ME 115(a): Solution to Homework #3
(Winter, 2009/2010)

Problem 1:

Part (a): Elements of $SU(2)$ have the form:

$$\begin{bmatrix} \mathbf{z} & \mathbf{w} \\ -\mathbf{w}^* & \mathbf{z}^* \end{bmatrix} = \begin{bmatrix} (a+ib) & (c+id) \\ -(c-id) & (a-ib) \end{bmatrix}$$

where $zz^* + ww^* = a^2 + b^2 + c^2 + d^2 = 1$. To show that the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

form a basis for $SU(2)$, let A, B, C, and D be real numbers. Then, the matrix formed by the product of A, B, C,and D with these matrices is:

$$A \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + B \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} + C \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + D \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = \begin{bmatrix} A+iB & C+iD \\ C-iD & A-iB \end{bmatrix}$$

is a matrix in $SU(2)$ for any choice of A, B, C, and D where $A^2 + B^2 + C^2 + D^2 = 1$. Thus these four basis matrices for $SU(2)$ are in 1-to-1 correspondence with the 1, i, j, and k basis elements for the quaternions. Thus, the scalar elements A, B, C, and D are in one-to-one correspondence with the scalar elements of unit quaternions. That is, let a unit quaternion be represented by $q = \lambda_1 + \lambda_2 i + \lambda_3 j + \lambda_4 k = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$. The correspondence is then:

$$\lambda_1 = A = Re(z) = \frac{z + z^*}{2} \quad (1)$$

$$\lambda_2 = B = Im(z) = \frac{i(z^* - z)}{2} \quad (2)$$

$$\lambda_3 = C = Re(w) = \frac{w + w^*}{2} \quad (3)$$

$$\lambda_4 = D = Im(w) = \frac{i(w^* - w)}{2} \quad (4)$$

Part (b): The unit quaternion elements are in one-to-one correspondence with the Euler parameters of a rotation: $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (\cos\frac{\phi}{2}, \omega_x \sin\frac{\phi}{2}, \omega_y \sin\frac{\phi}{2}, \omega_z \sin\frac{\phi}{2})$. ϕ is the rotation about an axis represented by a unit vector $\vec{\omega} = [\omega_x\ \omega_y\ \omega_z]^T$. A 2×2 complex matrix which represents an arbitrary rotation as a function of the z-y-x Euler angles can be developed as the product of 2×2 complex matrices which represent rotations about the z, y, and x axes. A rotation about the x-axis of amount γ has the 2×2 representation (since $\lambda_1 = \cos\frac{\gamma}{2}$,

$\lambda_2 = \sin\frac{\gamma}{2}$, $\lambda_3 = \lambda_4 = 0$):

$$\begin{bmatrix} (\cos\frac{\gamma}{2} + i\sin\frac{\gamma}{2}) & 0 \\ 0 & (\cos\frac{\gamma}{2} - i\sin\frac{\gamma}{2}) \end{bmatrix} = \begin{bmatrix} e^{i\frac{\gamma}{2}} & 0 \\ 0 & e^{-i\frac{\gamma}{2}} \end{bmatrix}$$

Similarly, a rotation of amount ϕ about the y-axis can be represented as:

$$\begin{bmatrix} \cos\frac{\phi}{2} & \sin\frac{\phi}{2} \\ -\sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{bmatrix}$$

while a rotation of amount ψ about the z-axis can be represented as:

$$\begin{bmatrix} \cos\frac{\psi}{2} & i\sin\frac{\psi}{2} \\ i\sin\frac{\psi}{2} & \cos\frac{\psi}{2} \end{bmatrix}$$

The product of these matrices yields the result.

