

Lecture Note: Quaternions, Exponential Map, and Quaternion Jacobians

Marc Toussaint

Learning & Intelligent Systems Lab, TU Berlin

March 1, 2024

SO(2) is the space of rotations in the plane \mathbb{R}^2 (or about a single axis), which can be described by an angle $\alpha \in [0, 2\pi]$. But there are good arguments to instead represent such rotations by a point on S^1 , which is the circle in \mathbb{R}^2 . E.g., topologically SO(2) is equal to S^1 and not to the interval $[0, 2\pi]$. Points in S^1 can also be described as $e^{i\alpha} \in \mathbb{C}$, which is the circle in the complex plane. Quaternions similarly represent rotations of SO(3) as points on S^3 , i.e., the sphere in \mathbb{R}^4 . And just as $e^{i\alpha} \in \mathbb{C}$ gives points on the circle S^1 , the exponential map of SO(3) gives defines quaternions on S^3 .

The following will first provide reference equations for quaternions, but then explain in more detail interpolation between rotations, which naturally introduces the exponential map directly for quaternions without need to go deeper into Lie groups. An emphasis is then on deriving the *quaternion Jacobian*, which describes the angular velocity that is induced when perturbing a quaternion parameter.

1 Reference

Let's define a quaternion $q \in \mathbb{R}^4$ as a 4-vector.¹ To represent a rotation, a quaternion needs to be normalized, i.e., $q \in S^3$. We denote the first entry by q_0 and the last three entries by $\bar{q} \equiv q_{1:3} \in \mathbb{R}^3$. Let's first summarize some basic properties for a normalized quaternion:

$$\text{rotation angle } \theta \in \mathbb{R}: \quad \theta = 2 \arccos(q_0) = 2 \arcsin(|\bar{q}|) \quad (1)$$

$$\text{normalized axis } \hat{w} \in S^2: \quad \hat{w} = \frac{\bar{q}}{|\bar{q}|} = \frac{1}{\sin(\theta/2)} \bar{q} \quad (2)$$

$$\text{rotation vector } w \in \mathbb{R}^3: \quad w = \log(q) = 2 \arccos(q_0) \frac{\bar{q}}{|\bar{q}|} \quad (3)$$

$$\text{quat. from vector } w = \theta \hat{w}: \quad q = \exp(w) = (\cos(\theta/2), \sin(\theta/2) \hat{w}) \quad (4)$$

$$\text{rotation matrix } R \in \mathbb{R}^{3 \times 3}: \quad R = \begin{pmatrix} 1 - q_{22} - q_{33} & q_{12} - q_{03} & q_{13} + q_{02} \\ q_{12} + q_{03} & 1 - q_{11} - q_{33} & q_{23} - q_{01} \\ q_{13} - q_{02} & q_{23} + q_{01} & 1 - q_{11} - q_{22} \end{pmatrix}, \quad q_{ij} := 2q_i q_j \quad (5)$$

$$\text{quaternion from matrix: } q_0 = \frac{1}{2} \sqrt{1 + \text{tr}R}, \quad q_1 = \frac{R_{32} - R_{23}}{4q_0}, \quad q_2 = \frac{R_{13} - R_{31}}{4q_0}, \quad q_3 = \frac{R_{21} - R_{12}}{4q_0} \quad (6)$$

$$\text{quaternion inverse: } q^{-1} = (q_0, -\bar{q}) \quad \text{or} \quad q^{-1} = (-q_0, \bar{q}) \quad (7)$$

$$\text{quaternion concatenation: } q \circ q' = (q_0 q'_0 - \bar{q}^\top \bar{q}', \quad q_0 \bar{q}' + q'_0 \bar{q} + \bar{q} \times \bar{q}') \quad (8)$$

$$\text{vector application, } x \in \mathbb{R}^3: \quad q \cdot x = (q \circ (0, x) \circ q^{-1})_{1:3} \quad (\text{or via matrix}) \quad (9)$$

Eqs. (3) and (4) follow from (1) and (2), but why they are defined as 'log' and 'exp' will become clear below. Concerning the application of a rotation quaternion q on a vector x , equation (9) is elegant,

¹In the original meaning, they are defined as $q = q_0 + q_1 i + q_2 j + q_3 k$, with j, k being extensions to the imaginary number i . But we just treat them as 4-vectors.

but computationally less efficient than first converting to matrix using (5) and then multiplying to the vector. Note that concatenating quaternions as well as applying a quaternion to a vector never requires to compute a sin or cos, as opposed, e.g. to the Rodriguez' equation (see below).

