1 Cayley transforms

1.1 A few general properties

1. We recall that $N \in \mathbb{R}^{n\times n}$ and by assumption that $|\mathrm{I} - N| \neq 0$. There are many ways to prove that the matrices $(\mathrm{I}+N)$ and $(\mathrm{I}-N)^{-1}$ commute (one way is just show that both matrices $(\mathrm{I}+N)$ and $(\mathrm{I}-N)^{-1}$ have same eigenvectors hence they will commute). One is to show that for any matrix A and any invertible matrix B, A and B commute if and only if A and B^{-1} commute ($AB = BA \Leftrightarrow B^{-1}A = AB^{-1}$). Then $(\mathrm{I}-N)(\mathrm{I}+N) = (\mathrm{I}-N^2) = (\mathrm{I}+N)(\mathrm{I}-N)$. Or by another a direct computation we have:

$$\begin{aligned}
(\mathrm{I}+N)(\mathrm{I}-N)^{-1} &= -(-2\mathrm{I}+(\mathrm{I}-N))(\mathrm{I}-N)^{-1} && (1)\\
&= -\mathrm{I}+2(\mathrm{I}-N)^{-1} && (2)\\
&= -(\mathrm{I}-N)^{-1}(-2\mathrm{I}+(\mathrm{I}-N)) && (3)\\
&= (\mathrm{I}-N)^{-1}(\mathrm{I}+N). && (4)
\end{aligned}$$

2. *Ad absurdum*, if $(\mathrm{I}+M)$ is not invertible, then there exists a non-trivial (i.e. non zero) vector $\mathbf{v} \neq \mathbf{0}$ such that $(\mathrm{I}+M)\,\mathbf{v} = \mathbf{0}$. This implies that

$$\begin{aligned}
M\,\mathbf{v} = -\mathbf{v} \quad &\Leftrightarrow \quad (\mathrm{I}+N)(\mathrm{I}-N)^{-1}\,\mathbf{v} = -\mathbf{v}, && (5)\\
&\Leftrightarrow \quad (\mathrm{I}-N)^{-1}(\mathrm{I}+N)\,\mathbf{v} = -\mathbf{v}, && (6)\\
&\Leftrightarrow \quad (\mathrm{I}+N)\,\mathbf{v} = -(\mathrm{I}-N)\,\mathbf{v}, && (7)\\
&\Leftrightarrow \quad \mathbf{v}+N\,\mathbf{v} = -\mathbf{v}+N\,\mathbf{v}, \quad \Leftrightarrow \quad \mathbf{v} = -\mathbf{v}, && (8)
\end{aligned}$$

a contradiction.

3. We simply rewrite the expression for M as an expression for N:

$$\begin{aligned}
M = (\mathrm{I}+N)(\mathrm{I}-N)^{-1} \quad &\Leftrightarrow \quad M(\mathrm{I}-N) = (\mathrm{I}+N) && (9)\\
&\Leftrightarrow \quad M-\mathrm{I} = (\mathrm{I}+M)N \overset{1.1.2}{\Leftrightarrow} N = (M+\mathrm{I})^{-1}(M-\mathrm{I}). && (10)
\end{aligned}$$

4. By mimicking the above computation we have

$$\begin{aligned}
M = (\mathrm{I}-P)(\mathrm{I}+P)^{-1} \quad &\Leftrightarrow \quad M(\mathrm{I}+P) = (\mathrm{I}-P) && (11)\\
&\Leftrightarrow \quad \mathrm{I}-M = (M+\mathrm{I})P && (12)\\
&\overset{1.1.2}{\Leftrightarrow} \quad P = (\mathrm{I}+M)^{-1}(\mathrm{I}-M) && (13)\\
&\Leftrightarrow \quad P = (\mathrm{I}-M)(\mathrm{I}+M)^{-1} && (14)
\end{aligned}$$