Part (c):

$$\phi = 2\cos^{-1}(a) = 2\cos^{-1}(\frac{\mathbf{z}+\mathbf{z}^*}{2}) \tag{5}$$

$$\omega_x = \frac{b}{\sqrt{b^2+c^2+d^2}} = \frac{(\mathbf{z}-\mathbf{z}^*)/2}{\sqrt{(\frac{\mathbf{z}-\mathbf{z}^*}{2})^2 + \mathbf{w}\mathbf{w}^*}} \tag{6}$$

$$\omega_y = \frac{c}{\sqrt{b^2+c^2+d^2}} = \frac{(\mathbf{w}+\mathbf{w}^*)/2}{\sqrt{(\frac{\mathbf{z}-\mathbf{z}^*}{2})^2 + \mathbf{w}\mathbf{w}^*}} \tag{7}$$

$$\omega_z = \frac{d}{\sqrt{b^2+c^2+d^2}} = \frac{(\mathbf{w}-\mathbf{w}^*)/2}{\sqrt{(\frac{\mathbf{z}-\mathbf{z}^*}{2})^2 + \mathbf{w}\mathbf{w}^*}} \tag{8}$$

Problem 2: Can every orthogonal matrix be represented by the exponential of a real matrix?

You should have either remembered or derived the fact that $det(e^C) = e^{tr(C)}$, where $tr(C)$ is the trace of determinant C. Note that if $tr(C)$ is real, than $e^{tr(C)}$ is always a positive number and therefore orthogonal matrices with determinant -1 can not be represented as a matrix exponential. This arises from the fact that the Orthogonal Group is a disconnected group. That is, the matrices with +1 determinant are all connected to each other, and similarly the ones with -1 determinant. But, the two subsets are disjoint.

Note, that if $tr(C) = \frac{\pi}{2}i$ (where $i^2 = -1$), then $det(e^C) = -1$. However, this can not be true if C is real. Recall that the trace of a matrix is equal to the sum of its eigenvalues. Let C be a $n \times n$ matrix. If n is even, then all of the eigenvalues of C must be complex cojugates and or an even number of real roots. Thus, the sum of the eigenvalues must be real. Similarly, if

n is odd, the eigenvalues will either be: (1) an odd number of real eigenvalues, or (2) an odd number of real eigenvalues and an even number of complex conjugate complex eigenvalues. In either case, the sum of the eigenvalues must be real number. Thus, if C is a real matrix (as specified in part (a)), then $tr(C)$ must be real, and therefore e^C can not represent orthogonal matrices with determinant -1.

Problem 3: (Problem 5(c) of chapter 2 in the MLS text). There are at least two ways to solve this problem. The simplest way is to use the result of part 5(b) quoted in the text:

$$R = \frac{1}{1+||a||^2}\begin{bmatrix} 1+a_1^2-a_2^2-a_3^2 & 2(a_1a_2-a_3) & 2(a_1a_3+a_2) \\ 2(a_1a_2+a_3) & 1-a_1^2+a_2^2-a_3^2 & 2(a_2a_3-a_1) \\ 2(a_1a_3-a_2) & 2(a_2a_3+a_1) & 1-a_1^2-a_2^2+a_3^2 \end{bmatrix} \tag{9}$$

where $||a||^2$ is shorthand notation for $||a||^2 = a_1^2+a_2^2+a_3^2$. Noting that

$$trace(R) = \frac{3-||a||^2}{1+||a||^2} \, .$$

This equation can be solved to obtain an expression for $||a||^2$:

$$||a||^2 = \frac{3-trace(R)}{1+trace(R)} = \frac{3-r_{11}-r_{22}-r_{33}}{1+r_{11}+r_{22}+r_{33}} \, .$$

Simple algebraic manipulation of the off-diagonal term of R yield

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \frac{1+||a||^2}{4}\begin{bmatrix} r_{32}-r_{23} \\ r_{13}-r_{31} \\ r_{21}-r_{12} \end{bmatrix}$$

If you didn't use the results of 5(b) in the text, then you would have started with Cayley's formula

$$R = (I-\hat{a})^{-1}(I+\hat{a}); \qquad \text{where } \hat{a} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

and derived Equation (9) from this expression.

