Unitary rotations

October 28, 2014

1 The special unitary 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 unitary transformations 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 complex representation of the Lorentz group, then finding the rotations as a subgroup. We end up showing that rotations may be accomplished using special (i.e, $\det U = 1$) unitary $\left(U^\dagger = U^{-1}\right)$ transformations in 2-dimensions, $SU(2)$.

1.1 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

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

so that $\alpha = \bar{\alpha} = a$, $\delta = \bar{\delta} = b$ and $\beta = \bar{\gamma}$, where a, b are real and an overbar denotes complex conjugation. We therefore may write

$$A = \begin{pmatrix} a & \bar{\beta} \\ \beta & b \end{pmatrix}$$

Introducing four real numbers, x, y, z, t, and setting $a = t + z, b = t - z$ and $\beta = x + iy$,

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

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. The most general, linear transformation of a matrix is by similarity transformation, so we consider any transformation of the form

$$A' = LAL^\dagger$$

where we use $L^\dagger$ on the right so that the new matrix is hermitian whenever A is,

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

L also 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 special linear group in 2 complex dimensions, $SL(2, C)$. Since these transformations preserve $\tau^2 = t^2 - z^2 - x^2 - y^2$, they are Lorentz transformations.

We now restrict to spatial rotations.

1.2 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. Since the Pauli matrices are all traceless, $tr\boldsymbol{\sigma} = 0$, but the identity is not, the condition we need is $trA = 0$. Such matrices take the form

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

and satisfy $\det A = -x^2 - y^2 - z^2$. We need only vanishing trace to have A represent a 3-dimensional vector.

Now consider a transformation of A. It is a rotation provided

$$tr\left(UAU^{\dagger}\right) = 0$$

whenever $trA = 0$. But the trace is cyclic, $tr\,(ABC) = tr\,(BCA) = tr\,(CAB)$, so that

$$tr\left(UAU^{\dagger}\right) = tr\left(U^{\dagger}UA\right)$$

If $U^{\dagger}U = 1$, then we have $trA' = trA$ and U is a rotation. The complex, 2-dimensional matrices with $U^{\dagger}U = 1$ are *unitary*, and since we have already asked for unit determinant, the rotation group is $SU\,(2)$.

1.3 The form of rotation matrices

We now find the general form of an $SU\,(2)$ transformation. Starting from an infinitesimal rotation, $U = 1+\varepsilon$, we require

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

which means the generator must be anti-hermitian. Let $\varepsilon = ih$, where h is Hermitian, $h = h^{\dagger}$. We also need the determinant of U to be 1. To first order, with

$$U = 1 + \begin{pmatrix} a+b & c-id \\ c+id & a-b \end{pmatrix}$$

and a, b, c, d all small, this gives

$$\begin{aligned} 1 &= \det U \\ &= \det \begin{pmatrix} 1+a+b & c-id \\ c+id & 1+a-b \end{pmatrix} \\ &= (1+a+b)(1+a-b) - (c-id)(c+id) \\ &= 1+a-b+a+b+O\left(\varepsilon^{2}\right) \\ &= 1+2a+O\left(\varepsilon^{2}\right) \end{aligned}$$

and therefore $a = 0$. This means that h is traceless, and U may be written as

$$U = 1 + \varepsilon\mathbf{n}\cdot\boldsymbol{\sigma}$$

with $\varepsilon \ll 1$ and $\mathbf{n}$ a unit vector.

1.4 Finite transformations

Taking the limit of many infinitesimal transformations,

$$\begin{aligned} U(\varphi,\mathbf{n}) &= \lim_{k\to\infty}\left(1+\varepsilon\mathbf{n}\cdot\boldsymbol{\sigma}\right)^{k} \\ &= \exp\left(\frac{i\varphi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}\right) \end{aligned}$$

where

$$\lim_{k\to\infty} \varepsilon k = \frac{\varphi}{2}$$

Now compute the exponential for a rotation,

$$\begin{aligned} U &= \exp\left(\frac{i\varphi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}\right) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{i\varphi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}\right)^k \end{aligned}$$

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. Then

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

Higher powers follow immediately,

$$\begin{aligned} \left(\frac{i\varphi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}\right)^{2m+1} &= (-1)^k\, i\left(\frac{\varphi}{2}\right)^{2k+1} \mathbf{n}\cdot\boldsymbol{\sigma} \\ \left(\frac{i\varphi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}\right)^{2k} &= (-1)^k \left(\frac{\varphi}{2}\right)^{2k} \mathbf{1} \end{aligned}$$

for $k = 0, 1, 2, \ldots$ and the exponential becomes

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

Now apply this to a 3-vector, written as

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

We have

$$\begin{aligned} X' &= \mathbf{x}'\cdot\boldsymbol{\sigma} \\ &= U(\mathbf{x}\cdot\boldsymbol{\sigma})U^\dagger \\ &= \left(\mathbf{1}\cos\frac{\varphi}{2} + i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\frac{\varphi}{2}\right)(\mathbf{x}\cdot\boldsymbol{\sigma})\left(\mathbf{1}\cos\frac{\varphi}{2} - i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\frac{\varphi}{2}\right) \\ &= (\mathbf{x}\cdot\boldsymbol{\sigma})\cos\frac{\varphi}{2}\cos\frac{\varphi}{2} + i(\mathbf{n}\cdot\boldsymbol{\sigma})(\mathbf{x}\cdot\boldsymbol{\sigma})\cos\frac{\varphi}{2}\sin\frac{\varphi}{2} - i(\mathbf{x}\cdot\boldsymbol{\sigma})(\mathbf{n}\cdot\boldsymbol{\sigma})\cos\frac{\varphi}{2}\sin\frac{\varphi}{2} \\ &\quad + (\mathbf{n}\cdot\boldsymbol{\sigma})(\mathbf{x}\cdot\boldsymbol{\sigma})(\mathbf{n}\cdot\boldsymbol{\sigma})\sin\frac{\varphi}{2}\sin\frac{\varphi}{2} \\ &= (\mathbf{x}\cdot\boldsymbol{\sigma})\cos\frac{\varphi}{2}\cos\frac{\varphi}{2} + in_i x_j[\sigma_i,\sigma_j]\cos\frac{\varphi}{2}\sin\frac{\varphi}{2} + (\mathbf{n}\cdot\boldsymbol{\sigma})(\mathbf{x}\cdot\boldsymbol{\sigma})(\mathbf{n}\cdot\boldsymbol{\sigma})\sin\frac{\varphi}{2}\sin\frac{\varphi}{2} \end{aligned}$$

Evaluating the products of Pauli matrices,

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

Substituting,

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

and equating coefficients,

$$\begin{aligned}
\mathbf{x}' &= \left(\mathbf{x}-\left(\mathbf{x}\cdot\mathbf{n}\right)\mathbf{n}\right)\cos\varphi-\mathbf{n}\times\mathbf{x}\sin\varphi+\left(\mathbf{x}\cdot\mathbf{n}\right)\mathbf{n}\\
&= \mathbf{x}_{\parallel}+\mathbf{x}_{\perp}\cos\varphi-\left(\mathbf{n}\times\mathbf{x}_{\perp}\right)\sin\varphi
\end{aligned}$$

which is the same transformation as we derived from $SO\left(3\right)$.

This means that as φ 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 $SU\left(2\right)$ 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,

$$\exp\left(\frac{i\varphi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}\right)=\mathbf{1}\cos\frac{\varphi}{2}+i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\frac{\varphi}{2}$$

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

$$\psi'=e^{\frac{i\varphi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}}\psi$$

This transformation preserves the hermitian norm of ψ, since

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

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. Also notice that as φ runs from 0 to 2π, ψ changes by only

$$\begin{aligned}\psi' &= e^{i\pi\mathbf{n}\cdot\boldsymbol{\sigma}}\psi \\ &= \left(\mathbf{1}\cos\pi + i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\pi\right)\psi \\ &= -\psi\end{aligned}$$

so while a vector rotates by 2π, a spinor changes sign. A complete cycle of an $SU(2)$ transformation therefore requires φ to run through 4π.