2 Continuously moving from \mathbf{I} to q – exponential and log mappings

Given a quaternion q , how can we continuously move from the $\mathbf{I} = (1, \bar{0})$ to q ? The answer is simply given by continuously increasing the rotation angle (1) from 0 to θ about a fixed normalized axis \hat{w} . Let $t \in [0, \theta]$ be the interpolation coefficient, we have

$$q(t) = (\cos(t/2), \sin(t/2) \hat{w}) . \quad (10)$$

This “motion” from \mathbf{I} to q is a geodesic (=shortest path) on S^3 , and it has constant absolute velocity:

$$\dot{q}(t) = (-\sin(t/2), \cos(t/2) \hat{w}) , \quad |\dot{q}(t)| = \frac{1}{2} . \quad (11)$$

At the origin (for $t = 0$), the velocity is $\dot{q}(0) = \frac{1}{2}(0, \hat{w})$, which shows how the rotation axis \hat{w} provides the tangent vector at \mathbf{I} . For $t > 0$ the velocity (11) “rotates along the tangent arc” on S^3 . We can show that $\dot{q}(t)$ can also be written as (using that \hat{w} and $\bar{q}(t)$ are colinear):

$$q(t) \circ \frac{1}{2}(0, \hat{w}) = \frac{1}{2}(0q_0(t) - \hat{w}^\top \bar{q}(t), 0\bar{q}(t) + q_0(t)\hat{w} + \hat{w} \times \bar{q}(t)) \quad (12)$$

$$= \frac{1}{2}(0 - \sin(t/2), \cos(t/2) \hat{w} + 0) = \dot{q}(t) . \quad (13)$$

The relation $\dot{q}(t) = q(t) \circ \frac{1}{2}(0, \hat{w})$ nicely makes explicit how the tangent vector $Frac12(0, \hat{w})$ at \mathbf{I} is being transported along the arc by multiplying it with $q(t)$ from the left.

Eq. (13) is a differential equation of the form $\dot{q}(t) = q(t)\alpha$, but with the multiplication being the group operation \circ . The solution $q(t)$ to this differential equation is *defined* as the exponential map.² Namely, for a vector $w = t\hat{w} \in \mathbb{R}^3$ we define

$$\exp(t\hat{w}) = (\cos(t/2), \sin(t/2) \hat{w}) , \quad \Rightarrow \quad \frac{\partial}{\partial t} \exp(t\hat{w}) = \exp(t\hat{w}) \circ \frac{1}{2}(0, \hat{w}) , \quad (14)$$

fully consistent to (10) and (13). Conversely, given a quaternion, we define the log as

$$\log(q) = 2 \operatorname{acos}(q_0) \frac{\bar{q}}{|\bar{q}|} . \quad (15)$$

3 Continuously moving from q_A to q_B – interpolation in quaternion space

Consider two quaternions q_A, q_B . When concatenating $q_B = q_A \circ q_{AB}$, you should think of q_{AB} as describing an “additional” rotation in the output coordinates of q_A . This is exactly as multiplying matrices $R_B = R_A R_{AB}$, where R_{AB} describes a rotation from orientation R_A to orientation R_B . The relative rotation $q_{AB} = q_A^{-1} \circ q_B$ retrieves exactly this relative rotation from q_A to q_B .

There are two ways to interpolate between orientations q_A and q_B . Let $t \in [0, 1]$. The *proper* interpolation uses the exponential map:

$$q(t) = q_A \circ \exp(t \log(q_A^{-1} \circ q_B)) . \quad (16)$$

²When describing $\text{SO}(3)$ as a Lie groups, analogous equations appear when tangent motion is integrated to continuously move to group elements. The solution is defined as $\exp(\cdot)$ of a tangent vector.

This first computes the relative transform $q_{AB} = q_A^{-1} \circ q_B$, computes its tangent vector representation $w_{AB} = \log(q_{AB})$, and interpolates from \mathbf{I} to q_{AB} as we described above. Finally it multiplies q_A from the left, making all this relative to orientation q_A . For completeness, the velocity is

$$\dot{q}(t) = q_A \circ \left[\frac{\partial}{\partial t} \exp(t \log(q_A^{-1} \circ q_B)) \right] \quad (17)$$

$$= q_A \circ \left[\exp(t w_{AB}) \circ \frac{1}{2}(0, w_{AB}) \right] \quad (18)$$

$$= q(t) \circ \frac{1}{2}(0, w_{AB}) , \quad |\dot{q}(t)| = \frac{1}{2}|w_{AB}| . \quad (19)$$

Again, notice how the tangent vector $\frac{1}{2}(0, w_{AB})$ at \mathbf{I} is being transported to become the local tangent vector (=velocity) by multiplying it with $q(t)$ from the left.

As an alternative, we can define the *embedded* interpolation as

$$q(t) = \text{normalize}((1-t)q_A + tq_B) , \quad (20)$$

where $\text{normalize}(q) = q/|q|$. This assumes that the two quaternions are *sign aligned*, i.e. $q_A^\top q_B \geq 0$. This interpolation seems native. However, it describes the exact same *path* as the proper interpolation: along the geodesic arc on S^3 from q_A to q_B . The only difference is that it does not have constant absolute velocity on S^3 . The non-normalized interpolation $(1-t)q_A + tq_B$ in \mathbb{R}^4 has constant velocity, but the projection to S^3 modulates velocities depending on the angle of the tangent space to the embedded interpolation. But note that, as $q_A^\top q_B \geq 0$, the angle between q_A and q_B in \mathbb{R}^4 is at most 90 degrees; and therefore the angle between the linear interpolation and the tangent spaces is at most 45 degrees – in this worst case, the velocities at start and end are by a factor $1/\sqrt{2}$ larger than in the middle.

The point here is that interpolation of rotations is very intuitive: “straight” interpolation on the sphere S^3 which can equally be described using (20) (with the exponential map) or the naive embedded interpolation (20). But only the proper interpolation with the exponential map has constant absolute velocity.

4 Angular Jacobian w.r.t. Quaternion Parameters

Assume we have a quaternion $q \in \mathbb{R}^4$ as a *parameter* of some optimization problem and need a Jacobian w.r.t. q . More concretely, we want to know the angular velocity $w = J\dot{q}$ that is induced by a velocity (or infinitesimal variation) of q . To relate to Eq. (19) we think of $q = q_A$ as the base orientation, and \dot{q} as moving from q_A to q_B . Eq. (19) for $t = 0$ becomes

$$\dot{q} = q \circ \frac{1}{2}(0, w) . \quad (21)$$

This equation relates \dot{q} to an angular vector $w \in \mathbb{R}^3$, where $\frac{1}{2}(0, w)$ is a tangent vector at \mathbf{I} that is transported to become a tangent vector \dot{q} at q . Conversely, if \dot{q} is given and tangential, then we can read out w by multiplying q^{-1} from the left,

$$q^{-1} \circ \dot{q} = \frac{1}{2}(0, w) , \quad w = 2 [q^{-1} \circ \dot{q}]_{1:3} . \quad (22)$$

However, in the case where \dot{q} is non-tangential, i.e., $\dot{q}^\top q \neq 0$, the change in length of the quaternion does not represent any angular velocity. One approach is to first tangentialize $\hat{\dot{q}} = \dot{q} - q q^\top \dot{q}$. However, when plugging in the tangentialized $\hat{\dot{q}}$ we get

$$w = 2 [q^{-1} \circ (\dot{q} - q(q^\top \dot{q}))]_{1:3} \quad (23)$$

$$= 2 [\dot{q} \circ q^{-1}]_{1:3} - 2 [q^{-1} \circ q]_{1:3} (q^\top \dot{q}) , \quad (24)$$

and the latter term is identically zero. So, our “tangentialization term” drops out anyway.

To construct the Jacobian, let’s assume $\dot{q} = e_i$ is one of the unit vectors $e_0, \dots, e_3 \in \mathbb{R}^4$. Then we have

$$w_i = 2 [q^{-1} \circ e_i]_{1:3} . \quad (25)$$

Further, if q itself was not normalized we want the angular Jacobian of $\tilde{q} = \frac{1}{|q|}q$ w.r.t. q . As the above relation between q and w is linear, we have

$$w_i = \frac{2}{|q|} [\tilde{q}^{-1} \circ e_i]_{1:3} . \quad (26)$$

Based on this we construct the **angular quaternion Jacobian** $J(q) \in \mathbb{R}^{3 \times 4}$ with the 4 columns

$$J_{\cdot i}(q) = \frac{2}{|q|} [\tilde{q}^{-1} \circ e_i]_{1:3} , \quad (27)$$

so that we have $w = J(q) \dot{q}$ for any (also non-normalized) q and any (also non-tangential) \dot{q} .

