Proceedings of Machine Learning Research vol vvv:1–19, 2024

Dynamics abstract harmonic analysis

Daniel Felipe Ordoñez-Apraez DANIEL.ORDONEZ@IIT.IT
Vladimir Kostic TODO@SAMPLE.COM
Giulio Turrisi GIULIO.TURRISI@IIT.IT
Pietro Novelli LISAN.AL.GAIB@DUNE.COM
Carlos Mastalli C.MASTALLI@HW.AC.UK
Massimiliano Pontil AN3@SAMPLE.COM
Claudio Semini CLAUDIO.SEMINI@IIT.IT
Computational Statistics and Machine Learning - Istituto Italiano di Tecnologia
Dynamic Legged Systems - Istituto Italiano di Tecnologia
Robot Motor Intelligence – Heriot-Watt University

Abstract

TODO: Bla, bla..

Keywords: Symmetric dynamical systems, harmonic analysis, Koopman operator

TODO:

- Motivation: Modeling and control of symmetric dynamical systems, exploiting/capturing both state and temporal symmetries of the dynamics.
- Challenges: Sample-efficiency, generalization of the model, interpretability, numerical efficiency.
- Trends: Koopman/Transfer operator theory, Neural Operators, Neural-ODEs, Model-based analytic methods (e.g., rigid-body dynamics).
- Methodology: Exploit the group representation of finite/infinite-dimensional state vector spaces, to decompose the space into a sum of orthogonal (orthogonal observable functions, not to be conflated with independent dynamics, which are only true in the infinite-dimensional case). Resulting in a new basis set of the state space, that poses several numerical benefits, including block-diagonalization of the Koopman operator (and state-dependent/instantaneous linearization of non-linear dynamics), reduced dimensionality of the subspaces, and reduced number of symmetry constraints. Plus the fact that on top of exploiting equivariance (which many people do these days) we exploit them gaining interpretability.
- Contributions:
  - Formalization of a method for learning models exploiting both state and temporal symmetries.
  - Gained interpretability of dynamic mode decomposition algorithms, by know doing DMD on each harmonics state-subspace.
  - Introduction of harmonic analysis of the state space in robotics.
  - Geometric intuition of all of these abstract concepts (i know this is not very relevant, but perhaps appreciated in the engineering world)
- Results:
  - Improvements in sample efficiency, robustness to noise, robustness to hyperparameters, generalization, and interpretability of Deep Learning based Koopman operator modeling.
  - Characterization of the low-dimensional synergetic nature of locomotion gaits in bipedal and quadrupedal systems. Basically that most relevant locomotion gaits evolve MOSTLY in a lower-dimensional subspace of the state space (the relevant isotypic spaces).

✤ Structure:

1. Introduction [2pg including abstract]
2. Preliminaries : 2.1. Modeling of dynamics, 2.2 Symmetries and Abstract Harmonic Analysis [2pg]
3. Symmetric Dynamical Systems : 3.1 Symmetric models, 3.2 Dynamic Harmonic Analysis [2pg] [mostly known with an insight that was not yet exploited which can be seen as contribution]
4. Symmetric Koopman-based data-driven models [1pg] [contribution]
5. Experiments [2.5pg] [contribution]
6. Conclusions, limitations and future work [0.5pg]

✤ Writing instructions: $\omega_t \in \Omega$ is idealized model of the dynamics (temporal evolution of a real object); $x \colon \Omega \to \mathcal{X}$ is a representation of a dynamics ω_t in a Hilbert space $\mathcal{X}$; $x_t \in \mathcal{X}$ is dynamical system (mathematical model of dynamics) approximating $\omega_t \in \Omega$

1. Introduction

Robotics demands dynamic modeling that goes beyond conventional approaches. Despite that, dynamics in robotics are often computed through highly-efficient algorithms for rigid-body systems. These algorithms, rooted in minimal-coordinate representation, leverage kinematic-induced structures to orchestrate recursive computations, as elegantly expounded by Featherstone (2007). These characteristics make them attractive for co-design, planning and estimation in robotics. Despite their merits, algorithms grounded in rigid body dynamics often grapple with the challenge of nonlinear optimization programs, as evidenced by various works (Mastalli et al., 2023; Kobilarov et al., 2015; Dinev et al., 2022). This can confine control policies and estimates to local minima, leading to inaccurate environmental assessments or unaccomplished tasks or sub-optimal robot designs—limitations that stifle the true potential of robotic systems. Moreover, while rigid body dynamics form a crucial subset, the broader landscape of robotics demands the anticipation of closed-loop dynamics in high-level decision-making scenarios, such as orchestrating intricate parkour maneuvers or enabling seamlessly human-robot interactions. To overcome the limitations mentioned earlier, this research is driven by the need to tackle the difficulties associated with building linear models using the Koopman theory. We aim to address these challenges by borrowing concepts from symmetric groups, isotypic decomposition, and harmonic analysis.