5. First note that if S is skew then $\mathrm{I}-S$ is invertible, because $(\mathrm{I}-S)\mathbf{v} = \mathbf{0} \Rightarrow \mathbf{v}\cdot\mathbf{v} = \mathbf{v}\cdot S\mathbf{v} = \mathbf{0}$ because $\mathbf{v}\cdot S\mathbf{v} \overset{S \text{ any matrix}}{=} \mathbf{v}\cdot S^T\mathbf{v} \overset{S=-S^T}{=} -\mathbf{v}\cdot S\mathbf{v} \Rightarrow \mathbf{v}\cdot S\mathbf{v} = \mathbf{0}$. Second, the proof of question 2.1 in Series 2 guarantees that all eigenvalues of $Q \in SO(n)$ lie on the unit circle in the complex plane and $|Q| = -1 \Rightarrow \lambda = -1$ is an eigenvalue. Third note that if Q is a Cayley transform, then by 1.1.2, $(\mathrm{I}+Q)$ is invertible and so Q can not have -1 as an eigenvalue.

This in turn implies $Q \notin O(n) \setminus SO(n)$ so that if Q is a Cayley transform and orthogonal, then it must be that $Q \in SO(n)$. Then we have

$$\begin{aligned} Q^T Q = \mathrm{I} \quad &\Leftrightarrow \quad (\mathrm{I}-S)^{-T}(\mathrm{I}+S)^T(\mathrm{I}+S)(\mathrm{I}-S)^{-1} = \mathrm{I}, &(15)\\ &\Leftrightarrow \quad (\mathrm{I}+S)^T(\mathrm{I}+S) = (\mathrm{I}-S)^T(\mathrm{I}-S), &(16)\\ &\Leftrightarrow \quad \mathrm{I} + S^T + S + S^T S = \mathrm{I} - S^T - S + S^T S, &(17)\\ &\Leftrightarrow \quad S^T = -S. &(18) \end{aligned}$$

Thus any matrix Q that is a Cayley transform of another matrix S, is in $SO(n)$ if and only if S is skew, i.e. $S = S^T$. However this does not say that all matrices in $SO(n)$ are Cayley transform of skew matrices because $\lambda = -1$ is possible, in which case $Q + \mathrm{I}$ is not invertible, $Q + \mathrm{I}$ singular implies by 1.1.2 of this series, that Q is not a Cayley transform of any matrix. On the other hand if $Q \in SO(n)$ does not have $\lambda = -1$ as an eigenvalue, then $(\mathrm{I} + Q)$ is invertible, so that N can be defined via the inverse transform as in 1.1.3.

$$N = (Q + \mathrm{I})^{-1}(Q - \mathrm{I}),$$

any such N has $(\mathrm{I} - N)$ invertible (adapt proof of 1.1.2) so $Q = (\mathrm{I}+N)(\mathrm{I}-N)^{-1}$ and therefore $S = -S^T$ because $Q \in SO(n)$.

Thus there is a bijective mapping between $n \times n$ skew matrices $S \in \mathbb{R}^{n\times n}$ & special orthogonal matrices $Q \in SO(n)$ such that $Q + I$ is invertible, i.e. $\lambda = -1$ is not a eigenvalue. In the case $n = 3$ we know $\lambda = -1$ arises when Q is a rotation through π. Not yet obvious, but one can show that for general n, $(Q(S) + \mathrm{I})$ becomes singular as the norm of the matrices S tends to infinity.

1.2 The case of $SO(3)$

1. For any matrix N the scalar geometric series expansion

$$\frac{1}{1-a} = 1 + a + a^2 + a^3 + ..., \quad \text{converge for } |a| < 1$$

motivates the Neumann series for matrices

$$(\mathrm{I} - N)^{-1} = \mathrm{I} + N + N^2 + N^3 + ...,$$

which certainly converges if $||N||$ is small and may converge in other cases. If N is a $n \times n$ matrix it satisfies a degree n characteristic polynomial, so that only the powers $0, 1, ...(n-1)$ of N can be linearly independent. This observation motivates the ansatz (i.e. guess) that there may be scalar coefficient $a_0, ..., a_{n-1}$ such that

$$(\mathrm{I} - N)^{-1} = a_0\mathrm{I} + a_1 N + a_2 N^2 + ...a_{n-1}N^{n-1}$$

In general each a_i will be defined as a (scalar) infinite series, which may or may not converge. But we can avoid any study of convergence just by trying to solve the finite dimensional system of equations for $a_0, ..., a_{n-1}$ that come from the scalar coefficients of $\mathrm{I}, N, ..., N^{n-1}$ in the matrix equation

$$(\mathrm{I} - N)(a_0\mathrm{I} + ... + a_{n-1}N^{n-1}) = \mathrm{I}_n$$

where the term N^n is eliminated in favour of $\mathrm{I}, ..., N^{n-1}$ using the characteristic polynomial $\mathbb{P}$. This technique can be applied quite generally, but here we consider the case $n = 3$ & $N = S = -S^T = [\mathbf{u}\times]$ with characteristic polynomial

$$S^3 = -u^2 S \quad \text{(from Qn. 1.4 of series 2).} \qquad (19)$$

For now on we denote $u := ||\mathbf{u}||$.

Accordingly we make the following *ansatz*: $(\mathrm{I} - S)^{-1} = \alpha \mathrm{I} + \beta S + \gamma S^2$. Then we look for three numbers α, β and γ such that

$$
\begin{aligned}
& (\mathrm{I} - S)\,(\alpha\,\mathrm{I} + \beta\,S + \gamma\,S^2) = \mathrm{I}, && (20)\\
\Leftrightarrow\quad & (\alpha - 1)\,\mathrm{I} + (\beta - \alpha)\,S + (\gamma - \beta)\,S^2 - \gamma S^3 = 0, && (21)\\
\overset{(19)}{\Leftrightarrow}\quad & (\alpha - 1)\,\mathrm{I} + (\beta - \alpha + \gamma\,u^2)\,S + (\gamma - \beta)\,S^2 = 0, && (22)
\end{aligned}
$$

where we used the characteristic polynomial in the last equation to eliminate the S^3 term. The condition (22) is satisfied by

$$
\alpha = 1, \qquad \beta = \gamma = \frac{1}{1+u^2}, \tag{23}
$$

Finally, we can now compute explicitly the Cayley transform

$$
\begin{aligned}
Q &= (\mathrm{I} + S)\,(\mathrm{I} - S)^{-1}, && (24)\\
&= (\mathrm{I} + S)\,\frac{1}{1+u^2}\Big((1+u^2)\,\mathrm{I} + S + S^2\Big), && (25)\\
&= \frac{1}{1+u^2}\Big((1+u^2)\,\mathrm{I} + (2+u^2)\,S + 2\,S^2 + S^3\Big) && (26)\\
&\overset{(19)}{=} \frac{1}{1+u^2}\Big((1+u^2)\,\mathrm{I} + 2\,S + 2\,S^2\Big), && (27)\\
& && (28)
\end{aligned}
$$

Or, equivalently as in

$$
\overset{(session\ 2,ex\ 1.3)}{=} \frac{1}{1+u^2}\Big((1-u^2)\,\mathrm{I} + 2\,[\mathbf{u}\times] + 2\,\mathbf{u}\otimes\mathbf{u}\Big). \tag{29}
$$

As for the second part of the question, a Rodrigues formula of series 2 directly gives the matrix of a right-handed rotation of angle ϕ and unit axis $\mathbf{n}$:

$$
R(\phi, \mathbf{n}) = \cos\phi\,\mathrm{I} + \sin\phi\,[\mathbf{n}\times] + (1 - \cos\phi)\,\mathbf{n}\otimes\mathbf{n}. \tag{30}
$$

We first rewrite equation (29) as

$$
Q = \frac{1-u^2}{1+u^2}\,\mathrm{I} + \frac{2u}{1+u^2}\left[\frac{\mathbf{u}}{u}\times\right] + \frac{2u^2}{1+u^2}\,\frac{\mathbf{u}}{u}\otimes\frac{\mathbf{u}}{u}. \tag{31}
$$

In order to match the expressions for R and Q we need to express the angle ϕ in terms of u. For that we recall the double angle formulas for cosinus and sinus for $\cos(\phi/2) \neq 0$ e.g. $\phi \in (-\pi, \pi)$,

$$
\cos\phi = \frac{1 - \tan^2\frac{\phi}{2}}{1 + \tan^2\frac{\phi}{2}}, \quad \text{and} \quad \sin\phi = \frac{2\tan\frac{\phi}{2}}{1+\tan^2\frac{\phi}{2}}. \tag{32}
$$

