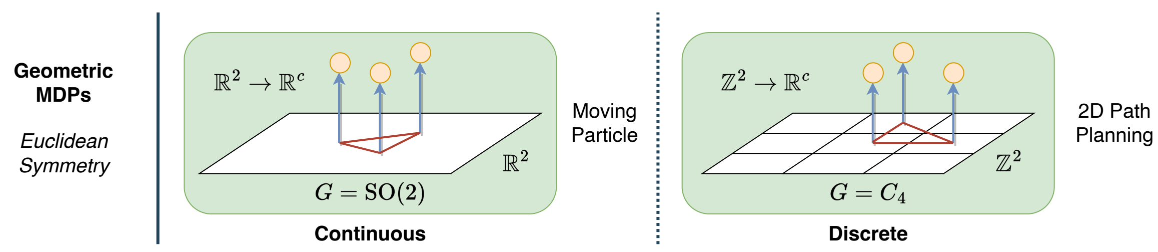
$E(3)$ *equivariance* can improve efficiency of *sampling-based* RL and planning.

Can Euclidean Symmetry be Leveraged in Reinforcement Learning and Planning?

Northeastern Khoury College of Computer and Information Sciences

Linfeng Zhao, Owen Howell, Jung Yeon Park, Xupeng Zhu, Robin Walters, and Lawson L.S. Wong
Khoury College of Computer Sciences, Northeastern University

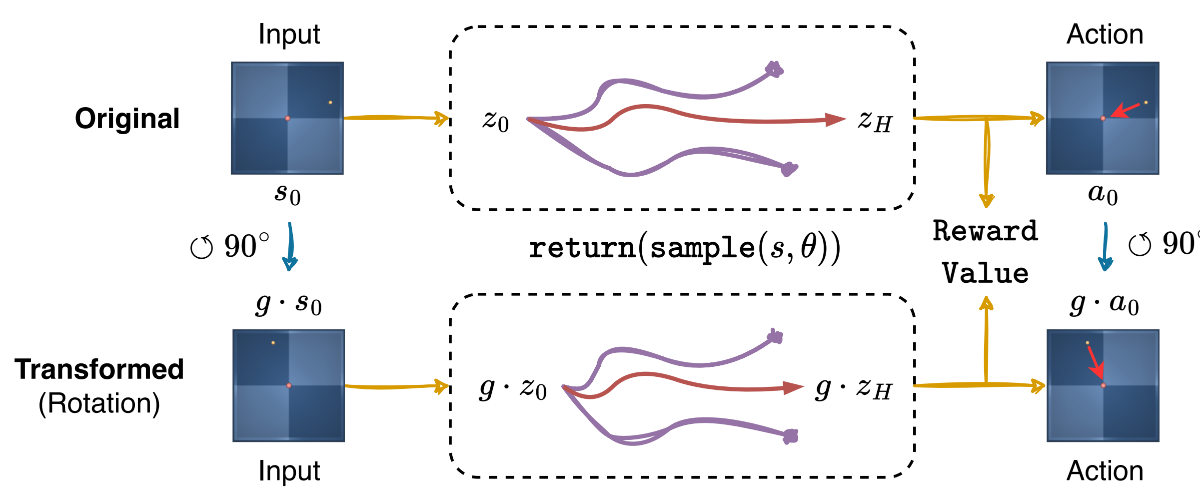


Overview & Geometric Properties of MDPs

- In robotic tasks, changes in reference frames typically do not influence the underlying physical properties of the system, known as invariance of physical laws.
- These transformations form Euclidean group. We identify such class of MDPs as "Geometric MDP".

Method: Equivariance in Sampling-based Planning

- The work generalizes previous work on symmetry in path planning on 2D grid [Zhao et al. ICLR'23] to continuous action space and symmetry group, necessitating sampling-based planning and RL.
- We identify the conditions to achieve equivariance in sampling-based planning: (1) invariant **return** function, and (2) the action samples $\mathbb{A}$ is closed under group G.
- The figure demonstrates equivariance in the procedure.



Original: $s_{t+1} = f(s_t, a_t)$ $\to$ Linearized at step t: $s_{t+1} = A_t \cdot s_t + B_t \cdot a_t$

$$s_{t+1} = A(p) \cdot s_t + B(p) \cdot a_t, \quad A : \mathcal{S} \times \mathcal{A} \to \mathbb{R}^{d_\mathcal{S} \times d_\mathcal{S}}, \quad B : \mathcal{S} \times \mathcal{A} \to \mathbb{R}^{d_\mathcal{S} \times d_\mathcal{A}}.$$