[Part regarding koopman in control - work in progress - i should add a part on forecasting, to pass the idea that a better model is the key] Recently, researchers started to exploit the Koopman Theory [cite] to find a transformation of the system's state space where the systems can evolve (approximately) linearly. The main idea is to lift the original system's state to a higher dimensional space (ideally infinite-dimensional), exploiting hand-crafted basis function [cite] or Neural Networks representation [cite]. As a consequence, a stronger linearization procedure (compared to the one obtainable with the classical Taylor's expansion) can be employed to approximate the nonlinear dynamics [it was a paper from asada, need to find it..], and standard linear control techniques, such as the Linear Quadratic Regulator (LQR) or linear Model Predictive Control (MPC) can be applied to more easily control the system [cite LQR/MPC]. Applications of such a technique can be widely found in the area of robotics. In [D Bruder, RSS, 2019] the authors learned a linear model of a soft robot by only employing input-output data gathered directly on the real system, showing impressive tracking performance of the resulting linear controller. Similar results were shown in [Giorgos Mamakoukas, TRO - Carl Folkestad, ICRA] for other classes of challenging underactuated systems.

Given the data-driven nature of such an approach, the main limitation lies in the quality of the underlying model, which is strongly dependent on the regression algorithm employed and the amount of collected data [some citations? brunton maybe and the technique of Vlad/Massi?]. In this paper, we tackle these challenges by exploiting the symmetries of a dynamical system to... **#here we should do a small introduction of the simmetries and why they reduce data/give better model..Then we can anticipate how we use them. @Daniel probably you can take care of this?**

Story: - world is nonlinear -¿ we do nonlinear controller, difficult and complex -¿ we linearize the system by Taylor for linear control -¿ strong but limiting, people start to use Koopman for a better "global" linear system identification, and then linear modeling/control/estimation -¿ data-driven, you need a good model, we can use symmetry to embeed the symmetries of the laws of physics governing the dynamcis of the actual system, obtaining better sample efficiency and improved generalization/performance (as now my model does not violates the laws of physics) (1). Furthermore, we can leverage symmetries to decompose the system state space. Why this is relevant?

- (LINEAR or NON-Linear dynamics, doesnt matter) Each isotypic subspace is (1) lower dimensional and (2) is constraint by a unique subset of symmetries. This is usefull becase it helps to identify the observable functions that I can use to "capture/model" the state of my system in this isotypic space. This are nor "any" obsrervable funtion, these are functions that have a known symmetry (another way to say it, these are functions constructed from simpler functions).
- For linear dynamics, this decomposition also entails the **decomposition of the dynamics**. Which is an inductive bias that leads me to decompose the problem of understanding/capturing/modeling how the system evolves in time, into several simpler (lower-dimensional) problems of understanding how my system evolves in each isotypic subspace.

if we notice that something is not useful (IN THE NONLINEAR SYSTEM), we can neglect some components, and we can reduce the size of the model (hence, less state variable when going to koopman, optimal control is easier! We can put additional important stuff in the future works..) (2) -¿ For (1) we show an experiment on a linear system, fom (2) we put experiment on this using a legged robot with a closed loop controller doing periodic gait, to show that for specific motion/system some component that is not relevant. In fact, we show in the figure that different motion/gait have different "power" on each decomposition. This is our contribution.

TODO: [Bad summary...] In this work, we introduce a theoretical and practical framework leveraging Abstract Harmonic Analysis (AHA) to decompose the *state space* of symmetric dynamical systems into an orthogonal finite sum of lower-dimensional subspaces, termed *harmonics state-subspaces*. Each subspace represents the range of the system's configurations respecting a unique subgroup of the system's symmetry group. This decomposition aims to translate the problem of modeling and control of the system's dynamics from a complex high-dimensional constrained optimization problem to a set of lower-dimensional, less-constrained (subgroup of the original symmetry constraints) problems. To motivate our analysis, we demonstrate that the harmonics state-subspaces are useful for characterizing the low-dimensional synergetic nature of locomotion gaits in bipedal and quadrupedal systems. Furthermore, we explore the implications of symmetry exploitation and AHA in modeling dynamical systems with Koopman operator theory. Since, for this modeling formalism, each harmonics state-subspace becomes an independent linear infinite-dimensional system that can be modeled with its own Koopman operator. The resultant decomposition of the Koopman operator leads to a data-driven modeling paradigm that adeptly captures both state and temporal symmetries.

