Continuous representations in neural networks:

- The mapping $g$ in the below diagram needs to be continuous.
- In other words, $g$ is an embedding ($X$ is homeomorphic to a subset of $\mathbb{R}$).
- Discontinuous representations are hard for neural networks to approximate.

Commonly used rotation representations are discontinuous.

- A 2D example:
  Polar angle representation

Note: For 3D rotations, all representations are discontinuous in the real Euclidean spaces of four or fewer dimensions.

Let's find continuous 3D rotation representations.

6D representation for 3D rotations

Mapping from $\text{SO}(3)$ (matrix) to a 6D representation (Stiefel Manifold):

$$g_{\text{GS}} \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \end{bmatrix} = \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix}$$

Mapping from 6D representation to $\text{SO}(3)$ (Gram-Schmidt-like orthogonalization):

$$f_{\text{GS}} \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

5D representation for 3D rotations

Eliminating one dimension of $g_{\text{GS}}$ by stereographically projecting $[a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}]^T$ to 3D.

$$v = \frac{v_i}{||v||}$$

An illustration of stereographic projection in 2D. We are given as input a point $p$ on the unit sphere $S^1$. We construct a ray from a fixed projection point $N_0 = (0, 1)$ through $p$ and find the intersection of this ray with the plane $y = 0$. The resulting point $p'$ is the stereographic projection of $p$.

- For the general case of the n dimensional rotation group $\text{SO}(n)$, please check our paper. We also discuss other groups such as the orthogonal group and similarity transforms.

Sanity Test:

Inverse Kinematics Test:

Worst two frames using quaternions.

Worst two frames using 6D representations.

IK results for the two frames with highest pose error from the test set for the network trained using quaternions, and the corresponding results on the same frames for the network trained on the 6D representation.