Last Time

- DDP vs. DIRCOL
- Quaternions

This Time:

- Optimization with Quaternions

---

* Quaternion Recap

- 4D unit vectors
- Multiplication rule:

$$q_1 * q_2 = \begin{bmatrix} s_1 \\ v_1 \end{bmatrix} * \begin{bmatrix} s_2 \\ v_2 \end{bmatrix} = \begin{bmatrix} s_1 s_2 - v_1^T v_2 \\ s_1 v_2 + s_2 v_1 + v_1 \times v_2 \end{bmatrix}$$

$$L(q_1) = \left[\begin{array}{c|c} s_1 & v_1^T \\ \hline v_1 & s_1 I + \hat{v}_1 \end{array}\right] \Rightarrow q_1 * q_2 = L(q_1) q_2$$

- Quaternion Conjugate

$$q^+ = \begin{bmatrix} s \\ -v \end{bmatrix} = Tq \quad , \quad T = \begin{bmatrix} 1 & 0 \\ 0 & -I \end{bmatrix}$$

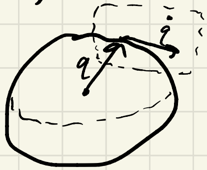
- Identity:

$$q_I = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- "Hat map" for Quaternions

$$\hat{\omega} = \begin{bmatrix} 0 \\ \omega \end{bmatrix} = H\omega \quad , \quad H = \begin{bmatrix} 0 \\ I \end{bmatrix}$$

* Geometry of Quaternions



q lives on a sphere in $\mathbb{R}^4$

$\dot{q}$ lives in the tangent plane to the sphere at q

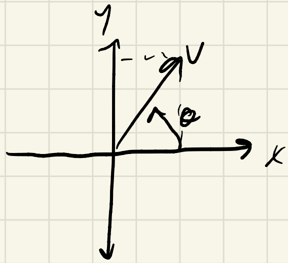
- Kinematics

$\dot{q} \in \mathbb{R}^4$, $\omega \in \mathbb{R}^3$, $\dot{q} = \frac{1}{2} q * \begin{bmatrix} 0 \\ \omega \end{bmatrix} = \frac{1}{2} L(q) H \omega$

$q * \begin{bmatrix} 0 \\ \omega \end{bmatrix}$

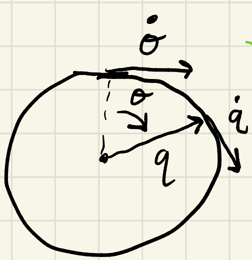
ω is always written down in the tangent plane at the identity then Kinematics rotate to the tangent plane at q

Analogy with unit complex numbers in 2D:



$$q = \underbrace{\cos(\theta)}_{\text{x-component}} + \underbrace{i\sin(\theta)}_{\text{y-component}} \Rightarrow q = \underbrace{\begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}}_{q^T q = 1}$$

$$\dot{q} = \frac{\partial q}{\partial \theta}\dot{\theta} = \begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix} \dot{\theta} = \underbrace{\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}}_{\text{rotation matrix}} \underbrace{\begin{bmatrix} 0 \\ \dot{\theta} \end{bmatrix}}_{\text{2D version of hat map}}$$



Kinematics rotates $\dot{\theta}$ from tangent plane at "north pole" to tangent plane at q

* Differentiating Quaternions

- Two key facts:

  1) Derivatives are really 3D "tangent vectors"
  2) Rotations compose by multiplication, not addition

- Infinitesimal Rotation

$$\delta q = \begin{bmatrix} \cos(\theta/2) \\ a\sin(\theta/2) \end{bmatrix} \approx \begin{bmatrix} 1 \\ \frac{1}{2}a\theta \end{bmatrix} \approx \begin{bmatrix} 1 \\ \frac{1}{2}\underbrace{\phi}_{\text{small axis-angle vector}} \end{bmatrix} \approx \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 \\ \phi \end{bmatrix}$$

- Compose with q:

$$q' = q\,\delta q = L(q)\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 \\ \phi \end{bmatrix}\right) = q + \underbrace{\tfrac{1}{2}L(q)H}_{G(q)\,\in\,\mathbb{R}^{4\times 3}}\phi$$

"attitude jacobian"

- Note: we can use any 3-parameter rotation representation for ϕ. They all linearize the same up to a constant:

$$q = \begin{bmatrix} \cos(\|\phi\|/2) \\ \frac{\phi}{\|\phi\|}\sin(\|\phi\|/2) \end{bmatrix}, \quad q = \begin{bmatrix} \sqrt{1-\phi^T\phi} \\ \phi \end{bmatrix}, \quad \frac{1}{\sqrt{1+\phi^T\phi}}\begin{bmatrix} 1 \\ \phi \end{bmatrix}$$

axis-angle — vector part of q — Gibbs/Rodrigues Vector

- We'll use the vector part of q since it is simplest/cheapest.