Figure 1: TODO: SKETCH of the idea of the teaser/header image showing the Klein4 group (the simplest non-so-trivial group). Left: Show the partition of the state space and the representation of the systems state as points in this plane representing the Ω. This image will present the intiution of finite state symmetries and connect the mathematical notation to intiution. Not sure if just to focus on the robot (?). The plane should show some trajectories of motion clearly depicting the equivariance of the dynamics, and illustrations of the notations of groups and actions. Right: Depiction of the harmonic analysis of both systems. Associating each of the Isotypic components to a subgroup of the. There should be notation stating the low-dimensionality of the Isotupic spaces and the association with a lower symmetry group. Clearly there is a lot to improve but these images are easy to tweek and improve.

TODO: Remark that, while AHA has been largely exploited in modeling molecular dynamics and particle physics, these fields of knowledge do not address the problem of control. Therefore, the implications for controllable robotic/dynamical systems have been relatively unexplored. The plot of the distribution of kinetic energy and work in the Isotypic spaces should convey the idea that symmetric dynamical systems tend to have regimes of motion where some isotypic spaces play a larger role than others. This translates to the system mostly evolving in a lower-dimensional subspace of the state space (the relevant isotypic spaces). This is relevant for the purpuse of sampling data for learning the dynamics of the system (i.e., avoid collecting a dataset of motion that only exites a few isotypic spaces). But also from the perspective of control, since we can focus on controlling the system's motion in the relevant isotypic spaces, which are lower-dimensional and less constrained than the full state space.

1.1. Related work

+ i fill that this part and introduction can be unified. TODO:

- Mention Salova et al. (2019) paper on approximating Isotypic Decomposition of the Koopman operator with a (non-learnable) dictionary of observale functions. This is the most similar paper to ours. Key differences: We learn the relevant functions for each isotypic space, instead of hand crafing the dictionary of funcitons and projecting them to the isotypic spaces. The paper is strongly focused on the numerical efficiency gained by the block-diagonal decomposition, but fails to see the role of the Isotypic spaces as the decomposition of the motion of the system into a linear combination of state harmonics (less constrained and lower-dimensional motions).
- Mention Point-Group symmetries in chemistry and Normal Vibrational Modes in Molecular Dynamics [Please help Pietro].
- Mention applications of Harmonic Analysis to operators in physics. Solving the Schrodinger equation, and ... where else has this been used? [Again please help Pietro]
- Mention Brunton's latest paper in symmetries in learnign dynamical systems. Connect the lie derivative to the Morphological/Structural symmetries. Not really sure yet how to connect this without extending considerable in the explanation.
- Cite Otto and Rowley (2021), for potential uses of Operator/Koopman theory on estimation and control.

In this work, we focus on dynamical system, evolving in a Euclidean space of n-dimensions $\mathbb{E}_n$, which are composed of n_b bodies (or particles) interacting with one another and the environment (see fig. 1). We should state this clearly. I believe presenting the more abstract formalism: dynamical system evolving in a symmetric domain (e.g., manifold, grid, mesh, graph), which inherits some of the symmetries of the domain is too abstract to fit with the narrative, and the fact that we only have as examples systems evolvign in the Euclidean space. For grids, we can consider images/videos or even more fun animals on the Lenia game..

2. Preliminaries

In this section, we provide a concise overview of the fundamental concepts from two scientific domains that form the basis of our work. The first encompass dynamical systems theory, while the second delves into abstract harmonic analysis specifically focusing on point group symmetries.

Modeling of temporally evolving phenomena Dynamical systems are mathematical abstractions of temporally evolving phenomena. They are typically described via a tuple $(\Omega, \mathbb{T}, \Phi_\Omega)$, where Ω represents an abstract set of system's states $\boldsymbol{\omega} \in \Omega$, $\mathbb{T}$ represents time (typically $[0, T)$, $T \in \mathbb{R} \cup \{\infty\}$, or $\mathbb{N}$), and $\Phi_\Omega \colon \Omega \times \mathbb{T} \to \Omega$ is the evolution map that satisfies $\Phi_\Omega(\boldsymbol{\omega}, 0) = \boldsymbol{\omega}$ and $\Phi_\Omega(\Phi_\Omega(\boldsymbol{\omega}, t_1), t_2) = \Phi_\Omega(\boldsymbol{\omega}, t_1 + t_2)$ for all $\boldsymbol{\omega} \in \Omega$ and $t_1, t_2 \in \mathbb{T}$. Hence, one can write the trajectory of the system as $\boldsymbol{\omega}_t := \Phi_\Omega(\boldsymbol{\omega}_0, t)$, $t \in \mathbb{T}$, $\boldsymbol{\omega}_0 \in \Omega$.