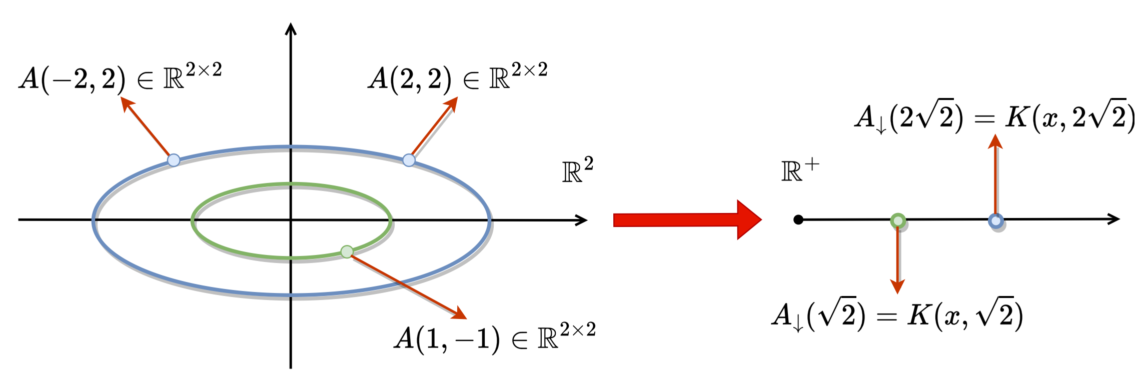
$$\forall g \in G, \quad A(g \cdot p) = \rho_\mathcal{S}(g) A(p) \rho_\mathcal{S}(g^{-1}), \quad B(g \cdot p) = \rho_\mathcal{S}(g) B(p) \rho_\mathcal{A}(g^{-1})$$

Theory: Linearization and Steerable Constraints

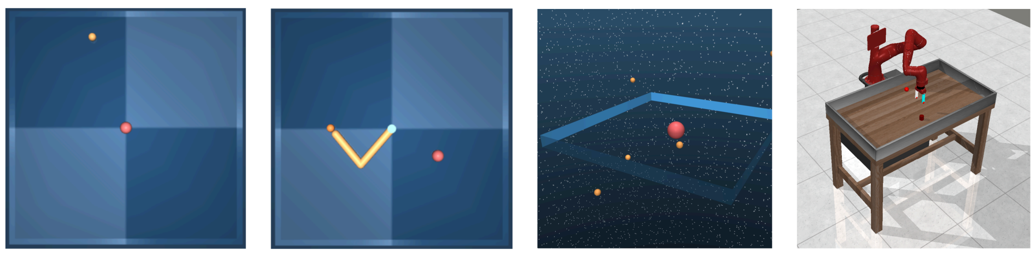
- For Geometric MDPs (with continuous group action), linearizing the dynamics and the group action results in a linear state-space model but with parameterized kernels.
- The kernels satisfy G-steerable kernel constraints.

Illustration of Steerable Kernel Constraints

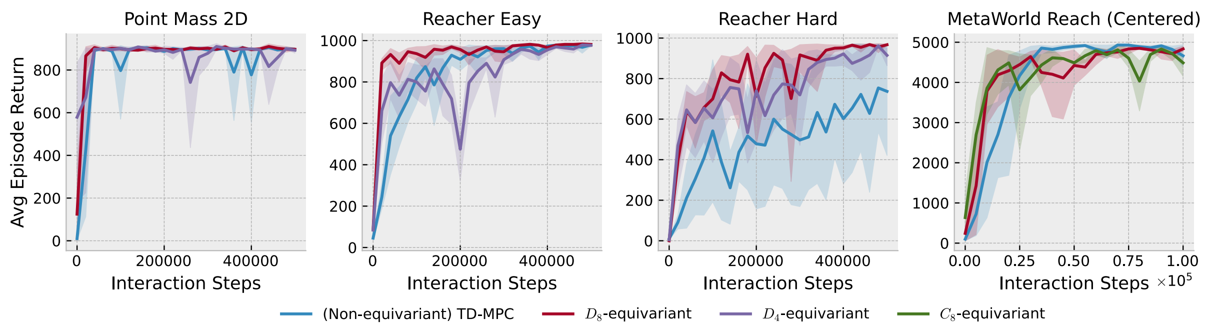
- The illustrative examples show how dimensionality of of the space of the matrices can be reduced.
- This demonstrates how a matrix-valued kernel $A : X \to \mathbb{R}^{2\times 2}$ is constrained by the SO(2)-steerable kernel constraints on a set of orbits $A(g \cdot p) = \rho_{\text{out}}(g) A(p) \rho_{\text{in}}(g^{-1})$.

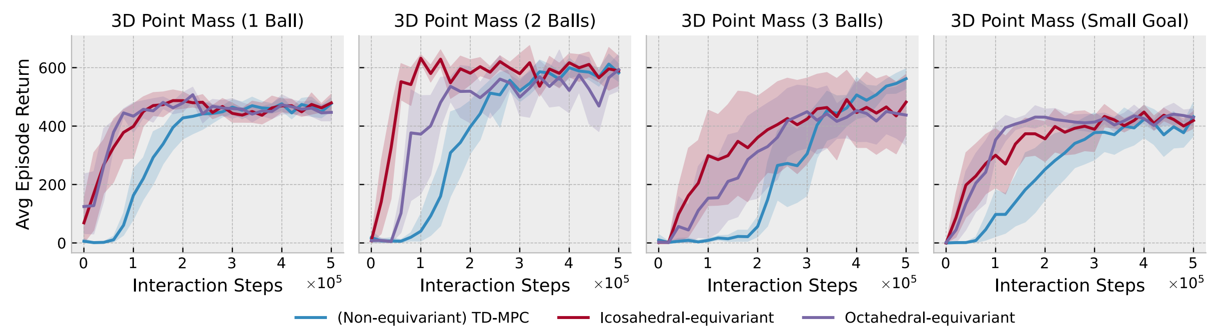


Empirical Evaluation



- We propose an equivariant model-based RL algorithm based on TD-MPC. We show that which components need to be equivariant.
- We run it on several tasks to demonstrate better sample efficiency.





Project and Paper Webpage: