$SU(2)$ and $SO(3)$

D.B. Westra

February 21, 2008

Abstract

We discuss some very basis properties of $SU(2)$ and $SO(3)$. In particular, we show the isomorphism $SU(2)/\mathbb{Z}_2 \cong SO(3)$. Along the way we discuss Pauli-matrices, Euler angles and give a parametrization of $SU(2)$.

1 The group of rotations

Consider the space $\mathbb{R}^3$ equipped with the standard inner product

$$x \cdot y = x_1 y_1 + x_2 y_2 + x_3 y_3 \, , \quad x = (x_1, x_2, x_3) \, , \; y = (y_1, y_2, y_3) \, .$$

The origin in $\mathbb{R}$ we denote O. A rotation is a smooth map $R : \mathbb{R} \to \mathbb{R}$ such that the origin in $\mathbb{R}$ such that R preserves the origin, angles, distances and orientations. Since a rotation preserves the origin, any combination of linearly dependent vectors is mapped to linearly dependent vectors. Consider two points A and B in $\mathbb{R}^3$. The origin O, A and B make up a triangle through the vectors $\overrightarrow{OA}$, $\overrightarrow{OB}$ and $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA}$. Now we consider the image of a rotation R. Since angles and distances have to be preserved the new triangle $\triangle OA'B'$, where A' is the image of A and B' is the image of B, is congruent to $\triangle OAB$ and all sides have the same length. Hence $\overrightarrow{A'B'} = \overrightarrow{OB'} - \overrightarrow{OA'}$. But then we must have that R is a linear map.

For a rotation to preserve angles and lengths it is sufficient and necessary to preserve the inner product. This follows from the observation that the angle between two vectors $\mathbf{a}$ and $\mathbf{b}$ is given by

$$\cos(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a} \cdot \mathbf{b}}{\sqrt{\mathbf{a} \cdot \mathbf{a}} \sqrt{\mathbf{b} \cdot \mathbf{b}}} \, .$$

In order to preserve the orientation a rotation needs to preserve the outer product (also called vector product). This can be seen as follows. Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ be three linearly independent vectors. The sign of the quantity $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$ is preserved under an orientation preserving map, since $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$ is precisely the determinant of the matrix formed by adjoining $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ to a 3×3 matrix. Since the inner product is already conserved it follows that the vector product is conserved as well.

Hence we arrive at a more manageable definition of a rotation as a linear transformation that preserve the inner product. If we identify a rotation R with the matrix representing R we see that R is a rotation if and only if R is a linear transformation $R : \mathbb{R} \to \mathbb{R}$ satisfying

$$R^T R = 1 \, , \quad \det R = 1$$

Note that 1 here denotes the 3×3 identity matrix in the first equation.

If R is a rotation we see in particular that $\det R \neq 0$. Therefore the rotations form a group; we denote this group by $SO(3)$ and call it the orthogonal group, or the rotation group.

If λ is an eigenvalue of a rotation R we see that $\lambda = \pm 1$. We want to show that there is always an eigenvector with eigenvalue 1. The characteristic polynomial of R has three roots λ_1, λ_2 and λ_3. The modulus of the roots has to be 1 and if there is a imaginary eigenvalue μ, then so is its conjugate $\bar{\mu}$ an eigenvalue. If all the three eigenvalues are real, then the only possibilities are that all three are 1 or that two are -1 and the third is 1. Let now λ_1 be imaginary and take $\lambda_2 = \bar{\lambda}_1$. Then $\lambda_3 = \frac{1}{\lambda_1 \bar{\lambda}_1}$ is real and positive and since it has to be of unit modulus $\lambda_3 = 1$. We see that there is always an eigenvalue 1. If R is a rotation

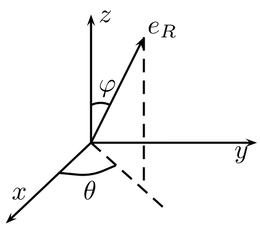

Figure 1: Rotation around e_R. Any axis is characterized by two angles φ and θ.

and not the identity there is just one eigenvector with eigenvalue one; we denote this eigenvector by e_R. We thus have $Re_R = e_R$ and if $R \neq 1$ then e_R is unique. The vector e_R determines a one-dimensional subspace of $\mathbb{R}$ that is left invariant under the action of R. We call this one-dimensional invariant subspace the axis of rotation.

If e_R is an axis of rotation we can always perform a change of coordinates such that e_R is the unit vector in the z-direction e_z. It is easy to see that this change of coordinates may be taken a rotation itself. Even more is true. Consider an arbitrary rotation with axis of rotation determined by e_R over an angle ψ and denote the rotation by $R(e_R, \psi)$. Call the angle between the plane in which e_R and the z-axis lie and the plane in xz-plane θ. Call the angle between e_R and the z-axis φ. See figure 1 The rotation can now be broken down into three rotations. First we use two rotations two go to a coordinate system with coordinates x', y' and z' in which the e_R points in the z'-direction, and then we rotate around the z'-axis over an angle ψ. The two rotation to go to the new coordinate system are: (a) a rotation around the z-axis around an angle θ to align the x-axis with the projection of e_R onto the xy-plane, (b) a rotation over an angle φ around the image of the y-axis under the first rotation. Hence we can write $R(e_R, \psi)$ as a product

$$R(e_R, \psi) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

This way we obtain a system of coordinates on the manifold $SO(3)$. The three angles are then called the Euler angles and they play an important role in the analysis of the spinning top and other mechanical description of rotations.

We note in particular the following: The group $SO(3)$ is generated by all elementary rotations $R_x(\alpha)$, $R_y(\beta)$ and $R_z(\gamma)$ given by

$$\begin{aligned} R_x(\alpha) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} , \quad R_y(\beta) = \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} , \\ R_z(\gamma) &= \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} . \end{aligned}$$

2 Complex rotations: $SU(2)$

After $\mathbb{C}$ the vector space $\mathbb{C}^2$ is the simplest example of a Hilbert space with an inner product given by

$$\langle \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} , \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \rangle = \bar{z}_1 w_1 + \bar{z}_2 w_2 .$$

The group $U(2)$ is the automorphism group of $\mathbb{C}^2$: it consists of linear transformations that preserve the inner product. Thus a 2×2 matrix U with complex entries represents an element of $U(2)$ if and only of $U^\dagger U = 1$. It follows that $\det U$ is a complex number of modulus 1. The group $U(1)$ consists of all complex numbers of modulus 1, which allows us to rephrase the previous finding as follows: Taking the determinant is a group homomorphism $\det : U(2) \to U(1)$. Indeed $\det UV = \det U \det V$ and $\det 1 = 1$. The kernel of the group homomorphism det is a subgroup of $U(2)$, which we provisionally denote N. It is even a normal subgroup, since if $\det U = 1$ and $V \in U(2)$ then $\det VUV^{-1} = \det V \det U \det V^{-1} = 1$. The image of det is the whole of $U(1)$ since if $u = e^{i\varphi} \in U(1)$ then the image of

$$U = \begin{pmatrix} e^{i\varphi/2} & 0 \\ 0 & e^{i\varphi/2} \end{pmatrix} \in U(2)$$

is clearly $e^{i\varphi}$. Hence $U(1) \cong U(2)/N$ as a group, and $U(2) = U(1) \times N$; $U(2)$ is not a simple group but is a direct product of N with the abelian group $U(1)$. The group N is called the special unitary group and denoted $SU(2)$.

Another way to see the group $SU(2)$ arise is to consider the subgroup of $U(2)$ consisting of linear transformation that leave the orientation invariant, or put in other words, that preserve the analytic two-form $\omega = dz_1 \wedge dz_2$. Equivalently, one can consider the subgroup of $U(2)$ that leave the symplectic form $\Omega : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ invariant, with

$$\Omega(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}) = z_1 w_2 - z_2 w_1 \,.$$

3 Parametrizing $SU(2)$

We now wish to show that $SU(2)$ is a real manifold that is isomorphic to the three sphere S^3. We do this by finding an explicit parametrization of $SU(3)$ in terms of two complex numbers x and y satisfying $|x|^2 + |y|^2 = 1$. If one splits up x and y in a real and imaginary parts, one sees that x and y define a point on S^3.

We write an element $U \in SU(2)$ as

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} .$$

Writing out the equation $U^\dagger U = 1$ and $\det U = 1$ one finds the following equations:

$$\begin{aligned} |a|^2 + |c|^2 = 1 \,, \qquad & |b|^2 + |d|^2 = 1 \,, \\ \bar{a}b + \bar{c}d = 0 \,, \qquad & ad - bc = 1 \,. \end{aligned}$$

We first assume $b = 0$ and find then that $ad = 1$ and $\bar{c}d = 0$, implying that $c = 0$ and U is diagonal with $a = \bar{d}$. Next we suppose $b \neq 0$ and use $a = -\frac{c\bar{d}}{\bar{b}}$ to deduce that $|b| = |c|$ and $|a| = |d|$; we thus have $b \neq 0 \Leftrightarrow c \neq 0$. We also see that we can use the ansatz

$$\begin{aligned} a &= e^{i\alpha} \cos\theta \,, \qquad b = e^{i\beta} \sin\theta \,, \\ c &= -e^{i\gamma} \sin\theta \,, \quad d = e^{i\delta} \cos\theta \,. \end{aligned}$$

Using again $a = -\frac{c\bar{d}}{\bar{b}}$ we see $\alpha + \delta = \beta + \gamma$ and writing out $ad - bc = 1$ we find $\alpha = -\delta$ and $\beta = -\gamma$. We thus see $a = \bar{d}$ and $b = -\bar{c}$. Hence the most general element of $SU(2)$ can be written as

$$U(x, y) = \begin{pmatrix} x & y \\ -\bar{y} & \bar{x} \end{pmatrix} , \quad \text{with} \;\; |x|^2 + |y|^2 = 1 \,.$$

The map $S^3 \to SU(2)$ given by $(x, y) \mapsto U(x, y)$ is clearly injective, and from the above analysis bijective. Furthermore the map is smooth. Hence we conclude that $SU(2) \cong S^3$ as a real manifold.

4 Pauli matrices

We introduce the Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

Note that the Pauli-matrices are precisely all the traceless Hermitian 2×2 complex matrices and make up a three-dimensional vector space. Therefore they provide a realization of the Lie algebra $su(2)$. Physicists in the advent of quantum mechanics wanted Lie algebra elements to be as Hermitian as possible, which was realized – if possible – by multiplying the Lie algebra elements by i. And since Pauli was a physicist, the Pauli-matrices are Hermitian and not anti-Hermitian as is often seen in the mathematics literature. Nowadays physicists work with both conventions ... It is easy to check that the Pauli-matrices satisfy the relations

$$\sigma^i\sigma^j = \delta^{ij} + i\epsilon^{ijk}\sigma^k \,,$$

where we also introduced the Levi–Civitá symbol, which is the totally anti-symmetric symbol on three indices with normalization $\epsilon^{123} = 1$. Hence if (ijk) is not a permutation of (123) it is zero and if (ijk) is a permutation of (123) then ϵ^{ijk} is the sign of the permutation. In particular we note the commutator relations[1]

$$[\sigma^i, \sigma^j] = \sigma^i\sigma^j - \sigma^j\sigma^i = 2i\epsilon^{ijk}\sigma^k \,,$$

which resembles the vector product in $\mathbb{R}^3$. We also note the identities

$$\mathrm{Tr}(\sigma^i) = 0 \,, \quad \mathrm{Tr}(\sigma^i\sigma^j) = 2\delta^{ij} \,, \quad \mathrm{Tr}([\sigma^i, \sigma^j]\sigma^k) = 4i\epsilon^{ijk} \,.$$

For every vector $\vec{x} \in \mathbb{R}$ we identify an element x of $su(2)$ as follows

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \leftrightarrow x = \sum_{i=1}^{3} x_i\sigma^i \,.$$

From now on we simply identify the elements with each other and thus write equality signs instead of arrows. We see that $\vec{x} \times \vec{y}$ corresponds to the element $\frac{1}{2i}[x, y]$;

$$\vec{x} \times \vec{y} = \frac{1}{2i}[x, y] \,.$$

And similarly we find

$$\vec{x} \cdot \vec{y} = \frac{1}{2}\mathrm{Tr}(xy) \,, \quad \vec{x} \times \vec{y} \cdot \vec{z} = \frac{1}{4i}\mathrm{Tr}([x, y]z) \,.$$

If U is an element of $SU(2)$ and $x \in su(2) \cong \mathbb{R}^3$ we see that

$$\mathrm{Tr}(UxU^{-1}) = 0 \,, \quad (UxU^{-1})^\dagger = (U^{-1})^\dagger xU^\dagger = UxU^{-1} \,,$$

and we conclude that U induces a map $\mathbb{R}^3 \to \mathbb{R}^3$. In fact, U defines a special orthogonal transformation since

$$\vec{x} \cdot \vec{y} \mapsto \mathrm{Tr}(UxU^{-1}UyU^{-1}) = \vec{x} \cdot \vec{y}$$

is invariant, so that U preserves the inner product on $\mathbb{R}^3$ and similarly $\vec{x} \times \vec{y} \cdot \vec{z}$ is invariant so that the action of U preserves the orientation. We thus found a map $R : SU(2) \to SO(3)$, whereby $U \in SU(2)$ gets mapped to the element $R(U)$ in $SO(3)$ corresponding to $x \mapsto UxU^{-1}$. Since $U_1(U_2xU_2^{-1})U_1^{-1} = (U_1U_2)x(U_1U_2)^{-1}$ the map $SU(2) \to SO(3)$ is a group homomorphism; $R(U_1U_2) = R(U_1)R(U_2)$.

Explicitly we find

$$\begin{aligned}
U(x, y)\sigma^1U(x, y)^{-1} &= \mathrm{Re}(x^2 - y^2)\sigma^1 - \mathrm{Im}(x^2 - y^2)\sigma^2 + 2\mathrm{Re}(x\bar{y})\sigma^3 \,, \\
U(x, y)\sigma^2U(x, y)^{-1} &= \mathrm{Im}(x^2 + y^2)\sigma^1 + \mathrm{Re}(x^2 + y^2)\sigma^2 + 2\mathrm{Im}(x\bar{y})\sigma^3 \,, \\
U(x, y)\sigma^3U(x, y)^{-1} &= -2\mathrm{Re}(xy)\sigma^1 + 2\mathrm{Im}(xy)\sigma^2 + (|x|^2 - |y|^2)\sigma^3 \,.
\end{aligned}$$

Therefore

$$R(U(x, y)) = \begin{pmatrix} \mathrm{Re}(x^2 - y^2) & \mathrm{Im}(x^2 + y^2) & -2\mathrm{Re}(xy) \\ -\mathrm{Im}(x^2 - y^2) & \mathrm{Re}(x^2 + y^2) & 2\mathrm{Im}(xy) \\ 2\mathrm{Re}(x\bar{y}) & 2\mathrm{Im}(x\bar{y}) & |x|^2 - |y|^2 \end{pmatrix} \,.$$

[1] The commutation relations provide the ultimate proof that the physicists' Lie algebra is not the same as that of the mathematicians, the transition being simply multiplying the elements by i.

We find

$$\begin{aligned}
R(U(\cos\alpha/2, -i\sin\alpha/2)) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} = R_x(\alpha)\,, \\
R(U(\cos\beta/2, -\sin\beta/2)) &= \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} = R_y(\beta)\,, \\
R(U(e^{-i\gamma/2}, 0)) &= \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} = R_z(\gamma)\,,
\end{aligned}$$

and hence the map $R : SU(2) \to SO(3)$ is surjective. Suppose now that $U(x, y)$ is mapped to the identity element in $SO(3)$. We see then that $x\bar{y} = 0$, so that either $x = 0$ or $y = 0$. Since also $|x|^2 - |y|^2 = 1$, we cannot have $x = 0$ and hence $y = 0$. Furthermore from $\mathrm{Re}x^2 = |x|^2 = 1$ we see $x = \pm 1$: indeed we see that $R(U(x, y)) = R(U(-x, -y))$. The kernel of $R : SU(2) \to SO(3)$ is thus given by ± 1 times the identity matrix, which is a $\mathbb{Z}_2$-subgroup of $SU(2)$. As any kernel of group homomorphisms, the kernel is a normal subgroup. All in all we have shown

$$SU(2) \cong SO(3)/\mathbb{Z}_2\,.$$