Rigid Bodies

October 22, 2012

A rigid body is defined as one in which the distance between any two points in the body remains constant. If we think of the body as made up of individual particles, the distances between any two of these particles is fixed. We may equally well consider a continuum approximation. In either case the complete orientation of the body requires six parameters. Beginning from an arbitrary origin, we may locate any one point, P_1, in the body by three coordinates. Picking any other point, P_2, the distance d_{12} between them is fixed, so P_2 lies on a sphere of radius d_{12} centered on P_1, and we may specify the position of P_2 on this sphere by giving two angles. Finally, with the positions of P_1 and P_2 fixed, any third point P_3 lies in a plane with P_1 and P_2. Since this plane contains the line connecting P_1 and P_2, the only freedom in specifying P_3 is a single angle, specifying the orientation of this plane. We therefore require $3+2+1$ parameters to fully specify the position and orientation of a rigid body.

Fix an orthonormal coordinate system, with basis $\mathbf{i},\mathbf{j},\mathbf{k}$, fixed in space, and a second set, $\mathbf{i}',\mathbf{j}',\mathbf{k}'$, fixed in the rigid body. Think of the origin of the second set as P_1 so that the origins of these two systems separated by the position vector $\mathbf{R}_1$ of the point P_1. Now let P_2 define the $\mathbf{i}'$ the only remaining difference between the basis vectors will be determined by the three angular variables required to specify P_2, P_3.

1 Change of basis

We now seek the relationship between two orthonormal bases with a common origin.

The first key fact is that the transformation is *linear*, and this is immdediate by the definition of a vector basis. Given a set of basis vectors $(\mathbf{i},\mathbf{j},\mathbf{k})$, every vector can be expanded as a linear combination

$$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$$

Since any other basis is comprised of vectors, the vectors of the new basis $(\mathbf{i}',\mathbf{j}',\mathbf{k}')$ may be expanded in the old:

$$\begin{aligned}
\mathbf{i}' &= a_{11}\mathbf{i} + a_{12}\mathbf{j} + a_{13}\mathbf{k} \\
\mathbf{j}' &= a_{21}\mathbf{i} + a_{22}\mathbf{j} + a_{23}\mathbf{k} \\
\mathbf{k}' &= a_{31}\mathbf{i} + a_{32}\mathbf{j} + a_{33}\mathbf{k}
\end{aligned}$$

and conversely,

$$\begin{aligned}
\mathbf{i} &= b_{11}\mathbf{i}' + b_{12}\mathbf{j}' + b_{13}\mathbf{k}' \\
\mathbf{j} &= b_{21}\mathbf{i}' + b_{22}\mathbf{j}' + b_{23}\mathbf{k}' \\
\mathbf{k} &= b_{31}\mathbf{i}' + b_{32}\mathbf{j}' + b_{33}\mathbf{k}'
\end{aligned}$$

We consider two types of transformation: passive and active. A passive transformation is one in which any given vector remains fixed while we transform the basis. An active transformation is one in which we leave the basis fixed, but transform all vectors. Thus, for a passive transformation and an arbitrary vector $\mathbf{v}$, we expand in each basis:

$$\begin{aligned}
\mathbf{v} &= v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k} \\
&= v_1'\mathbf{i}' + v_2'\mathbf{j}' + v_3'\mathbf{k}'
\end{aligned}$$

where the basis vectors are related as above. For an active transformation, we consider two vectors

$$\begin{aligned} \mathbf{v} &= v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k} \\ \mathbf{v}' &= v_1'\mathbf{i} + v_2'\mathbf{j} + v_3'\mathbf{k} \end{aligned}$$

where the components are related by

$$\begin{aligned} \mathbf{v} &= v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k} \\ \mathbf{v}' &= v_1\mathbf{i}' + v_2\mathbf{j}' + v_3\mathbf{k}' \end{aligned}$$

1.1 Einstein summation convention

All of this is much easier in index notation. Let the three basis vectors be denoted by $\hat{\mathbf{e}}_i$, $i = 1, 2, 3$, so that

$$\begin{aligned} \hat{\mathbf{e}}_1 &= \hat{\mathbf{i}} \\ \hat{\mathbf{e}}_2 &= \hat{\mathbf{j}} \\ \hat{\mathbf{e}}_3 &= \hat{\mathbf{k}} \end{aligned}$$

and similarly for $\hat{\mathbf{e}}_i^{'}$. Then the basis transformations above may be written as

$$\begin{aligned} \hat{\mathbf{e}}_i^{'} &= \sum_{j=1}^{3} a_{ij}\hat{\mathbf{e}}_j \\ \hat{\mathbf{e}}_i &= \sum_{j=1}^{3} b_{ij}\hat{\mathbf{e}}_j^{'} \end{aligned}$$

and the vector expansions as

$$\mathbf{v} = \sum_{j=1}^{3} v_j\hat{\mathbf{e}}_j$$

It is easy to see that this will lead us to write $\sum_{j=1}^{3}$ millions of times. The Einstein convention avoids this by noting that when there is a sum there is also a repeated index – j, in the cases above. Also, we almost never repeat an index that we do not sum, so we may drop the summation sign. Thus,

$$\begin{aligned} \sum_{j=1}^{3} a_{ij}\hat{\mathbf{e}}_j &\Longrightarrow a_{ij}\hat{\mathbf{e}}_j \\ \sum_{j=1}^{3} b_{ij}\hat{\mathbf{e}}_j^{'} &\Longrightarrow b_{ij}\hat{\mathbf{e}}_j^{'} \\ \sum_{j=1}^{3} v_j\hat{\mathbf{e}}_j &\Longrightarrow v_j\hat{\mathbf{e}}_j \end{aligned}$$

