^2$ and $\mathbf{J}\bullet\mathbf{K}$.

In the Lorentz Group "vector representation" known as 1/2⊗1/2 the generators are represented as 4x4 matrices. When these are exponentiated, one obtains the finite rotation and boost matrices used in special relativity. For example, with space-time vectors ordered x^μ = (ct,x,y,z) one has

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(r) & -\sin(r) \\ 0 & 0 & \sin(r) & \cos(r) \end{pmatrix} = \exp(-irJ_1) \qquad \text{where } (J_1)^\mu{}_\nu = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \qquad \text{// rotation } R_x(r)$$

$$\begin{pmatrix} \cosh(b) & \sinh(b) & 0 & 0 \\ \sinh(b) & \cosh(b) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \exp(-ibK_1) \quad \text{where } (K_1)^\mu{}_\nu = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{// boost } B_x(b) \tag{G.5.6}$$

The **Poincare Group** is basically the Lorentz group SO(3,1) bolted onto the group T(4) of translations in four directions ct,x,y,z. It thus has 10 generators J_i, K_i and P_μ where these last four are momentum and energy. The Poincare algebra has two Casimir operators whose eigenvalues are associated with mass and spin. For example, the mass Casimir is $P^\mu P_\mu$. The irreducible representations of the Poincare Group are associated with "elementary particles" which have well defined mass and spin.

The group **SU(3)** has 8 generators and 2 Casimirs. In one key representation the generators are represented by 3x3 matrices which act on 3-vectors. Instead of having up and down states as with spin-1/2 noted above, these vectors have up, down and strange (sideways) states called u,d.s which are associated with quarks. The full symmetry group for the Standard Model of elementary particles is SU(3) x SU(2) x U(1) in which SU(3) plays its part.

Special Groups have Traceless Generators

It is desirable that "rotation matrices" have unit determinant because such matrices then do not change the "length" of a vector on which they act. When the representation matrices are restricted to have unit determinant, they are called "special" and the group name is prefixed by the letter S, as in SO(3) for the rotation group. For the 3x3 representation of the rotation group we know that the matrices are real orthogonal which means $RR^T = 1$ which in turn means $[\det(R)]^2 = 1$ and in SO(3) we select only those R with det(R) = +1. A "rotation" which just negates z (reflection) still satisfies $RR^T = 1$ but has det(R) = -1.

An elegant theorem concerning exponentiated square matrices is this (proved below):

$$\det(e^A) = e^{tr(A)} \tag{G.5.7}$$

where det is the determinant and $tr(A) \equiv \Sigma_i A_{ii}$ is the "trace" or "spur" of the matrix A -- the sum of the diagonal elements. If we want the matrix $\exp(-i\,\theta\,\hat{\mathbf{n}} \bullet \mathbf{J})$ to have unit determinant so it is "Special", the exponent must be traceless, and in this case that means that the generators J_i must all be traceless. One can see from examples (G.1.2) and (G.5.3) and (G.5.6) that this is indeed the case.

Offhand, the matrix identity $\det(e^{\mathbf{A}}) = e^{\mathbf{tr(A)}}$ seems very unlikely and almost too simple. For that reason, we include here a straightforward proof which we feel is "one for the Book".

Proof of (G.5.7) There are no implied sums in this proof!

Write $A = \Sigma_{ij} A_{ij}\, s^{(ij)}$ where $[s^{(ij)}]_{ab} \equiv \delta_{ia}\delta_{jb}$. To verify,

$$A_{ab} = \Sigma_{ij} A_{ij}\, [s^{(ij)}]_{ab} = \Sigma_{ij} A_{ij} \delta_{ia}\delta_{jb} = A_{ab} \quad .$$

The matrix $s^{(ij)}$ is all zeros except for a single 1 located in row a and column b. Using $e^{x+y\cdots} = e^x e^y \ldots$ we can write

$$e^{\mathbf{A}} = \exp(\Sigma_{ij} A_{ij}\, s^{(ij)}) = \Pi_{i,j} \exp(A_{ij}\, s^{(ij)}) \quad .$$