- This lets us differentiate any quaternion function by simply inserting $G(q)$ in the right places:

$f(q) : \mathbb{H} \to \mathbb{R}$ (gradient of scalar-valued function)

$\mathbb{H}$: quaternions ("Hamilton")

$$\nabla f = \frac{\partial f}{\partial q} \frac{\partial q}{\partial \phi} = \frac{\partial f}{\partial q} G(q)$$

$f(q) : \mathbb{H} \to \mathbb{H}$ (jacobian of quaternion-valued function)

$$\phi' = \underbrace{\left[G(f(q))^T \frac{\partial f}{\partial q} G(q) \right]}_{\nabla f \in \mathbb{R}^{3\times 3}} \phi \qquad q' = f(q)$$

$$q' * \begin{bmatrix} \sqrt{1-\phi'^T\phi'} \\ \phi' \end{bmatrix} = f\left(q * \begin{bmatrix} \sqrt{1-\phi^T\phi} \\ \phi \end{bmatrix}\right)$$

transform output $\delta q' \to \phi'$

transform input $\phi \to \delta q$

- Hessian of $f(q) : \mathbb{H} \to \mathbb{R}$

$$\nabla^2 f(q) = G(q)^T \frac{\partial^2 f}{\partial q^2} G(q) + \underbrace{\overset{3\times 3}{I} \overset{\text{scalar}}{\left(\frac{\partial f}{\partial q} q \right)}}_{\text{comes from } \frac{\partial G}{\partial q}}$$

- Now we can do Newton's method with quaternions (and DDP and SQP)

* Example: Pose Estimation

- Given a bunch of vectors to known features in the environment, determine the robot's attitude

- Called "Wahba's Problem"

$$\underset{q}{\operatorname{argmin}} \quad J(q) = \sum_{n}^{m} \| \underbrace{{}^{N}x_n}_{\substack{\text{known vectors} \\ \text{in inertial frame} \\ \text{(e.g. from a map)}}} - Q(q)\, \underbrace{{}^{B}x_n}_{\substack{\text{measured vectors} \\ \text{in body frame} \\ \text{(e.g. from camera)}}} \|_2^2 = \| \overbrace{r(q)}^{\text{"residual"}} \|_2^2$$

$$r(q) = \underbrace{\begin{bmatrix} {}^{N}x_1 - Q\,{}^{B}x_1 \\ {}^{N}x_2 - Q\,{}^{B}x_2 \\ \vdots \end{bmatrix}}_{3m \times 1} \Rightarrow \underbrace{\nabla r(q)}_{3m \times 3} = \underbrace{\frac{\partial r}{\partial q}}_{3m \times 4} \underbrace{G(q)}_{4 \times 3}$$

- Gauss-Newton Method:

do:

$$\nabla r(q) = \frac{\partial r}{\partial q} G(q)$$

$$\phi = -\left[(\nabla r^T \nabla r)^{-1} \nabla r^T \right] r(q)$$

$$q \leftarrow q * \begin{bmatrix} \sqrt{1 - \phi^T \phi} \\ \phi \end{bmatrix}$$

← multiplicative update

In general, perform line search here

while $\| \phi \| > tol$

* Background: Gauss-Newton for least-squares problems

$$\min_{x} J(x) = \frac{1}{2}\|r(x)\|_2^2 = \frac{1}{2} r(x)^T r(x)$$

$$\frac{\partial J}{\partial x} = \underbrace{r(x)^T \frac{\partial r}{\partial x}}.$$

throw this term out

$$\frac{\partial J}{\partial x^2} = \left(\frac{\partial r}{\partial x}\right)^T \left(\frac{\partial r}{\partial x}\right) + \cancel{\left(I \otimes r(x)^T\right)} \frac{\partial^2 vec(r)}{\partial x^2}$$

$$\Rightarrow \left(\frac{\partial J}{\partial x^2}\right)^{-1} \nabla J \approx \left[\left(\frac{\partial r}{\partial x}\right)^T \left(\frac{\partial r}{\partial x}\right)\right]^{-1} \frac{\partial r}{\partial x}^T r(x)$$