Our prototypical example is rigid-body robot dynamics, where the domain of possible states Ω is formed by all feasible configurations (positions and velocities) of the robot's n_b rigid bodies in Euclidean space, i.e., $\Omega \subseteq \mathbb{SE}_1(3) \times \cdots \times \mathbb{SE}_{n_b}(3)$. Being $\mathbb{SE}_i(3)$ the i-th body configuration in 3-dimensions. The flow map Φ_Ω describes the dynamics governed by Newtonian mechanics. The classical model of this behavior is the Lagrangian dynamics model $(\mathcal{X}, \mathbb{T}, \Phi_\mathcal{X})$, for which the state vector space $\mathcal{X} \doteq \mathcal{Q} \times \mathcal{T}_{\boldsymbol{q}}\mathcal{Q}$ is composed of the space of generalized positions $\mathcal{Q}$ and velocity coordinates $\mathcal{T}_{\boldsymbol{q}}\mathcal{Q}$, and the map $\Phi_\mathcal{X}$ is defined by the Lagrangian equations of motion (Lanczos, 2012). Furthermore, we can consider each generalized position $\boldsymbol{q} = [q_1, \cdots, q_{|\mathcal{Q}|}]$ and velocity $\dot{\boldsymbol{q}} = [\dot{q}_1, \cdots, \dot{q}_{|\mathcal{T}_{\boldsymbol{q}}\mathcal{Q}|}]$ coordinate as a scalar observable function $q_i, \dot{q}_i : \Omega \to \mathbb{C}$, measuring, at every state $\boldsymbol{\omega}$, the position and velocity of one of the system's degrees of freedom.

This traditional approach selects a representation $\chi : \Omega \to \mathcal{X}$ that injectively maps the state $\boldsymbol{\omega}$ from the abstract set Ω to its representation $\chi(\boldsymbol{\omega})$ in a vector space $\mathcal{X}$, so that the temporal evolution $\boldsymbol{x}_t := \chi(\boldsymbol{\omega}_t)$, $t \in \mathbb{T}$, is numerically modeled by $\Phi_\mathcal{X}$. This helps to develop efficient algorithms for simulation, estimation, and control. In this context, see e.g. (Mezić, 2021), an optimal model of dynamics $(\boldsymbol{\omega}_t)_{t \in \mathbb{T}}$ minimizes the prediction error

$$\mathrm{err}(\Phi_\mathcal{X}, \boldsymbol{\omega}_0) := \int_\mathbb{T} ||\Phi_\mathcal{X}(\chi(\boldsymbol{\omega}_0), \mathrm{t}) \ominus \chi(\boldsymbol{\omega}_\mathrm{t})||^2 \mathrm{dt}, \tag{1}$$

where $\ominus$ is the difference operator needed to handle the state manifolds. We adopt this notation from (Mastalli et al., 2020), but it is generally used when optimizing over manifolds in (Gabay, 1982).

Depending on the complexity of the flow Φ_Ω, the tasks of identifying optimal models $(\mathcal{X}, \mathbb{T}, \Phi_\mathcal{X})$ can be analytically and computationally challenging. In simpler terms, the most manageable case arises when the dynamics in the vector space $\mathcal{X}$ are both autonomous and linear. This occurs when a system can be expressed as $\boldsymbol{x}_{t+\Delta t} = A_{\Delta t}\boldsymbol{x}_t$ for every $t, \Delta t \in \mathbb{T}$, with the linear operator $A_{\Delta t}$ dependent on Δt. In this scenarios, identifying the optimal model is straightforward: $\Phi_\mathcal{X}(\boldsymbol{x}_0, t) = A_1^t \boldsymbol{x}_0$ with $\mathbb{T} = \mathbb{N}_0$. Extending this concept to continuous time ($\mathbb{T} = [0, \infty) \subseteq \mathbb{R}$), the flow map is defined as $\Phi_{\mathcal{X}}(\boldsymbol{x}_0, t) = \exp(A_0\, t)\boldsymbol{x}_0$, where the linear operators A_1 and $A_0 := \lim_{\Delta t \to 0^+}(I - A_{\Delta t})/\Delta t$ are typically known. To simplify notations, whenever we discuss linear models we will write $(\mathcal{X}, \mathbb{T}, A)$, where linear operator A defining $\Phi_{\mathcal{X}}$ is either A_0 or A_1 depending on $\mathbb{T}$.