The repeated index is called a dummy index, and it does not matter what letter we choose for it,

$$v_j\hat{\mathbf{e}}_j = v_k\hat{\mathbf{e}}_k$$

as long as we do not use an index that we have used elsewhere in the same expression. Thus, in the basis change examples above, we cannot use i as the dummy index because it is used to distinguish three independent equations:

$$\begin{aligned} \hat{\mathbf{e}}_1^{'} &= a_{1j}\hat{\mathbf{e}}_j \\ \hat{\mathbf{e}}_2^{'} &= a_{2j}\hat{\mathbf{e}}_j \\ \hat{\mathbf{e}}_3^{'} &= a_{3j}\hat{\mathbf{e}}_j \end{aligned}$$

Such an index is called a free index. Free indices must match in every term of an expression.

Since the basis is orthonormal, we know that the dot product is given by

$$\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij}$$

where δ_{ij} is the Kronecker delta, equal to 1 if $i = j$ and to 0 if $i \neq j$. Notice that the expression above represents nine separate equations. If we repeat the index, we have a single equation

$$\begin{aligned} \hat{\mathbf{e}}_k \cdot \hat{\mathbf{e}}_k &= \delta_{kk} \\ &= 3 \end{aligned}$$

Be sure you understand why the result is 3.

We can find the relationship between the matrices a_{ij} and b_{ij}, since, substituting one basis change into the other,

$$\begin{aligned} \hat{\mathbf{e}}_i' &= a_{ij}\hat{\mathbf{e}}_j \\ &= a_{ij}\left(b_{jk}\hat{\mathbf{e}}_k'\right) \\ &= a_{ij}b_{jk}\hat{\mathbf{e}}_k' \end{aligned}$$

Taking the dot product with $\hat{\mathbf{e}}_m'$ (notice that we cannot use i, j or k), we have

$$\begin{aligned} \hat{\mathbf{e}}_i' &= a_{ij}b_{jk}\hat{\mathbf{e}}_k' \\ \hat{\mathbf{e}}_m' \cdot \hat{\mathbf{e}}_i' &= \hat{\mathbf{e}}_m' \cdot \left(a_{ij}b_{jk}\hat{\mathbf{e}}_k'\right) \\ \delta_{mi} &= a_{ij}b_{jk}\hat{\mathbf{e}}_m' \cdot \hat{\mathbf{e}}_k' \\ &= a_{ij}b_{jk}\delta_{mk} \\ &= a_{ij}b_{jm} \end{aligned}$$

and since $\delta_{mi} = \delta_{im}$ is the identity matrix, 1, this shows that

$$AB = 1$$

so that the matrix B with components b_{ij} is inverse to A,

$$B = A^{-1}$$

1.2 Passive transformation

Consider a passive transformation from $\hat{\mathbf{e}}_j$ to $\hat{\mathbf{e}}_i'$. Substituting for the relationship between the basis vectors, we have

$$\begin{aligned} v_i'\hat{\mathbf{e}}_i' &= v_i\hat{\mathbf{e}}_i \\ &= v_i\left(b_{ij}\hat{\mathbf{e}}_j'\right) \\ &= \left(v_ib_{ij}\right)\hat{\mathbf{e}}_j' \end{aligned}$$

Continuing with the passive transformation, we take the dot product of both sides of our equation with each of the three basis vectors, $\hat{\mathbf{e}}_k'$:

$$\begin{aligned} v_i'\hat{\mathbf{e}}_i' \cdot \hat{\mathbf{e}}_k' &= \left(v_ib_{ij}\right)\hat{\mathbf{e}}_j' \cdot \hat{\mathbf{e}}_k' \\ v_i'\delta_{ik} &= v_ib_{ij}\delta_{jk} \\ v_k' &= v_ib_{ik} \end{aligned}$$

1.3 Active transformation

Now consider an active rotation of a vector $\mathbf{v}$ to a new vector $\mathbf{v}'$. To see clearly what is happening, first suppose we have two bases, $\hat{\mathbf{e}}_i$, and $\hat{\mathbf{e}}_i^{'}$ which differ only by a rotation around the z axix through an angle θ,

$$\begin{aligned} \mathbf{i}' &= \mathbf{i}\cos\theta + \mathbf{j}\sin\theta \\ \mathbf{j}' &= -\mathbf{i}\sin\theta + \mathbf{j}\cos\theta \\ \mathbf{k}' &= \mathbf{k} \end{aligned}$$

so that

$$a_{ij} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Now, suppose we want $\mathbf{v}$ lies in the x-direction and we want $\mathbf{v}'$ to be rotated by an angle θ. Then we have

$$\begin{aligned} \mathbf{v} &= v\mathbf{i} \\ \mathbf{v}' &= v\mathbf{i}' \\ &= v\left(\mathbf{i}\cos\theta + \mathbf{j}\sin\theta\right) \\ &= \left(\mathbf{i}v\cos\theta + \mathbf{j}v\sin\theta\right) \end{aligned}$$

so the components of $\mathbf{v}'$ in the unprimed basis are

$$v_i^{'} = (v\cos\theta, v\sin\theta, 0)$$

or, in terms of a_{ij},

$$v_i^{'} = v_j a_{ji}$$

This is the general relationship between $\mathbf{v}$ and $\mathbf{v}'$ since in general, if

$$\hat{\mathbf{e}}_i^{'} = a_{ij}\hat{\mathbf{e}}_j$$

and we reqire

$$\begin{aligned} \mathbf{v} &= v_i\hat{\mathbf{e}}_i \\ \mathbf{v}' &= v_i\hat{\mathbf{e}}_i^{'} \\ &= v_i'\hat{\mathbf{e}}_i \end{aligned}$$

it follows that

$$\begin{aligned} v_i^{'}\hat{\mathbf{e}}_i &= v_i\hat{\mathbf{e}}_i^{'} \\ v_i^{'}\hat{\mathbf{e}}_i &= v_i a_{ij}\hat{\mathbf{e}}_j \\ v_k^{'} &= v_i a_{ik} \end{aligned}$$

Notice that the transformation law is exactly inverse to the transformation of the basis.

1.4 Transpose

Let A have components a_{ij}. Then the transpose of A, called A^t, has components a_{ji}:

$$\begin{aligned} [A]_{ij} &= a_{ij} \\ \left[A^t\right]_{ij} &= a_{ji} \end{aligned}$$

For our active transformation, define the transformation O by

$$O = A^t$$

where A is our active transformation matrix,

$$v_k^{'} = v_i a_{ik}$$

Then we have

$$\begin{aligned} v_k^{'} &= v_i [O]_{ki} \\ &= [O]_{ki} v_i \end{aligned}$$

and this is the usual form of transformation, with the vector on the right. We may also write this as

$$\mathbf{v}' = O\mathbf{v}$$

2 Orthogonal transformations

2.1 Defining property

The squared length of a vector is given by taking the dot product of a vector with itself,

$$v^2 = \mathbf{v} \cdot \mathbf{v}$$

An *orthogonal transformation* is a linear transformation of a vector space that preserves lengths of vectors. This defining property may therefore be written as a linear transformation,

$$\mathbf{v}' = O\mathbf{v}$$

such that

$$\mathbf{v}' \cdot \mathbf{v}' = \mathbf{v} \cdot \mathbf{v}$$

Write this definition in terms of components using index notation. Setting the components of the transformation to $[O]_{ij} = O_{ij}$

$$v_i^{'} = O_{ij} v_j$$

we have

$$\begin{aligned} \mathbf{v}' \cdot \mathbf{v}' &= \mathbf{v} \cdot \mathbf{v} \\ v_i^{'} v_i^{'} &= v_i v_i \\ (O_{ij} v_j)(O_{ik} v_k) &= v_i v_i \end{aligned}$$

Notice that we change the name of the dummy indices so that we never have more than two indices repeated. Each term in the expression $(O_{ij} v_j)(O_{ik} v_k)$, for each value of i, j, k is just a real number, so we may rearrange the terms in any order we like. We may also write the dot product as a double sum, $v_i v_i = \delta_{jk} v_j v_k$. Then

$$\begin{aligned} v_i v_i &= O_{ij} v_j O_{ik} v_k \\ \delta_{jk} v_j v_k &= O_{ij} O_{ik} v_j v_k \end{aligned}$$

or

$$0 = (O_{ij} O_{ik} - \delta_{jk}) v_j v_k$$

To go further we need to derive a useful general result. Suppose $A_{ij} = -A_{ji}$ are the components of any antisymmetric matrix, and $S_{ij} = S_{ji}$ are the components of an arbitrary symmetric matrix. Consider the trace of the product,

$$tr(AS)$$

Since the product is given by $A_{ij}S_{jk}$, its trace is found by setting $k = i$ and summing,

$$A_{ij}S_{ji}$$

We can evaluate this, just from the symmetries,

$$\begin{aligned} A_{ij}S_{ji} &= A_{ij}S_{ij} \\ &= -A_{ji}S_{ij} \end{aligned}$$

Since dummy indices can be named arbitrarily, we may write $A_{ji}S_{ij} = A_{ij}S_{ji}$, giving

$$\begin{aligned} A_{ij}S_{ji} &= -A_{ji}S_{ji} \\ 2A_{ij}S_{ji} &= 0 \\ A_{ij}S_{ji} &= 0 \end{aligned}$$

Therefore, the full contraction of a symmetric object and and antisymmetric always vanishes.

Returning to our defining property of orthogonal transformations,

$$0 = (O_{ij}O_{ik} - \delta_{jk})\, v_j v_k$$

and recognizing that the matrix

$$M_{ij} = v_i v_j$$

is symmetric but otherwise arbitrary, we see that we can make no claim about the antisymmetric part of $(O_{ij}O_{ik} - \delta_{jk})$, since such a contraction vanishes identically in any case. What we can conclude is that the contraction of $(O_{ij}O_{ik} - \delta_{jk})$ with the arbitrary symmetric matrix $v_j v_k$ requires the vanishing of the symmetric part,

$$(O_{ij}O_{ik} - \delta_{jk}) + (O_{ik}O_{ij} - \delta_{kj}) = 0$$

This is enough, because

$$\begin{aligned} O_{ij}O_{ik} &= O_{ik}O_{ij} \\ \delta_{jk} &= \delta_{kj} \end{aligned}$$

and we have

$$\begin{aligned} O_{ij}O_{ik} - \delta_{jk} &= 0 \\ O_{ij}O_{ik} &= \delta_{jk} \end{aligned}$$

and since the components of O^t are just $[O^t]_{ij} = O_{ji}$,

$$\begin{aligned} O^t_{ji}O_{ik} &= \delta_{jk} \\ O^t O &= 1 \end{aligned}$$

and we see that the defining property of an orthogonal transformation is that the transpose equal the inverse:

$$O^t = O^{-1}$$

2.2 Infinitesimal generators

There are two useful very ways to write the components of a 3-dimensional vector space. The first is as the usual real triples,

$$\mathbf{v} = (v_1, v_2, v_3)$$

while the second is as linear combinations of traceless, Hermitian 2×2 matrices,

$$\begin{aligned} \mathbf{v} &= v_1\sigma_1 + v_2\sigma_2 + v_3\sigma_3 \\ &= v_i\sigma_i \end{aligned}$$

where σ_i are the three Pauli matrices,

$$\begin{aligned} \sigma_1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \sigma_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ \sigma_2 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

Prove, as an exercise, that these are the only traceless, Hermitian 2×2 matrics. This means that there is a one to one correspondence between 2-dimensional traceless Hermitian matrics and 3-dimensional real vectors,

$$\begin{aligned} \mathbf{v} &= v_i\sigma_i \\ &= \begin{pmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 \end{pmatrix} \end{aligned}$$

Notice that linear combinations of these matrices remain traceless and Hermitian. The squared length of the vector $\mathbf{v}$ may be written as the negative of the determinant,

$$\begin{aligned} v^2 &= -\det \mathbf{v} \\ &= -\det \begin{pmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 \end{pmatrix} \\ &= -\left[-v_3^2 - (v_1 + iv_2)(v_1 - iv_2)\right] \\ &= \left(v_1^2 + v_2^2 + v_3^2\right) \end{aligned}$$

These facts give us two ways to describe orthogonal transformations. For the usual real representation of vectors, we have already seen that a real, linear transformation satisfying $O^t = O^{-1}$ is an orthogonal transformation. For the matrix representation, we require a similarity transformation,

$$\mathbf{v}' = U\mathbf{v}U^{-1}$$

which preserves three properties:

1. Vanishing trace,
$$tr(\mathbf{v}') = tr\left(U\mathbf{v}U^{-1}\right) = 0$$

2. Hermitian,
$$\begin{aligned} \mathbf{v}'^\dagger &= \left(U\mathbf{v}U^{-1}\right)^\dagger \\ &= U^{-1\dagger}\mathbf{v}^\dagger U^\dagger \\ &= U^{-1\dagger}\mathbf{v}U^\dagger \end{aligned}$$

3. Determinant,
$$\det \mathbf{v}' = \det \mathbf{v}$$

For both of these representations, the conditions on the components, while algebraic, are complicated. Fortunately, we can linearize the conditions and still recover the full form of the transformations. We consider both representations.

2.3 SO(3): The Special (i.e., det O = 1) Orthogonal group in 3 dimensions

For the real, 3-dimensional representation of the rotations, we require

$$O^t = O^{-1}$$

Notice that the identity satisfies this condition, so we may consider linear transformations near the identity which also satisfy the condition. Let

$$O = 1 + \varepsilon$$

where $[\varepsilon]_{ij} = \varepsilon_{ij}$ are all small, $|\varepsilon_{ij}| \ll 1$ for all i, j. Keeping only terms to first order in ε_{ij}, we have:

$$\begin{aligned} O^t &= 1 + \varepsilon^t \\ O^{-1} &= 1 - \varepsilon \end{aligned}$$

where we see that we have O^{-1} right by computing

$$\begin{aligned} OO^{-1} &= (1+\varepsilon)(1-\varepsilon) \\ &= 1 - \varepsilon^2 \\ &\approx 1 \end{aligned}$$

correct to first order in ε. Now we impose our condition,

$$\begin{aligned} O^t &= O^{-1} \\ 1 + \varepsilon^t &= 1 - \varepsilon \\ \varepsilon^t &= -\varepsilon \end{aligned}$$

so that the matrix ε must be antisymmetric.

Next, we write the most general antisymmetric 3×3 matrix as a linear combination of a convenient basis,

$$\begin{aligned} \varepsilon &= w_i J_i \\ &= w_1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} + w_2 \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + w_3 \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & w_2 & -w_3 \\ -w_2 & 0 & w_1 \\ w_3 & -w_1 & 0 \end{pmatrix} \end{aligned}$$

Notice that the components of the three matrices J_i are neatly summarized by

$$[J_i]_{jk} = \varepsilon_{ijk}$$

where ε_{ijk} is the totally antisymmetric Levi-Civita tensor. The matrices J_i are called the *generators* of the transformations.

Knowing the generators is enough to recover an arbitrary rotation. Starting with

$$O = 1 + \varepsilon$$

we may apply O repeatedly, taking the limit

$$\begin{aligned} O(\theta) &= \lim_{n \longrightarrow \infty} O^n \\ &= \lim_{n \longrightarrow \infty} (1+\varepsilon)^n \\ &= \lim_{n \longrightarrow \infty} (1 + w_i J_i)^n \end{aligned}$$

where the limit is taken in such a way that if w is the length of the infinitesmal vector w_i, so that $w_i = wn_i$, where n_i is a unit vector, then

$$\lim_{n \longrightarrow \infty} nw = \theta$$

where θ is finite. Using the binomial expansion, $(a+b)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} a^{n-k} b^k$ we have

$$\begin{aligned}
\lim_{n \longrightarrow \infty} O^n &= \lim_{n \longrightarrow \infty} (1 + w_i J_i)^n \\
&= \lim_{n \longrightarrow \infty} \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} (1)^{n-k} (w_i J_i)^k \\
&= \lim_{n \longrightarrow \infty} \sum_{k=0}^{n} \frac{n\,(n-1)\,(n-2)\cdots(n-k+1)}{k!} \frac{1}{n^k} \left((nw)\, n_i J_i\right)^k \\
&= \lim_{n \longrightarrow \infty} \sum_{k=0}^{n} \frac{\frac{n}{n}\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)\cdots\left(\frac{n-k+1}{n}\right)}{k!} (\theta n_i J_i)^k \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} (\theta n_i J_i)^k \\
&\equiv \exp(\theta n_i J_i)
\end{aligned}$$

We define the exponential of a matrix by the power series for the exponential, applied using powers of the matrix.

To find the form of a general rotation, we now need to find powers of $n_i J_i$. This turns out to be straightforward:

$$\begin{aligned}
[n_i J_i]_{jk} &= n_i \varepsilon_{ijk} \\
\left[(n_i J_i)^2\right]_{mn} &= (n_i \varepsilon_{imk})(n_j \varepsilon_{jkn}) \\
&= n_i n_j \varepsilon_{imk} \varepsilon_{jkn} \\
&= -n_i n_j (\delta_{ij}\delta_{mn} - \delta_{in}\delta_{jm}) \\
&= -(n_i n_j \delta_{ij})\,\delta_{mn} + n_m n_n \\
&= n_m n_n - \delta_{mn} \\
\left[(n_i J_i)^3\right]_{mn} &= (n_m n_k - \delta_{mk})\, n_i \varepsilon_{ikn} \\
&= n_m n_k n_i \varepsilon_{ikn} - \delta_{mk} n_i \varepsilon_{ikn} \\
&= -n_i \varepsilon_{imn} \\
&= -[n_i J_i]_{mn}
\end{aligned}$$

The powers come back to $n_i J_i$ with only a sign change, so we can divide the series into even and odd p This is the matrix for a rotation through an angle θ around an axis in the direction of $\mathbf{n}$.

We can now compute the exponential explicitly:

$$\begin{aligned}
[O(\theta, \hat{\mathbf{n}})]_{mn} &= [\exp(\theta n_i J_i)]_{mn} \\
&= \left[\sum_{k=0}^{\infty} \frac{1}{k!} (\theta n_i J_i)^k\right]_{mn} \\
&= \left[\sum_{l=0}^{\infty} \frac{1}{(2l)!} (\theta n_i J_i)^{2l}\right]_{mn} + \left[\sum_{l=0}^{\infty} \frac{1}{(2l+1)!} (\theta n_i J_i)^{2l+1}\right]_{mn} \\
&= \left[1 + \sum_{l=1}^{\infty} \frac{1}{(2l)!} (\theta n_i J_i)^{2l}\right]_{mn} + \left[\sum_{l=0}^{\infty} \frac{1}{(2l+1)!} (\theta n_i J_i)^{2l+1}\right]_{mn}
\end{aligned}$$

$$
\begin{aligned}
&= \delta_{mn} + \sum_{l=1}^{\infty} \frac{(-1)^l}{(2l)!} \theta^{2l} (\delta_{mn} - n_m n_n) + \sum_{l=0}^{\infty} \frac{(-1)^l}{(2l+1)!} \theta^{2l+1} [n_i J_i]_{mn} \\
&= \delta_{mn} + (\cos\theta - 1)(\delta_{mn} - n_m n_n) + \sin\theta [n_i J_i]_{mn}
\end{aligned}
$$

where we get $(\cos\theta - 1)$ because the $l = 0$ term is missing from the sum.

To see what this means, let O act on an arbitrary vector $\mathbf{v}$, and write the result in normal vector notation,

$$
\begin{aligned}
[O(\theta, \hat{\mathbf{n}})]_{mn} v_n &= (\delta_{mn} + (\cos\theta - 1)(\delta_{mn} - n_m n_n) + \sin\theta [n_i J_i]_{mn}) v_n \\
&= \delta_{mn} v_n + (\cos\theta - 1)(\delta_{mn} v_n - n_m n_n v_n) + \sin\theta [n_i J_i]_{mn} v_n \\
&= v_m + (\cos\theta - 1)(v_m - n_m (n_n v_n)) + \sin\theta \varepsilon_{imn} n_i v_n
\end{aligned}
$$

Now define the components of $\mathbf{v}$ parallel and perpendicular to the unit vector $\mathbf{n}$:

$$
\begin{aligned}
\mathbf{v}_{\|} &= (\mathbf{v} \cdot \mathbf{n}) \mathbf{n} \\
\mathbf{v}_{\perp} &= \mathbf{v} - (\mathbf{v} \cdot \mathbf{n}) \mathbf{n}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
O(\theta, \hat{\mathbf{n}}) \mathbf{v} &= \mathbf{v} - (\mathbf{v} - (\mathbf{v} \cdot \mathbf{n}) \mathbf{n}) + (\mathbf{v} - (\mathbf{v} \cdot \mathbf{n}) \mathbf{n}) \cos\theta - \sin\theta (\mathbf{n} \times \mathbf{v}) \\
&= \mathbf{v}_{\|} + \mathbf{v}_{\perp} \cos\theta - \sin\theta (\mathbf{n} \times \mathbf{v})
\end{aligned}
$$

This expresses the rotated vector in terms of three mutually perpendicular vectors, $\mathbf{v}_{\|}, \mathbf{v}_{\perp}, (\mathbf{n} \times \mathbf{v})$. The direction $\mathbf{n}$ is the axis of the rotation. The part of $\mathbf{v}$ parallel to $\mathbf{n}$ is therefore unchanged. The rotation takes place in the plane perpendicular to $\mathbf{n}$, and this plane is spanned by $\mathbf{v}_{\perp}, (\mathbf{n} \times \mathbf{v})$. The rotation in this plane takes $\mathbf{v}_{\perp}$ into the linear combination $\mathbf{v}_{\perp} \cos\theta - (\mathbf{n} \times \mathbf{v}) \sin\theta$, which is exactly what we expect for a rotation of $\mathbf{v}_{\perp}$ through an angle θ. The rotation $O(\theta, \hat{\mathbf{n}})$ is therefore a rotation by θ around the axis $\hat{\mathbf{n}}$.

2.4 SU(2): The Special (det U = 1) Unitary ($U^\dagger = U^{-1}$) group in 2 dimensions

It turns out that all orthogonal groups ($SO(n)$, rotations in n real dimensions) may be written as special cases of rotations in a related complex space. For $SO(3)$, it turns out that rotations in a complex, 2-dimensional space work. To see why this is, we show that we can write a real, 3-dimensional vector as a complex hermitian matrix. We establish this by first studying the Lorentz group, then finding the rotations as a subgroup.

2.4.1 Lorentz transformations

First, notice that matrices form a vector space. We can add linear combinations of them to form new matrices, and the same is true of hermitian matrices. Any real linear combination of hermitian matrices is also hermitian, since for any real a, b and hermitian A, B we have

$$
\begin{aligned}
C &= aA + bB \\
C^\dagger &= (aA + bB)^\dagger \\
&= (aA)^\dagger + (bB)^\dagger \\
&= aA^\dagger + bB^\dagger \\
&= aA + bB \\
&= C
\end{aligned}
$$

Next, we notice that the space of 2-dim hermitian matrices is 4-dimensional. Let A be hermitian. Then

$$
\begin{aligned}
A &= \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \\
&= A^\dagger \\
&= \begin{pmatrix} \bar{\alpha} & \bar{\gamma} \\ \bar{\beta} & \bar{\delta} \end{pmatrix}
\end{aligned}
$$

so that $\alpha = \bar{\alpha}$, $\delta = \bar{\delta}$ and $\beta = \bar{\gamma}$. We therefore may write

$$\begin{aligned} A &= \begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix} \\ &= t\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + x\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + z\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= t\mathbf{1} + \mathbf{x}\cdot\boldsymbol{\sigma} \end{aligned}$$

where we choose the identity and the Pauli matrices as a basis for the 4-dim space.

Now consider the determinant of A,

$$\begin{aligned} \det A &= \det\begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix} \\ &= (t+z)(t-z) - (x+iy)(x-iy) \\ &= t^2 - z^2 - x^2 - y^2 \end{aligned}$$

This is the proper length of a 4-vector in spacetime, which means that any transformation which preserves the hermiticity and determinant of A is a Lorentz transformation.

It is now easy to write the Lorentz transformations. Since matrices transform by similarity transformation, we consider any transformation of the form

$$A' = LAL^\dagger$$

This is hermitian whenever A is and preserves the determinant provided

$$\begin{aligned} 1 &= \det A' \\ &= \det\left(LAL^\dagger\right) \\ &= \det L \det A \det L^\dagger \\ &= \det L \det L^\dagger \\ &= |\det L|^2 \end{aligned}$$

so that $\det L = \pm 1$. The positive determinant transformations preserve the direction of time and are called orthochronos, forming the group $SL(2, C)$, i.e. unit determinant (special), linear transformations in 2 complex dimensions.

2.4.2 Rotations as SU(2)

The rotation group is the subset of the Lorentz transformations which do not involve the time, t. We therefore may look for those Lorentz transformations with $t' = t$. Since we may write our 4-vector, before and after, as

$$\begin{aligned} A &= t\mathbf{1} + \mathbf{x}\cdot\boldsymbol{\sigma} \\ A' &= t'\mathbf{1} + \mathbf{x}'\cdot\boldsymbol{\sigma} \end{aligned}$$

we need the Lorentz transformations which leave the identity unchanged. Now consider an infinitesimal Lorentz transformation, $L = 1 + \varepsilon$. It will be a rotation provided

$$\begin{aligned} (1+\varepsilon)\,1\,(1+\varepsilon)^\dagger &= 1 \\ 1 + \varepsilon + \varepsilon^\dagger + O\left(\varepsilon^2\right) &= 1 \\ \varepsilon + \varepsilon^\dagger &= 0 \\ \varepsilon^\dagger &= -\varepsilon \end{aligned}$$

which means the generator must be anti-hermitian. A general anti-hermitian matrix may be written as

$$\begin{aligned} \varepsilon &= i\left(a_0 1 + \mathbf{a}\cdot\boldsymbol{\sigma}\right) \\ \varepsilon^\dagger &= -i\left(a_0 1 + \mathbf{a}\cdot\boldsymbol{\sigma}\right) \end{aligned}$$

We must still require unit determinant. To do this, recall the methods of the previous sub-Section, in which we showed that a general rotation is related to an infinitesimal rotation by

$$O = \exp\left(w_i J_i\right)$$

where J_i are the infinitesimal generators. The same argument holds here, so for the Lorentz transformations. Since the generators are given by the property

$$\varepsilon^\dagger = -\varepsilon$$

the infintesimal matrix ε is anti-hermitian, and we may write it as i times a general hermitian matrix,

$$\begin{aligned} \varepsilon &= i\left(a_0 1 + \mathbf{a}\cdot\boldsymbol{\sigma}\right) \\ \varepsilon^\dagger &= -i\left(a_0 1 + \mathbf{a}\cdot\boldsymbol{\sigma}\right) \end{aligned}$$

A general rotation must therefore be of the form

$$U = \exp\left(i\left(a_0 1 + \mathbf{a}\cdot\boldsymbol{\sigma}\right)\right)$$

In this form we may check the determinant using the fact that when $A = e^B$,

$$\det A = e^{trB}$$

To have unit determinant we therefore demand

$$\begin{aligned} 1 &= \det U \\ &= \exp\left(itr\left(a_0 1 + \mathbf{a}\cdot\boldsymbol{\sigma}\right)\right) \\ &= e^{2ia_0} \end{aligned}$$

and set $a_0 = 1$.

Finally, notice that

$$\varepsilon^\dagger = -\varepsilon$$

means that

$$\begin{aligned} A^{-1} &= 1 - \varepsilon \\ &= 1 + \varepsilon^\dagger \\ &= A^\dagger \end{aligned}$$

so that the rotation group is $SU(2)$, the group of unit determinant, unitary matrices in 2-dimensions.

2.4.3 The form of rotation matrices

Now compute the exponential for a rotation,

$$\begin{aligned} U &= \exp\left(i\mathbf{a}\cdot\boldsymbol{\sigma}\right) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!}\left(i\mathbf{a}\cdot\boldsymbol{\sigma}\right)^k \end{aligned}$$

We need to compute powers of the Pauli matrices. For this it is helpful to have the product

$$\sigma_i \sigma_j = \delta_{ij} \mathbf{1} + i\varepsilon_{ijk}\sigma_k$$

which you are invited to prove. Let $\mathbf{a} = a\mathbf{n}$, where $\mathbf{n}$ is a unit vector. Then

$$\begin{aligned}
(\mathbf{a} \cdot \boldsymbol{\sigma})^2 &= a^2 (\mathbf{n} \cdot \boldsymbol{\sigma})^2 \\
&= (n_i \sigma_i)(n_j \sigma_j) \\
&= n_i n_j \sigma_i \sigma_j \\
&= n_i n_j (\delta_{ij} \mathbf{1} + i\varepsilon_{ijk}\sigma_k) \\
&= n_i n_j \delta_{ij} \mathbf{1} + i\varepsilon_{ijk} n_i n_j \sigma_k \\
&= a^2 (\mathbf{n} \cdot \mathbf{n}) \mathbf{1} + i (\mathbf{n} \times \mathbf{n}) \cdot \boldsymbol{\sigma} \\
&= a^2 \mathbf{1}
\end{aligned}$$

Higher powers follow immediately,

$$\begin{aligned}
(\mathbf{a} \cdot \boldsymbol{\sigma})^{2k+1} &= a^{2k+1} \mathbf{n} \cdot \boldsymbol{\sigma} \\
(\mathbf{a} \cdot \boldsymbol{\sigma})^{2k} &= a^{2k} \mathbf{n} \cdot \boldsymbol{\sigma}
\end{aligned}$$

and the exponential becomes

$$\begin{aligned}
U &= \sum_{k=0}^{\infty} \frac{1}{k!} (i\mathbf{a} \cdot \boldsymbol{\sigma})^k \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} a^{2k} (\mathbf{n} \cdot \boldsymbol{\sigma})^{2k} + i \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} a^{2k+1} (\mathbf{n} \cdot \boldsymbol{\sigma})^{2k+1} \\
&= \mathbf{1} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} a^{2k} + i\mathbf{n} \cdot \boldsymbol{\sigma} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} a^{2k+1} \\
&= \mathbf{1} \cos a + i\mathbf{n} \cdot \boldsymbol{\sigma} \sin a
\end{aligned}$$

Now apply this to a 3-vector, written as

$$X = \mathbf{x} \cdot \boldsymbol{\sigma}$$

We have

$$\begin{aligned}
X' &= \mathbf{x}' \cdot \boldsymbol{\sigma} \\
&= U (\mathbf{x} \cdot \boldsymbol{\sigma}) U^\dagger \\
&= (\mathbf{1} \cos a + i\mathbf{n} \cdot \boldsymbol{\sigma} \sin a)(\mathbf{x} \cdot \boldsymbol{\sigma})(\mathbf{1} \cos a - i\mathbf{n} \cdot \boldsymbol{\sigma} \sin a) \\
&= (\mathbf{x} \cdot \boldsymbol{\sigma}) \cos a \cos a + i (\mathbf{n} \cdot \boldsymbol{\sigma})(\mathbf{x} \cdot \boldsymbol{\sigma}) \cos a \sin a - i (\mathbf{x} \cdot \boldsymbol{\sigma})(\mathbf{n} \cdot \boldsymbol{\sigma}) \cos a \sin a \\
&\quad + (\mathbf{n} \cdot \boldsymbol{\sigma})(\mathbf{x} \cdot \boldsymbol{\sigma})(\mathbf{n} \cdot \boldsymbol{\sigma}) \sin a \sin a \\
&= (\mathbf{x} \cdot \boldsymbol{\sigma}) \cos a \cos a + i n_i x_j [\sigma_i, \sigma_j] \cos a \sin a + (\mathbf{n} \cdot \boldsymbol{\sigma})(\mathbf{x} \cdot \boldsymbol{\sigma})(\mathbf{n} \cdot \boldsymbol{\sigma}) \sin a \sin a
\end{aligned}$$

Evaluating the products of Pauli matrices,

$$\begin{aligned}
i n_i x_j [\sigma_i, \sigma_j] &= i n_i x_j (2i\varepsilon_{ijk}\sigma_k) \\
&= -2 (\mathbf{n} \times \mathbf{x}) \cdot \boldsymbol{\sigma} \\
(\mathbf{n} \cdot \boldsymbol{\sigma})(\mathbf{x} \cdot \boldsymbol{\sigma})(\mathbf{n} \cdot \boldsymbol{\sigma}) &= (\mathbf{n} \cdot \boldsymbol{\sigma}) x_i n_j (\delta_{ij} 1 + i\varepsilon_{ijk}\sigma_k) \\
&= (\mathbf{n} \cdot \boldsymbol{\sigma}) ((\mathbf{x} \cdot \mathbf{n}) 1 + i (\mathbf{x} \times \mathbf{n}) \cdot \boldsymbol{\sigma}) \\
&= (\mathbf{x} \cdot \mathbf{n})(\mathbf{n} \cdot \boldsymbol{\sigma}) + i (\mathbf{n} \cdot \boldsymbol{\sigma})(\mathbf{x} \times \mathbf{n}) \cdot \boldsymbol{\sigma}
\end{aligned}$$

$$
\begin{aligned}
&= (\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) + in_i(\mathbf{x}\times\mathbf{n})_j\,\sigma_i\sigma_j \\
&= (\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) + in_i(\mathbf{x}\times\mathbf{n})_j\,(\delta_{ij}1 + i\varepsilon_{ijk}\sigma_k) \\
&= (\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) + i\mathbf{n}\cdot(\mathbf{x}\times\mathbf{n})\,1 - (\mathbf{n}\times(\mathbf{x}\times\mathbf{n}))\cdot\boldsymbol{\sigma} \\
&= (\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) - (\mathbf{x}(\mathbf{n}\cdot\mathbf{n}) - \mathbf{n}(\mathbf{x}\cdot\mathbf{n}))\cdot\boldsymbol{\sigma} \\
&= 2(\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) - \mathbf{x}\cdot\boldsymbol{\sigma}
\end{aligned}
$$

Substituting,

$$
\begin{aligned}
\mathbf{x}'\cdot\boldsymbol{\sigma} &= (\mathbf{x}\cdot\boldsymbol{\sigma})\cos a\cos a - 2(\mathbf{n}\times\mathbf{x})\cdot\boldsymbol{\sigma}\cos a\sin a + (2(\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma}) - \mathbf{x}\cdot\boldsymbol{\sigma})\sin a\sin a \\
&= (\mathbf{x}\cdot\boldsymbol{\sigma})(\cos a\cos a - \sin a\sin a) - (\mathbf{n}\times\mathbf{x})\cdot\boldsymbol{\sigma}2\cos a\sin a + 2(\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma})\sin a\sin a \\
&= (\mathbf{x}\cdot\boldsymbol{\sigma})\cos 2a - (\mathbf{n}\times\mathbf{x})\cdot\boldsymbol{\sigma}\sin 2a + (\mathbf{x}\cdot\mathbf{n})(\mathbf{n}\cdot\boldsymbol{\sigma})(1-\cos 2a) \\
&= [\mathbf{x}\cos 2a - \mathbf{n}\times\mathbf{x}\sin 2a + (\mathbf{x}\cdot\mathbf{n})\,\mathbf{n}\,(1-\cos 2a)]\cdot\boldsymbol{\sigma} \\
&= [(\mathbf{x} - (\mathbf{x}\cdot\mathbf{n})\,\mathbf{n})\cos 2a - \mathbf{n}\times\mathbf{x}\sin 2a + (\mathbf{x}\cdot\mathbf{n})\,\mathbf{n}]\cdot\boldsymbol{\sigma}
\end{aligned}
$$

and equating coefficients,

$$
\begin{aligned}
\mathbf{x}' &= (\mathbf{x} - (\mathbf{x}\cdot\mathbf{n})\,\mathbf{n})\cos 2a - \mathbf{n}\times\mathbf{x}\sin 2a + (\mathbf{x}\cdot\mathbf{n})\,\mathbf{n} \\
&= \mathbf{x}_{\parallel} + \mathbf{x}_{\perp}\cos 2a - (\mathbf{n}\times\mathbf{x}_{\perp})\sin 2a \\
&= \mathbf{x}_{\parallel} + \mathbf{x}_{\perp}\cos\theta - (\mathbf{n}\times\mathbf{x}_{\perp})\sin\theta
\end{aligned}
$$

which is the same transformation as we derived from $SO(3)$ once we identify $2a = \theta$. This means that as a runs from 0 to 2π, the 3-dim angle only runs from 0 to π, and a complete cycle requires a to climb to 4π.

There are important things to be gained from the $SU(2)$ representation of rotations. First, it is much easier to work with the Pauli matrices than it is with 3×3 matrices. Although the generators in the 2- and 3-dimensional cases are simple, the exponentials are not. The exponential of the J_i matrices is rather complicated, while the exponential of the Pauli matrices may again be expressed in terms of the Pauli matrices,

$$
\exp ia\mathbf{n}\cdot\boldsymbol{\sigma} = \mathbf{1}\cos a + i\mathbf{n}\cdot\boldsymbol{\sigma}\sin a
$$

and this is a substantial simplification of calculations.

More importantly, there is a crucial physical insight. The transformations U act on our hermitian matrices by a similarity transformation, but they also act on some 2-dimensional vector space. Denote a vector in this space as $\psi = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, so that the transformation of ψ is given by

$$
\psi' = U\psi
$$

This transformation preserves the hermitian norm of ψ, since

$$
\begin{aligned}
\psi'^{\dagger}\psi' &= \left(\psi^{\dagger}U^{\dagger}\right)(U\psi) \\
&= \psi^{\dagger}\left(U^{\dagger}U\right)\psi \\
&= \psi^{\dagger}\psi
\end{aligned}
$$

The complex vector ψ is called a spinor, with its first and second components being called "spin up" and "spin down". While spinors were not discovered physically until quantum mechanics, their existence is predictable classically from the properties of rotations.