Again, note that this angular vector w is to be interpreted *relative* to q , in the output space of q . To translate this to a world coordinate angular Jacobian, we need to transform by q again. To give a more complete example: if we have a pose $q_B = q_A \circ \text{normalize}(q)$ that is parameterized by arbitrary (non-zero) $q \in \mathbb{R}^4$, then \dot{q} translates to a world coordinate angular velocity $w = R_B J(q) \dot{q}$, with R_B the rotation matrix of pose B .

5 Random rotations & “Gaussians”

To get a random rotation, we want to uniformly sample from S^3 . This is particularly easy by first sampling $q \sim \mathcal{N}(0, 1)$ in \mathbb{R}^4 and then normalize.

To sample a rotation “Gaussian around a mean” quaternion q , we have two options (akin to the interpolation options): The tangent space sampling:

$$q' = \exp(w) \circ q , \quad w = \mathcal{N}(0, \sigma^2) \in \mathbb{R}^3 , \quad (28)$$

or

$$q' = \text{normalize}(q + \delta/2) , \quad \delta = \mathcal{N}(0, \sigma^2) \in \mathbb{R}^4 . \quad (29)$$

In the limit of small σ^2 these become the same, as the Gaussian projected to the tangent space is the same as the Gaussian in the tangent space.

A Similarly: Rotation vector and angular velocity

To recap, a rotation vector is $w \in \mathbb{R}^3$, with rotation axis $\hat{w} = \frac{w}{|w|}$, and rotation angle $\theta = |w|$ (in radians, using the right thumb convention).

The application of a rotation described by $w \in \mathbb{R}^3$ on a vector $x \in \mathbb{R}^3$ is given as (Rodrigues’ formula)

$$w \cdot x = \cos(\theta) x + \sin(\theta) (\hat{w} \times x) + (1 - \cos(\theta)) \hat{w}(\hat{w}^\top x) . \quad (30)$$

This directly also implies conversion to a rotation matrix, as the rows of the corresponding rotation matrix are simply $w \cdot e_i$ for the unit vectors $e_{1:3} \in \mathbb{R}^3$. To simplify the notation, we define the skew matrix of a vector $w \in \mathbb{R}^3$ as

$$\text{skew}(w) = \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix} . \quad (31)$$

This allows to express the cross product as matrix multiplication, $w \times v = \text{skew}(w) v$. The rotation matrix $R(w)$ corresponding to a given rotation vector w then is:

$$R(w) = \exp(\text{skew}(w)) \quad (32)$$

$$= \cos(\theta) \mathbf{I} + \sin(\theta)/\theta \operatorname{skew}(w) + (1 - \cos(\theta))/\theta^2 w w^\top \quad (33)$$

The exp function is called exponential map (generating a group element (=rotation matrix) via an element of the Lie algebra (=skew matrix)). The other formula is called Rodrigues' formula: the first term is a diagonal matrix (\mathbf{I} is the 3D identity matrix), the second term is the skew symmetric part, the last term is the symmetric part ($w w^\top$ is also called outer product).

Angular velocity & derivative of a rotation matrix: We represent angular velocities by a vector $w \in \mathbb{R}^3$, with rotation axis \hat{w} , and rotation velocity $|w|$ (in radians per second). When a body's orientation at time t is described by a rotation matrix $R(t)$ and the body's angular velocity is w , then

$$\dot{R}(t) = \operatorname{skew}(w) R(t) . \quad (34)$$

(That's intuitive to see for a rotation about the x -axis with velocity 1.) Some insights from this relation: Since $R(t)$ must always be a rotation matrix (fulfill orthogonality and determinant 1), its derivative $\dot{R}(t)$ must also fulfill certain constraints; in particular it can only live in a 3-dimensional sub-space. It turns out that the derivative \dot{R} of a rotation matrix R must always be a skew symmetric matrix $\operatorname{skew}(w)$ times R – anything else would be inconsistent with the constraints of orthogonality and determinant 1.

Note also that, assuming $R(0) = I$, the solution to the differential equation $\dot{R}(t) = \operatorname{skew}(w) R(t)$ can be written as $R(t) = \exp(t \operatorname{skew}(w))$, where here the exponential function notation is used to denote a more general so-called exponential map, as used in the context of Lie groups. It also follows that $R(w)$ from (32) is the rotation matrix you get when you rotate for 1 second with angular velocity described by w .