When the system's domain $\mathcal{X}$ is an Euclidean space $\mathbb{R}^d$, linear models can be efficiently solved by spectral decomposition of matrix A. This, in turns, facilitates the prediction steps needed in optimal control algorithms. However, Lagrangian models are typically nonlinear, leading to many roadblocks in developing efficient control algorithms. As stated in the introduction, in many cases is not possible to resort to a linearization procedure to synthetize an optimal controller, and one is forced to solve the optimal control problem in the original nonlinear form with iterative methods, which can exhibit slow convergence rendering them unsuitable for real-time control (see e.g cite???).

Luckily, relying on the transfer operator theory of Markov processes and, in particular, Koopman operator theory, one is able to represent large class of dynamical systems in $(\boldsymbol{\omega}_t)_{t\in\mathbb{T}} \subseteq \Omega$ as autonomous linear systems in a Hilbert space of functions $\mathcal{L}^2_\mu(\Omega)$, where μ is a measure on Ω that is preserved by the dynamics, [cite Lasota, Mezic etc ???] [If we consider stochastic processes I think we could cite our papers]. Namely, starting from a (potentially stochastic) Markov process $(\boldsymbol{\omega}_t)_{t\in\mathbb{T}}$, one can define a bounded linear operator $A_{\Delta t}\colon \mathcal{X} \to \mathcal{X}$ simply by

$$A_{\Delta t}\boldsymbol{x} := \mathbb{E}[\boldsymbol{x}(\Phi_\Omega(\boldsymbol{\omega}, \Delta t)) \,|\, \boldsymbol{\omega} = \cdot],\ \boldsymbol{x} \in \mathcal{L}^2_\mu(\Omega),$$

see [cite Lasota]. This operator is known as the forward transfer operator, or in case when $(\boldsymbol{\omega}_t)_{t\in\mathbb{T}}$ is deterministic flow, the Koopman operator. Using it, we can define the optimal linear autonoumous model $(\mathcal{X}, \mathbb{T}, \Phi_{\mathcal{X}})$ by $\Phi_{\mathcal{X}}(\boldsymbol{x}, t) := A_t\boldsymbol{x}$, $\boldsymbol{x} \in \mathcal{X} := \mathcal{L}^2_\mu(\Omega)$ and $t \in \mathbb{T}$.

Following, the transfer operator approach we are assured to have optimal linear autonomous models, however, the problem of having practical efficient algorithms is then mitigated to finding the appropriate representation $\chi\colon \boldsymbol{\omega}_t \mapsto \boldsymbol{x}_t$, [cite ???]. As we shall see in sec 4, this is where machine learning and, in particular, deep learning helps. [Add a ref?]

Finally, we remark that in the context of robotics, since the representation into generalized coordinates of the Lagrangian model is injective and bounded, the injectivity of the representation for the transfer/Koopman operator model is assured. So, having this in mind, in the following, we assume that $\mathcal{X}$ is a general (finite or infinite dimensional) separable Hilbert space where we will numerically model the dynamic.

Symmetry groups and their representations In this study, we aim to exploit symmetries present in robotic systems, e.g. robot Solo's body is symmetric w.r.t. lateral axis, see fig. 1. To this end, in the remainder of this section, we will briefly formalize the main aspects of the symmetry groups and their representations in vector spaces.

These symmetries describe transformations of state or time that yield states that are functionally equivalent under Newtonian physics. This equivalence implies that when symmetrical states are subjected to identical forces, these will evolve similarly in time (see fig. 1-left). Mathematically, the set of symmetries of a system is defined as a group $\mathbb{G} = \{e, g_1, g_2, \dots\}$ that contains the identity transformation e, and is closed under composition $g_a \cdot g_b \in \mathbb{G} \mid \forall\, g_a, g_b \in \mathbb{G}$, with $(\cdot) : \mathbb{G} \times \mathbb{G} \to \mathbb{G}$. Unless stated otherwise, we assume the group $\mathbb{G}$ to be a finite group.

TODO: [Make a note on defining $\diamond$ on abstract sets Ω and on $\mathbb{T}$, we can be bit informal.]

As the state of a dynamical system is represented as a point in some separable (finite or infinite dimensional) Hilbert space $\mathcal{H}$, we study symmetries via a group representation $\rho_{\mathcal{H}} : \mathbb{G} \to \mathbb{U}(\mathcal{H})$

that maps a group element g (symmetry) to its action $\rho_{\mathcal{H}}(g)$ on $\mathcal{H}$, which is a unitary matrix/operator transformation of the vector space $\mathcal{H}$, i.e. an element of $\mathbb{U}(\mathcal{H})$. Then, for any point $\boldsymbol{h} \in \mathcal{H}$, we get a g-symmetric point $g \diamond \boldsymbol{h}$ by applying an action $\rho_{\mathcal{H}}(g)$ of symmetry $g \in \mathbb{G}$ as $g \diamond \boldsymbol{h} := \rho_{\mathcal{H}}(g)\boldsymbol{h}$. Clearly, this operation $\diamond$ depends on the group representation $\rho_{\mathcal{H}}$. However, one can easily see that from one group representation $\rho_{\mathcal{H}}$, we can easily derive another one $\rho'_{\mathcal{H}}(\cdot) \doteq \mathsf{Q}\rho_{\mathcal{H}}(\cdot)\mathsf{Q}^*$, where Q is some unitary operator. Recalling that $\mathcal{H}$ can be identified by the Hilbert space ℓ^2 by fixing a specific basis, this just means that these two group representations are essentially the same, that is they are (isomorphic (equivalent up to the change of basis) which we write as $\rho'_{\mathcal{H}} \sim \rho_{\mathcal{H}}$. Of particular interest are so called irreducible unitary representations for which the only subspaces invariant under all group actions are trivial, that is if $\rho_{\mathcal{H}} : \mathbb{G} \to \mathbb{U}(\mathcal{H})$, V a subspace of $\mathcal{H}$ and $\rho_{\mathcal{H}}(g)V \subset V$ for every $g \in \mathbb{G}$, then either $V = \{0\}$ or $V = \mathcal{H}$.

To give an intuition to what follows, let's consider the quadruped robot in fig. 1 represented in the state space $\Omega \subseteq \mathbb{R}^{12}$???. The finite symmetry group of the states is the Klein group $\mathbb{G} = \mathbb{K}_4 = \{e, g_s, g_t, g_r\}$ that can be decomposed, using the conjugacy relation between group elements, into four conjugacy classes that are the trivial subgroup $\mathbb{G}_1 = \{e\}$, and 3 subgroups isomorphic to a reflection group: $\mathbb{G}_2 = \{e, g_s\}$, $\mathbb{G}_3 = \{e, g_t\}$, and $\mathbb{G}_4 = \{e, g_r\}$. These decomposition of a group into subgroups allows one to introduce a particular unitary group representation on an arbitrary separable Hilbert space that is in the focus of this paper. This is summarized in the following version of Peter-Weyl theorem, see (Knapp, 1986, Theorem 1.12) and (Golubitsky et al., 2012, Theorem-2.5), a key result of representation theory.

Theorem 1 (Isotypic decomposition) *Let $\mathbb{G}$ a compact symmetry group, and $(\mathbb{G}_i)_{i=1}^k$ the subgroups associated with the group's k conjugacy classes. Let $\mathcal{H}$ be a separable Hilbert space and $\rho_{\mathcal{H}} : \mathbb{G} \to \mathbb{U}(\mathcal{H})$ be a unitary group representation. Then the following hold.*

(i) Space $\mathcal{H}$ can be decomposed into a direct sum of k mutually orthogonal subspaces $(\mathcal{H}_i)_{i=1}^k$, i.e. $\mathcal{H} = \mathcal{H}_1 \oplus^{\perp} \mathcal{H}_2 \oplus^{\perp} \cdots \oplus^{\perp} \mathcal{H}_k$, where $\rho_{\mathcal{H}} \sim \rho_{\mathcal{H}_1} \oplus \rho_{\mathcal{H}_2} \oplus \cdots \oplus \rho_{\mathcal{H}_k}$ and $\rho_{\mathcal{H}_i} \colon \mathbb{G}_i \to \mathbb{U}(\mathcal{H}_i)$ is a unitary representation of a subgroup $\mathbb{G}_i$ on the subspace $\mathcal{H}_i$, $i = 1, \dots, k$.

(ii) For each $i \in \{1, \dots, k\}$ there exists a subspace $\mathcal{H}_{i,1}$, $d_i := \dim(\mathcal{H}_{i,1}) < \infty$, an irreducible unitary representation $\bar{\rho}_i \colon \mathbb{G}_i \to \mathbb{U}(\mathcal{H}_{i,1})$ and at most countably many subspaces $\mathcal{H}_{i,j}$, $j = 2, \dots m_i$, isometrically-isomorphic to $\mathcal{H}_{i,1}$, so that $\mathcal{H}_i = \mathcal{H}_{i,1} \oplus^{\perp} \cdots \oplus^{\perp} \mathcal{H}_{k,m_i}$ and $\rho_{\mathcal{H}_i} \sim \bigoplus_{j=1}^{m_i} \bar{\rho}_i$.