Then using det(XY..) = det(X)det(Y).... ,

$$\det(e^{\mathbf{A}}) = \det \{\Pi_{i,j} \exp(A_{ij}\, s^{(ij)}) \} = \Pi_{i,j} \det [\exp(A_{ij}\, s^{(ij)})] \quad . \qquad \text{(G.5.8)}$$

[Case i ≠ j :] We first note that $[s^{(ij)}]^2 = 0$:

$$[s^{(ij)}]^2_{ac} = \Sigma_b [s^{(ij)}]_{ab}\, [s^{(ij)}]_{bc} = \Sigma_b \delta_{ia}\delta_{jb}\, \delta_{ib}\delta_{jc} = \delta_{ia}\delta_{ji}\delta_{jc} = 0 \text{ since } i \neq j \ .$$

Then $[s^{(ij)}]^n = 0$ for $n \geq 2$. In this case we have

$$\exp(A_{ij}\, s^{(ij)}) = \Sigma_{n=0}^{\infty} (A_{ij})^n\, [\, s^{(ij)}]^n/n! = 1 + A_{ij}\, [\, s^{(ij)}]$$

and so

$$\det [\exp(A_{ij}\, s^{(ij)})] = \det [\, 1 + A_{ij}\, s^{(ij)}] = 1 \ .$$

This is so because the matrix in question has all 1's on the diagonal and one non-vanishing off-diagonal element at (i,j). In fact, *any* triangular matrix (one side all zeros) with 1's on the diagonal has det = 1.

[Case i=j :] In this case we have $[s^{(ii)}]^2 = [s^{(ii)}]$:

$$[s^{(ii)}]^2_{ac} = \Sigma_b [s^{(ii)}]_{ab}\, [s^{(ii)}]_{bc} = \Sigma_b \delta_{ia}\delta_{ib}\, \delta_{ib}\delta_{ic} = \delta_{ia}\delta_{ii}\delta_{ic} = \delta_{ia}\delta_{ic} = [s^{(ii)}]_{ac} \quad .$$

Note that $[s^{(ii)}]_{ab} = \delta_{ia}\delta_{ib} = 1$ only when a = b = i, so $s^{(ii)}$ is an all-zero matrix with a single 1 at location i on the diagonal.

Since $[s^{(ii)}]^2 = s^{(ii)}$ it follows that $[s^{(ii)}]^n = s^{(ii)}$ for $n \geq 1$. Then,

$$\exp(A_{ii}\, s^{(ii)}) = \Sigma_{n=0}^{\infty} (A_{ii})^n\, [\, s^{(ii)}]^n \,/\, n! = 1 + s^{(ii)}\ [\Sigma_{n=1}^{\infty} (A_{ii})^n \,/\, n!]$$

$= 1 + s^{(ii)} [-1 + \Sigma_{n=0}^{\infty} (A_{ii})^n/n!] = 1 + s^{(ii)} [-1 + \exp(A_{ii})]$.

This last item is the unit matrix with the i^{th} diagonal 1 replaced by $\exp(A_{ii})$. Therefore,

$\det [\exp(A_{ii}\, s^{(ii)})] = \det \{ 1 + s^{(ii)} [-1 + \exp(A_{ii})] \} = \exp(A_{ii})$.

Now go back to (G.5.8),

$\det(e^{A}) = \Pi_{i,j} \det [\exp(A_{ij}\, s^{(ij)})]$

$= (\Pi_{i \neq j} \det [\exp(A_{ij}\, s^{(ij)})]) * (\Pi_{i=j} \det [\exp(A_{ij}\, s^{(ij)})])$

$= (1*1*1*1.....$ $....*1*1) * (\exp(A_{11})\exp(A_{22}) .....$)

$= 1 * \exp(A_{11} + A_{22} + ...) = \exp (tr(A))$

$= e^{tr(A)}$. QED