Problem 4: Problem 6(a,d,e) in Chapter 2 of MLS.

Part (a): Let Q and P be unit quaternions—i.e., $QQ^* = PP^* = 1$, where Q^* is the quaternion conjugate of Q. The set of unit quaternions is a group if you can show that:

(i) quaternion multiplication is a binary group operation (i.e., the product of two elements in the group yields a group element),

(ii) the group operation (quaternion multiplication) is associative

(ii) the set contains an identity element;

(iii) every group element has an inverse element, and the inverse element is in the group.

Each of these requirements can be shown as follows:

(i) The product of unit quaternions, QP, is a unit quaternion: $QP(QP)^* = QPP^*Q^* = QQ^* = 1$.

(ii) If Q_1, Q_2, and Q_3 are unit quaternions, then you must show that $Q_1\ (Q_2\ Q_3) = (Q_1\ Q_2)\ Q_3$. While there are many ways to show this, one can simply observe that since the four basis elements of a unit quarternion, $(1, i, j, k)$, satisfy the associative multiplication requirement, the quaternions themselves will satisfy this requirement.

(iii) The identity quaternion is: $e = (1, 0, 0, 0) = 1 + 0 \cdot i + 0 \cdot j + 0 \cdot k$.

(iv) The inverse of any unit quaternion Q is:

$$Q^{-1} = \frac{Q^*}{QQ^*} = Q^*$$

which is also a unit quaternion since

$$Norm(Q^{-1}) = Norm(Q^*) = Q^*(Q^*)^* = Q^*Q = (QQ^*)^* = 1^* = 1.$$

Part (d):

(i) If $A_1, A_2 \in SO(3)$, then each of the 9 elements in the product matrix $A_1\ A_2$ requires 3 multiplications and 2 additions. Hence, the product $A_1\ A_2$ requires a total of 27 multiplications and 18 additions.

(ii) Let q_1 and q_2 be quaternions, with respective real and vector parts q_{1R}, q_{2R} and $\vec{q}_{1P}$, $\vec{q}_{2P}$. The real part of the quaternion product, $q_{1R}q_{2R} - \vec{q}_{1P} \cdot \vec{q}_{2P}$, requires 4 multiplications and 3 additions (where the subtraction is counted as an addition). The pure part, $\vec{q}_{3P} = q_{1R}\vec{q}_{2P} + q_{2R}\vec{q}_{1P} + \vec{q}_{1P} \times \vec{q}_{2P}$, can be evaluated in 12 multiplications and 9 additions. Thus, the quaternion product requires a total of 16 multiplications and 12 additions. It is therefore more efficient than the equivalent matrix multiplication.

(iii) The rotation of a vector by multiplication of a 3×3 rotation matrix times a 3×1 vector requires only 9 multiplications and 6 additions.

(iv) The number of multiplications and additions for the equivalent quaternion operation will depend upon the form which one uses for the quaternion vector rotation. Using the identity $1 = q_R^2 + q_P \cdot q_P$, it is possible to show that the vector part of $q\tilde{x}q^{-1}$ in part (b) above can be rearranged to the form:

$$\vec{x} + 2[q_P \times (q_P \times \vec{x}) + q_R(q_P \times \vec{x})]$$

Since $q_P \times \vec{x}$ need only be evaluated once, this takes only 18 multiplications and 12 additions. However, no matter what form one tries, the quaternion approach will always take more operations than the matrix/vector approach for vector rotation.

Part (e): Let a body rotate about a fixed axis $\vec{\omega}$ by an angle $\phi(t)$. Let the unit quaternion $Q(t)$ denote the orientation of this rotating body. This quaternion can be represented in the "vector" form

$$Q(t) = \left(\cos(\frac{\phi(t)}{2}), \vec{\omega}\sin(\frac{\phi(t)}{2})\right) . \tag{10}$$

Thus, the time rate of change of this quaternion is:

$$\dot{Q}(t) = \left(-\frac{\dot{\phi}(t)}{2}\sin(\frac{\phi(t)}{2}), \frac{\dot{\phi}(t)}{2}\vec{\omega}\cos(\frac{\phi(t)}{2})\right) . \tag{11}$$

Recall that if $Q_1 = (Q_{1,r}, \vec{Q}_{1,v})$ and $Q_2 = (Q_{2,r}, \vec{Q}_{2,v})$ (in the convenient "vector" notation for a quaternion), then

$$Q_1 \cdot Q_2 = (Q_{1,r}Q_{2,r} - \vec{Q}_{1,v} \cdot \vec{Q}_{2,v}, Q_{1,r}\vec{Q}_{2,v} + Q_{2,r}\vec{Q}_{1,v} + \vec{Q}_{1,v} \times \vec{Q}_{2,v}) . \tag{12}$$

Substituting the appropriate forms of Equations (10) and (11) into Equation (12) yields:

$$\begin{aligned} \dot{Q} \cdot Q^* &= \left(-\frac{\dot{\phi}}{2}\sin(\frac{\phi}{2}), \frac{\dot{\phi}}{2}\vec{\omega}\cos(\frac{\phi}{2})\right)\left(\cos(\frac{\phi}{2}), -\vec{\omega}\sin(\frac{\phi}{2})\right) & (13) \\ &= \left(0, \frac{\dot{\phi}}{2}\vec{\omega}\right) & (14) \end{aligned}$$

Problem 5: Problem 8(b,c) in Chapter 2 of MLS.

Part(b):

$$\begin{aligned} e^{g\Lambda g^{-1}} &= I + \frac{1}{1!}g\Lambda g^{-1} + \frac{1}{2!}(g\Lambda g^{-1})^2 + \frac{1}{3!}(g\Lambda g^{-1})^3 + \cdots \\ &= I + \frac{1}{1!}g\Lambda g^{-1} + \frac{1}{2!}(g\Lambda^2 g^{-1}) + \frac{1}{3!}(g\Lambda^3 g^{-1}) + \cdots \\ &= g(I + \frac{1}{1!}\Lambda + \frac{1}{2!}\Lambda^2 + \frac{1}{3!}\Lambda^3 + \cdots)g^{-1} \\ &= g e^{\Lambda} g^{-1} \end{aligned}$$

part (c): Assuming that Λ is constant and θ is a function of time:

$$
\begin{aligned}
\frac{d}{dt}e^{\Lambda\theta} &= \frac{d}{dt}(I + \frac{1}{1!}\theta\Lambda + \frac{1}{2!}\theta^2\Lambda^2 + \cdots) \\
&= \frac{1}{1!}\dot{\theta}\Lambda + \frac{1}{2!}\dot{\theta}\theta\Lambda^2 + \cdots \\
&= \dot{\theta}\Lambda(I + \frac{1}{1!}\theta\Lambda + \cdots) = \dot{\theta}\Lambda e^{\Lambda\theta}
\end{aligned}
$$

Problem 6: Problem 9(a,b) in Chapter 2 of MLS.

Part (a): Part (a) of this problem first asked you to show that:

$$\text{image}(\hat{\omega}) \;=\; N_\omega^\perp$$

where $N_\omega^\perp$ is the orthogonal complement to the space spanned by $\vec{\omega}$. First note that $N_\omega^\perp$ is the subspace of $\mathbb{R}^3$ spanned by all vectors orthogonal to $\vec{\omega}$:

$$N_\omega^\perp \;=\; \{\vec{z} \in \mathbb{R}^3 \mid \vec{z} \cdot \vec{\omega} = 0\}.$$

Next, note that image($\hat{\omega}$) is:

$$\text{image}(\hat{\omega}) = \{\vec{v} \mid \vec{v} = \hat{\omega}\vec{z};\ \forall\ \vec{z}\ \in\ \mathbb{R}^3\}\ .$$

In order to show that image($\hat{\omega}$) $= N_\omega^\perp$, it is sufficient to show that the following two conditions hold: (i) image($\hat{\omega}$) $\subset N_\omega^\perp$; and (ii) $N_\omega^\perp \subset$ image($\hat{\omega}$).

Let $\vec{v}^*$ be a vector in image($\hat{\omega}$). Then $\vec{v}^* = \hat{\omega}\vec{z}$ for some vector $\vec{z}$. But, since $\vec{\omega}\cdot\vec{v}^* = \vec{\omega}\cdot(\hat{\omega}\vec{z}) = 0$, then $\vec{v}^*\ \in\ N_{\hat{\omega}}^\perp$. Hence, we can conclude that:

$$\text{image}(\hat{\omega}) \subset N_\omega^\perp.$$

To show the other inclusion, consider a vector $\vec{y}\ \in\ N_\omega^\perp$. It must be true that $\vec{y}\cdot\vec{\omega} = 0$. Let $\vec{p} = \vec{y} \times \vec{\omega}$. Then

$$\hat{\omega}\vec{p} = \hat{\omega}(\vec{y} \times \vec{\omega}) = -\hat{\omega}^2\vec{y} = -\vec{\omega} \times (\vec{\omega} \times \vec{y}) = \vec{y}\ .$$

Thus, we have that $\vec{y}\ \in$ image($\hat{\omega}$). Since the choice of the vector $\vec{y}$ was arbitrary, we have that $N_\omega^\perp\ \subset$ image($\hat{\omega}$). Hence, $N_\omega^\perp\ =$ image $\hat{\omega}$.

Part (a) of this problem also asked you to show that:

$$\text{kernel}(\hat{\omega}) \;=\; N_\omega$$

where N_ω is the space spanned by $\vec{\omega}$: $N_\omega = \{\vec{z} \mid \vec{z} = \alpha\vec{\omega},\ \forall\ \alpha\ \in\ \mathbb{R}\}$. Recall that the kernel of a linear map (a matrix) A is the set: kernel(A) $= \{\vec{z} \mid A\vec{z} = 0\}$. Hence:

$$
\begin{aligned}
\text{kernel}(\hat{\omega}) &= \{\vec{v} \in \mathbb{R}^3 \mid \hat{\omega}\vec{v} = 0\} \\
&= \{\vec{v} \in \mathbb{R}^3 \mid \vec{\omega} \times \vec{v} = 0\} \\
&= \{\vec{v} = \lambda\vec{\omega} \mid \lambda\ \in\ \mathbb{R}\} \\
&= N_\omega
\end{aligned}
$$

Part (b): Recall from the problem statement that in order to show that an operator P_V on a linear vector space V is a projection, one must show:

condition 1: $P_V(x) = x$ for any $x \in V$.

condition 2: $P_V(V) = V$ for the entire space V.

To show that $P_{N_\omega} = \vec{\omega}\vec{\omega}^T$ is a projection map, we must show that $P_{N_\omega}(x \in N_\omega) = x$ and image$(P_{N_\omega}) = N_\omega)$. To show the first condition, let $\vec{v} = \lambda\vec{\omega}$ for some $\lambda \in \mathbb{R}$. The vector $\vec{v}$ clearly lies in N_ω. Then:

$$P_{N_\omega}(\vec{v}) = \vec{\omega}\vec{\omega}^T\vec{v} = \vec{\omega}\vec{\omega}^T(\lambda\vec{\omega}) = \lambda\vec{\omega} = \vec{v}\ .$$

To show the second requirement,

$$\begin{aligned} \text{image}(P_{N_\omega}) &= \{\vec{\omega}\vec{\omega}^T\vec{z} \mid \vec{z} \in \mathbb{R}^3\} \\ &= \{\alpha\vec{\omega} \mid \alpha \in \mathbb{R}\} \\ &= N_\omega \end{aligned}$$

To show that operator $P_{N_\omega^\perp} = (I - \vec{\omega}\vec{\omega}^T)$ is projection, we must similarly show that conditions 1 and 2 hold. To show the first condition, let $\vec{v} \in N_\omega^\perp$: $\vec{v}^T\vec{\omega} = 0$. Then

$$P_{N_\omega^\perp}(\vec{v}) = (I - \vec{\omega}\vec{\omega}^T)\vec{v} = \vec{v} - \vec{\omega}(\vec{\omega}^T\vec{v}) = \vec{v}\ .$$

To show the second condition, let $\vec{z} \in$ image$(P_{N_\omega^\perp})$. There must exist a vector $\vec{y} \in \mathbb{R}^3$ such that $\vec{z} = (I - \vec{\omega}\vec{\omega}^T)\vec{y}$. For any such vector $\vec{z}$:

$$\vec{\omega} \cdot \vec{z} = \vec{\omega} \cdot (I - \vec{\omega}\vec{\omega}^T)\vec{y} = \vec{\omega} \cdot \vec{y} - (\vec{\omega} \cdot \vec{\omega})\vec{\omega}^T\vec{y} = 0\ .$$

Thus, image$(P_{N_\omega^\perp}) \subset N_\omega^\perp$. Next we want to show that $N_\omega^\perp \subset$ image$(P_{N_\omega^\perp})$. To do this, let $\vec{v} \in N_\omega^\perp$. Then

$$P_{N_\omega^\perp}(\vec{v}) = (I - \vec{\omega}\vec{\omega}^T)\vec{v} = \vec{v} - \vec{\omega}(\vec{\omega}^T\vec{v}) = \vec{v} - 0 = \vec{v}\ .$$

Since this is true for any $\vec{v} \in N_\omega^\perp$, then $N_\omega^\perp \subset$ image$(P_{N_\omega^\perp})$. Hence, $N_\omega^\perp =$ image$(P_{N_\omega^\perp})$, and thus $P_{N_\omega^\perp}$ is a projection.

Problem 7: Let Z-X-Y Euler angles be denoted by ψ, ϕ, and γ. First let's develop an expression for the net rotation due to Z-Y-X rotations. First find expressions for each of the individual rotation matrices:

$$R_\psi = Rot(\vec{z}, \psi) = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad R_\phi = Rot(\vec{y}, \phi) = \begin{bmatrix} \cos\phi & 0 & \sin\psi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{bmatrix}$$

$$R_\gamma = Rot(\vec{x}, \gamma) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{bmatrix}$$

where the notation $R(\vec{p}, \alpha)$ means rotation by angle α about axis $\vec{p}$. The final expression is obtained by multiplying these matrices:

$$R(\psi, \phi, \gamma) = R_\psi R_\phi R_\gamma = \begin{bmatrix} c\psi c\phi & c\psi s\phi s\gamma - s\psi c\gamma & c\psi s\phi c\gamma + s\psi s\gamma \\ s\psi c\phi & s\psi s\phi s\gamma + c\psi c\gamma & s\psi s\phi c\gamma - c\psi s\gamma \\ -s\phi & c\phi s\gamma & c\phi c\gamma \end{bmatrix}$$

where $s\psi = \sin\psi$, $c\phi = \cos\phi$, etc.

To find an expression for the angles ψ, ϕ, and γ as a function of the a_{ij} in the matrix

$$R = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

note that $\sin\phi = -a_{31}$. There are two solutions to this equation: $\phi_1 = \sin^{-1}(a_{31})$ and $\phi_2 = \pi - \phi_1$. Manipulation of the (a_{11}, a_{21}) and (a_{32}, a_{33}) terms yields:

$$\psi = Atan2[\frac{a_{21}}{\cos\phi}, \frac{a_{11}}{\cos\phi}]; \qquad \gamma = Atan2[\frac{a_{32}}{\cos\phi}, \frac{a_{33}}{\cos\phi}].$$