Next we make the following *ansatz*: $\phi = 2\,\mathrm{Arctan}(u)$, which has a unique root in $[0, \pi)$ for any finite $u \geq 0$. To verify the latter *ansatz* we just need to observe that $u = \tan(\frac{\phi}{2})$ and that

$$
\cos\phi = \frac{1-u^2}{1+u^2}, \quad (1 - \cos\phi) = \frac{2u^2}{1+u^2}, \quad \text{and} \quad \sin\phi = \frac{2u}{1+u^2}. \tag{33}
$$

Finally we obtain that

$$
R(2\,\mathrm{Arctan}(u), \mathbf{u}/u) = \frac{1}{1+u^2}\left[(1-u^2)\,\mathrm{I} + 2\,u^2\,\frac{\mathbf{u}}{u}\otimes\frac{\mathbf{u}}{u} + 2\,u\,[\frac{\mathbf{u}}{u}\times]\right] \overset{(31)}{=} Q. \tag{34}
$$

2. We use the inverse transform in 1.1.3 to find S in terms of Q. To this intent, we first compute $(\mathrm{I}+Q)^{-1}$. Since $Q \in SO(3)$ has eigenvalues $1, e^{\pm i\theta}$ the characterstic polynomial can be computed from $(\lambda-1)(\lambda-e^{i\theta})(\lambda-e^{-i\theta})$. The Cayley-Hamilton theorem implies

$$Q^3 - t\,Q^2 + t\,Q - \mathrm{I} = 0, \tag{35}$$

where $t := \mathrm{Trace}(Q) = (1+2\cos(\theta))$. Multiplying (35) by Q^T yields

$$Q^2 = t\,Q - t\,\mathrm{I} + Q^T. \tag{36}$$

Next mimicking the approach of the last question we try to find three numbers α, β and γ such that

$$(\mathrm{I}+Q)(\alpha\,Q^T + \beta\,\mathrm{I} + \gamma\,Q) = \mathrm{I}. \tag{37}$$

(Note that $Q^T = Q^{-1}$, so we assume an ansatz in Q^{-1}, Q^0, Q^1, rather than Q^0, Q^1, Q^2 which turns out to be easier here). To this intent compute

$$\begin{aligned}
\text{Eq. (37)} \quad &\Leftrightarrow \quad \alpha\,Q^T + (\alpha+\beta-1)\,\mathrm{I} + (\gamma+\beta)\,Q + \gamma\,Q^2 = 0,\\
&\overset{(36)}{\Leftrightarrow} \quad (\alpha+\gamma)\,Q^T + (\alpha+\beta-1-t\,\gamma)\,\mathrm{I} + (\gamma\,(1+t)+\beta)\,Q = 0,
\end{aligned} \tag{38}$$

$$\Leftrightarrow \quad \begin{cases} \alpha+\gamma = 0, \\ \alpha+\beta-t\,\gamma = 1, \\ \beta+(1+t)\,\gamma = 0, \end{cases} \quad \Leftrightarrow \quad \begin{cases} \alpha = \dfrac{1}{2\,(1+t)}, \\ \beta = \dfrac{1+t}{2\,(1+t)}, \\ \gamma = \dfrac{-1}{2\,(1+t)}. \end{cases} \tag{39}$$

Accordingly, the final result in (39) provides the three coefficients that we were looking for so that

$$(\mathrm{I}+Q)^{-1} = \frac{1}{2\,(1+t)}\Big(Q^T + (1+t)\,\mathrm{I} - Q\Big). \tag{40}$$

Now, making use of the inverse transform 1.1.3 of this sheet, we compute

$$\begin{aligned}
S \quad &= \quad (\mathrm{I}+Q)^{-1}\,(Q-\mathrm{I}) = \frac{1}{2\,(1+t)}\Big(Q^T + (1+t)\,\mathrm{I} - Q\Big)\,(Q-\mathrm{I}), && (41)\\
&= \quad \frac{1}{2\,(1+t)}\Big(-Q^T - t\,\mathrm{I} + (2+t)\,Q - Q^2\Big), && (42)\\
&\overset{(36)}{=} \quad \frac{Q-Q^T}{1+t} && (43)
\end{aligned}$$

as required. We note that from Session 2 we already knew that $S = [\mathbf{u}\times]$ with $(Q-Q^T)\mathbf{u} = 0$, but equation (42) explicitly gives both the direction and scale $\frac{1}{1+t}$.