As this results indicates, the group structure, conjugacy classes in particular, is carried over to unitary group representations on an arbitrary Hilbert space $\mathcal{H}$. Indeed, according to thm. 1(i) $\mathcal{H}$ is decomposed into finite number of mutually orthogonal subspaces $\mathcal{H}_i$, each, according to thm. 1(ii), related to a conjugacy class of the group via its irreducible representation $\bar{\rho}_i$. Hence, the name isotypic subspaces (same-type) for $\mathcal{H}_i$'s and isotypic components for $\rho_{\mathcal{H}_i}$'s. The number m_i of "copies" (isometrically-isomorphic subspaces) for each isotypic component is called its multiplicity, and the dimension $d_[i]$ is the minimal dimension to define $\mathbb{G}_i$-symmetric elements. If $\mathcal{H}$ is finite-dimensional, then all multiplicities are finite, while in the case of infinite-dimensional Hilbert spaces, isotypic subspaces can be infinite-dimensional yielding isotypic components with infinite multiplicity. Going back to our example of robot Solo, computing the isotypic decomposition of space $\mathbb{R}^{12}$ for the Klein group, allows one to project the state of the robot into each isotypic subspace, see fig. 1(right).

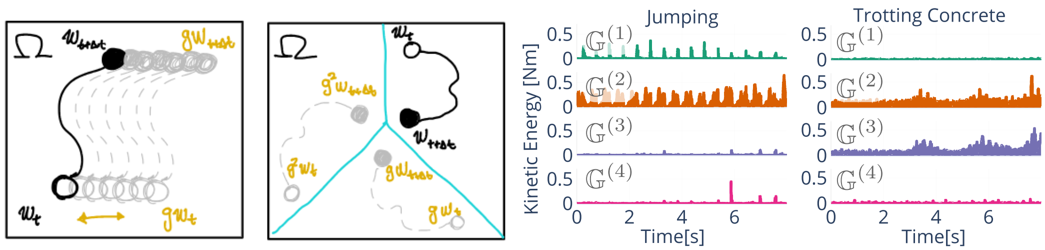

Figure 2: Representation of the studied symmetry groups: a) symmetries of the set of states, b) symmetries of the of time $\mathbb{T}$. to use the images we modified online, to put a) and b) under the images. Then put some explanations.

3. Symmetries of dynamical systems

In this section we show how point group symmetries, and in particular abstract harmonic analysis, can be used to build faithful models of dynamics.

To that end, consider the dynamical system $(\Omega, \mathbb{T}, \Phi_\Omega)$. There exist two types of symmetries the system might possess: state symmetries ($\mathbb{G}_\Omega$) and temporal symmetries ($\mathbb{G}_\mathbb{T}$). constraining the system's evolution:

$$\underbrace{g \diamond \boldsymbol{\omega}_t = \Phi_\Omega(g \diamond \boldsymbol{\omega}_0, t)}_{\text{state symmetry}} \mid \forall\, g \in \mathbb{G}_\Omega \text{ and } \underbrace{\boldsymbol{\omega}_{g \diamond t} = \Phi_\Omega(\boldsymbol{\omega}_0, g \diamond t)}_{\text{temporal symmetry}} \mid \forall\, g \in \mathbb{G}_\mathbb{T}, \tag{2}$$

respectively, where equations hold for all $t \in \mathbb{T}$ and $\boldsymbol{\omega}_0 \in \Omega$. As eq. (2) indicates, symmetries are formalized as equivariance constraints that describe properties of both the set of possible states (Ω) and of the manner in which states evolve in time (Φ_Ω). These properties translate to valuable geometric biases in modeling and control, which we now briefly mention.

State Symmetries: A symmetry group on the set of states Ω describe an equivalence relationship between any state $\boldsymbol{\omega} \in \Omega$ and its set of symmetric states: the *state orbit* $\mathbb{G}_\Omega\boldsymbol{\omega} = \{g \diamond \boldsymbol{\omega} | g \in \mathbb{G}_\Omega\}$ (see fig. 1). This relationship effectively describes how all symmetric states $\boldsymbol{\omega} \in \mathbb{G}_\Omega\boldsymbol{\omega}$ will evolve with a unique (or equivalent) trajectory of motion, up to a symmetry transformation $g \in \mathbb{G}_\Omega$.

