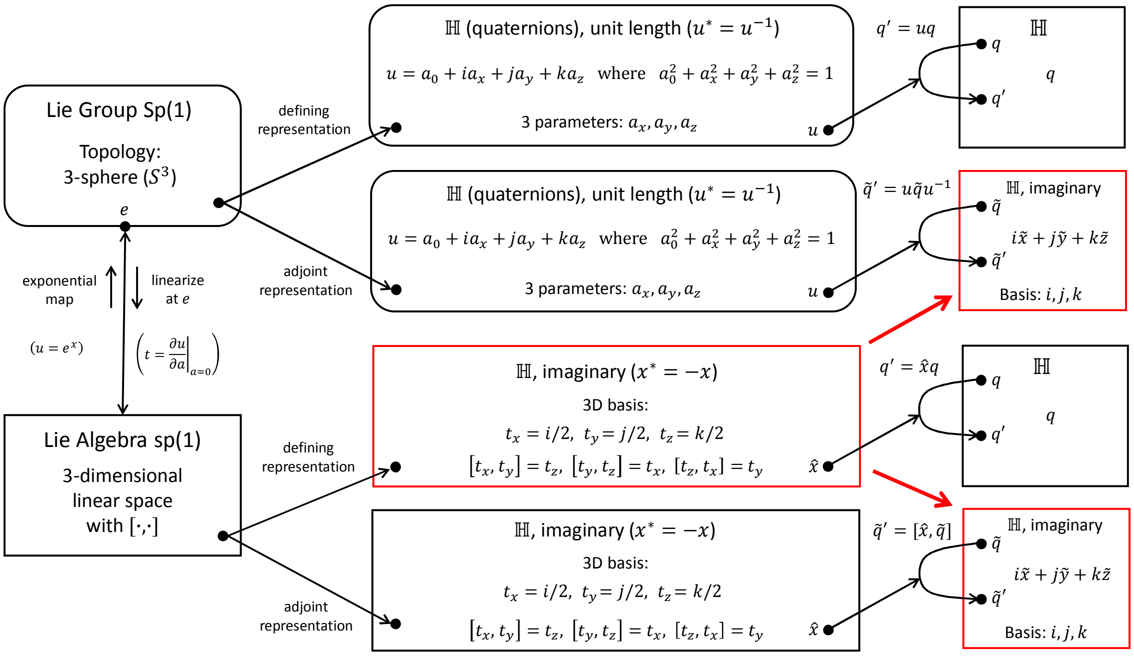
3.31 Sp(1): The Group of Quaternions of Unit Length



We studied transformations that preserve the (Euclidean) inner product $x^T y$ of two real vectors ($x, y \in \mathbb{R}^n$). We also studied transformations that preserve the (Hermitian) inner product $\psi^\dagger \phi$ of two complex vectors ($\psi, \phi \in \mathbb{C}^n$). What's next? Transformations that preserve the inner product $p^\dagger q$ of two *quaternionic* vectors ($p, q \in \mathbb{H}^n$, where the letter H stands for Hamilton who discovered the quaternions)! Transformations of the first type are called orthogonal (= real unitary) and form the group O(n), those of the second type are called unitary and form the group U(n), finally, those of the third type are called quaternionic unitary and form the group U(n, $\mathbb{H}$). The latter group is more commonly known as the *compact symplectic group* Sp(n). (See *John Baez: Symplectic, Quaternionic, Fermionic*, https://math.ucr.edu/home/baez/symplectic.html for an explanation of this name.)

The quaternions are a generalization of the complex numbers: instead of $a + ib$ we have $q = w + ix + jy + kz$, where $w, x, y, z \in \mathbb{R}$, that is, the imaginary part now has *three* pieces. Quaternionic conjugation changes the sign of the entire imaginary part: $q^* = w - ix - jy - kz$. To multiply two quaternions, we need to know all possible products of the three imaginary units i, j, and k. Each unit multiplied by itself yields minus one, $i^2 = j^2 = k^2 = -1$, just like for the complex numbers. The remaining six mixed products anticommutes and are given by $ij = k, ji = -k, jk = i, kj = -i, ki = j, ik = -j$.

The defining representation of Sp(1) consists of the quaternions u that preserve the inner product, $(up)^*(uq) = p^*q$, for any pair of quaternions p and q. Rewriting as $p^*u^*uq = p^*q$, we see that this is the case if $u^*u = 1$. Thus, u must be a *unit-length quaternion*: $u = a_0 + ia_x + ja_y + ka_z$, where $a_0^2 + a_x^2 + a_y^2 + a_z^2 = 1$ and $a_i \in \mathbb{R}$. These unit-length quaternions act on the general quaternions, $q = w + ix + jy + kz$, in the representation space: $q' = uq$ (see the upper branch of the diagram).

What is so special about Sp(1)? It turns out that Sp(1) is isomorphic to SU(2)! Thus, Sp(1) provides another way of looking at spinorial rotations: rather than "unitary 2×2 matrices with determinant one", we can also use "unit quaternions".

To show that Sp(1) and SU(2) are isomorphic, we need to establish a dictionary translating between the two groups. If we identify the quaternion $q = w + ix + jy + kz$ with the complex 2×2 matrix

$$Q = wI - i(x\sigma_x + y\sigma_y + z\sigma_z) = \begin{pmatrix} w - iz & -y - ix \\ y - ix & w + iz \end{pmatrix},$$

quaternion multiplication and matrix multiplication do the same thing! The three imaginary units of the quaternions behave just like the three Pauli matrices (times $-i$). For example, $ij = k$ translates to $(-i\sigma_x)(-i\sigma_y) = (-i\sigma_z)$; similarly, $k^2 = -1$ translates to $(-i\sigma_z)(-i\sigma_z) = -I$. Both translations are true statements. Next, we identify the unit quaternion $u = a_0 + ia_x + ja_y + ka_z$, which is an element of Sp(1), with the complex matrix

$$U = a_0 I - i(a_x\sigma_x + a_y\sigma_y + a_z\sigma_z) = \begin{pmatrix} a_0 - ia_z & -a_y - ia_x \\ a_y - ia_x & a_0 + ia_z \end{pmatrix} \quad \text{where} \;\; a_0^2 + a_x^2 + a_y^2 + a_z^2 = 1.$$

But this is exactly the SU(2) matrix we encountered earlier when discussing 3-sphere parameters! For example, the unit quaternion $u(\theta) = \cos(\theta/2) + k\sin(\theta/2)$ translates to the SU(2) matrix

$$U(\theta) = \begin{pmatrix} \cos(\theta/2) - i\sin(\theta/2) & 0 \\ 0 & \cos(\theta/2) + i\sin(\theta/2) \end{pmatrix} = \begin{pmatrix} \exp(-i\theta/2) & 0 \\ 0 & \exp(i\theta/2) \end{pmatrix},$$

which rotates a spinor about the z-axis. (Dictionary: $a_0 = \cos(\theta/2)$, $a_x = a_y = 0$, $a_z = \sin(\theta/2)$)

With this correspondence established, we can translate a unit quaternion acting on an arbitrary quaternion, $q' = uq$, to an SU(2) matrix U acting on a matrix Q of the form defined above, $Q' = UQ$. But there is one small problem: we are not used to an SU(2) matrix that acts on a *matrix*, $Q' = UQ$, we expect it to act on a *spinor*, $\psi' = U\psi$. To get rid of this discrepancy, we keep only the first column of Q, that is, we identify q with the spinor $\psi = \begin{pmatrix} w - iz \\ y - ix \end{pmatrix}$. The second column of Q is redundant.

Let's have a look at the Lie algebra sp(1). Taking the derivatives of $u = \sqrt{1 - a_x^2 - a_y^2 - a_z^2} + ia_x + ja_y + ka_z$ with respect to a_x, a_y, a_z and then evaluating the results at the identity ($a_x = a_y = a_z = 0$ and thus $a_0 = 1$) yields the three basis generators i, j, k. Thus, the Lie algebra consists of purely imaginary quaternions (= quaternions with no real part). For example, the generator of rotation about the z-axis is $\hat{x} = k \in$ sp(1) and translates to $X = -i\sigma_z \in$ su(2). Finally, normalizing the basis generators such that they satisfy the commutation relations of su(2) yields $i/2$, $j/2$, $k/2$ (see the diagram).

Can we use quaternions to rotate real 3D vectors (as opposed to spinors)? Yes, we know from our discussion of SU(2) that the *adjoint representation* can do that! We take as our representation space a copy of the Lie algebra (red arrows in the diagram), which is a 3-dimensional vector space over the reals with the basis i, j, k. (In contrast, the representation space of the defining representation can be regarded as either as a 1-dimensional vector space over the *quaternions* [basis: 1] or, equivalently, as a 4-dimensional vector space over the *reals* [basis: $1, i, j, k$]). To rotate the real 3D vector $(\tilde{x}, \tilde{y}, \tilde{z})$, we pack it into the imaginary part of the quaternion $\tilde{q} = i\tilde{x} + j\tilde{y} + k\tilde{z}$; then, we conjugate it with the unit quaternion u that encodes the rotation axis and the rotation angle: $\tilde{q}' = u\tilde{q}u^{-1}$; finally, we unpack the rotated vector from the imaginary part of $\tilde{q}' = i\tilde{x}' + j\tilde{y}' + k\tilde{z}'$. Interestingly, this method has numerical advantages over the more straightforward method of using real 3×3 matrices and thus finds practical applications in computer graphics.