1.3 The case of $SE(3)$

First we note that if $\mathbb{R}^{(n+1)\times(n+1)} \ni X = \begin{pmatrix} M & \mathbf{m} \\ 0 & 1 \end{pmatrix}$ with $M \in \mathbb{R}^{n\times n}$ then $det(X) = det(M)$ and X is invertible if and only if M is invertible. Moreover

$$X^{-1} = \begin{pmatrix} M^{-1} & -M^{-1}\mathbf{m} \\ 0 & 1 \end{pmatrix} \tag{44}$$

Next, we prove that if $\mathcal{S}$ is of the form

$$\mathcal{S} = \begin{pmatrix} [\mathbf{u}\times] & \mathbf{v} \\ 0 & 0 \end{pmatrix}, \tag{45}$$

then its Cayley transform is in $SE(3)$. This is because

$$\begin{aligned}
(\mathcal{I}+\mathcal{S})(\mathcal{I}-\mathcal{S})^{-1} &= \begin{pmatrix} \mathrm{I}+[\mathbf{u}\times] & \mathbf{v} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{I}-[\mathbf{u}\times] & -\mathbf{v} \\ 0 & 1 \end{pmatrix}^{-1}, & (46)\\
&= \begin{pmatrix} \mathrm{I}+[\mathbf{u}\times] & \mathbf{v} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} (\mathrm{I}-[\mathbf{u}\times])^{-1} & (\mathrm{I}-[\mathbf{u}\times])^{-1}\mathbf{v} \\ 0 & 1 \end{pmatrix} & (47)\\
&= \begin{pmatrix} Q & Q\mathbf{v}+\mathbf{v} \\ 0 & 1 \end{pmatrix}, & (48)
\end{aligned}$$

where $Q := (\mathrm{I}+[\mathbf{u}\times])(\mathrm{I}-[\mathbf{u}\times])^{-1} \in SO(3)$ because it is the Cayley transform of a skew matrix. Accordingly, the RHS of (48) is in $SE(3)$.

Next, we show that if $\mathcal{Q}$ is the Cayley transform of some matrix $\mathcal{S} \in \mathbb{R}^{4\times 4}$ and $\mathcal{Q} \in SE(3)$, then $\mathcal{S}$ is of the form (45). If $\mathcal{Q} \in SE(3)$ then there exist a matrix $Q \in SO(3)$ and a vector $\mathbf{q}$ such that

$$\mathcal{Q} = \begin{pmatrix} Q & \mathbf{q} \\ 0 & 1 \end{pmatrix}. \tag{49}$$

Furthermore, because $\mathcal{Q}$ is a Cayley transform, $\mathcal{Q}+\mathcal{I}$ must be invertible. This in turn implies that $Q+\mathrm{I}$ is invertible. Then after defining $S = (Q+\mathrm{I})^{-1}(Q-\mathrm{I})$ a similar argument to that developed in 1.1.2 (make sure you can do it) shows that $\mathrm{I}-S$ is invertible and a similar argument to 1.1.3 shows that Q is the Cayley transform of S. Finally, 1.1.5 ensures that S must be skew so there exists a vector $\mathbf{u}$ such that $S = [\mathbf{u}\times]$.

The result 1.1.3 implies

$$\begin{aligned}
\mathcal{S} &= (\mathcal{Q}+\mathcal{I})^{-1}(\mathcal{Q}-\mathcal{I}) = \begin{pmatrix} Q+\mathrm{I} & \mathbf{q} \\ 0 & 2 \end{pmatrix}^{-1} \begin{pmatrix} Q-\mathrm{I} & \mathbf{q} \\ 0 & 0 \end{pmatrix}\\
&= \begin{pmatrix} (Q+\mathrm{I})^{-1} & -\frac{1}{2}(Q+\mathrm{I})^{-1}\mathbf{q} \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} Q-\mathrm{I} & \mathbf{q} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} [\mathbf{u}\times] & (Q+\mathrm{I})^{-1}\mathbf{q} \\ 0 & 0 \end{pmatrix}.
\end{aligned}$$

Notice that all the hard work was in the $SO(3)$ case addressed in question 1.2. Using the $SO(3)$ results and equation (44) means there is little remaining work for the $SE(3)$ case.