The implications of state symmetries vary depending on whether $\mathbb{G}_\Omega$ is a finite or continuous group. Finite state symmetry groups, imply that the set of states $\Omega = \cup_{g \in \mathbb{G}_\Omega}\{g \diamond \Omega \,/\, \mathbb{G}_\Omega\}$ is defined as the union of a finite number of symmetry-transformed copies of the quotient set $\Omega/\mathbb{G}_\Omega$, representing the set of *unique* system's states (see fig. 1 and **??**). This finite decomposition of Ω is the property that enable the use of abstract harmonic analysis to decompose any state vector space representation of Ω (see **??**), and thus, the property we aim to exploit in this work.

Continuous (or Lie group) state symmetries are well studied and extensively exploited in simulation, state estimation, modeling and control (Chirikjian and Kyatkin, 2000). In a nutshell, for modeling, these symmetries help to identify *ignorable* measurements of the system's state, irrelevant for characterizing Φ_Ω. Commonly, these measurements are ignorable because of their dependence to the arbitrary selection of a coordinate reference frame.

Temporal Symmetries: A symmetry group of $\mathbb{T}$ can imply that the system's dynamics are either periodic and/or time-reversible. That is, for systems featuring periodic stable trajectories (limit-cycles) with period p, $\mathbb{T}$ becomes a ring $\mathbb{T} \doteq \mathbb{R}/p\mathbb{R}$ or $\mathbb{Z}/p\mathbb{Z}$, and consequently $\mathbb{G}_\mathbb{T}$ becomes a cyclic group $\mathbb{G}_{\mathbb{T}} \doteq \mathbb{C}_p$. While for time-reversible systems, we have that $\mathbb{T} \doteq \mathbb{R}$ or $\mathbb{Z}$, and the $\mathbb{G}_{\mathbb{T}}$ becomes the reflection group $\mathbb{G}_{\mathbb{T}} \doteq \mathbb{C}_2$.

Moving forward, our focus will be on studying the implications of finite state symmetry groups. Unless specified otherwise, $\mathbb{G}$ will denote a state symmetry group $\mathbb{G}_{\Omega}$. Formally, systems with a state symmetry group $\mathbb{G}$ are defined as follows.

Definition 1 (Symmetric dynamical systems) *A dynamical system* $(\Omega, \mathbb{T}, \Phi_{\Omega})$ *is* $\mathbb{G}$*-symmetric, if* $\mathbb{G}$ *is a symmetry group of* Ω*, and the system's dynamics are* $\mathbb{G}$*-equivariant* $\Phi_{\Omega}(g \diamond \boldsymbol{\omega}, t) = g \diamond \Phi_{\Omega}(\boldsymbol{\omega}, t) \mid \forall\, g \in \mathbb{G}, t \in \mathbb{T}$.

When designing a numerical model $(\mathcal{X}, \mathbb{T}, \Phi_{\mathcal{X}})$ of a symmetric dynamical system on a vector space $\mathcal{X}$, we can either ignore or inherit the group structure. This, depends on the existence of the group representation $\rho_{\mathcal{X}} : \mathbb{G} \to \mathbb{U}(\mathcal{X})$, translating the symmetric relationship between states $\mathbb{G}\boldsymbol{\omega} \doteq \{g \diamond \boldsymbol{\omega} \mid \forall\ g \in \mathbb{G}\}$ to the points on $\mathcal{X}$ representing these states : $\mathbb{G}\boldsymbol{x}(\boldsymbol{\omega}) \doteq \{g \diamond \boldsymbol{x}(\boldsymbol{\omega}) \doteq \rho_{\mathcal{X}}(g)\boldsymbol{x}(\boldsymbol{\omega}) = \boldsymbol{x}(g \diamond \boldsymbol{\omega}) \mid \forall\ g \in \mathbb{G}\}$. Following def. 1, we will denote numerical models with symmetric vector spaces and $\mathbb{G}$-equivariant dynamics as symmetric models.

A familiar example of a symmetric model for $\mathbb{G}$-symmetric robotic systems, such as robot Solo in fig. 1, is Lagrangian dynamics. This model $(\mathcal{X}, \mathbb{T}, \Phi_{\mathcal{X}})$ defines the vector space $\mathcal{X} \doteq \mathcal{Q} \times \mathcal{T}_{\boldsymbol{q}}\mathcal{Q}$ as the space of generalized coordinates (with the associated group representation $\rho_{\mathcal{X}} \doteq \rho_{\mathcal{Q}} \oplus \rho_{\mathcal{T}_{\boldsymbol{q}}\mathcal{Q}}$), and the dynamics $\Phi_{\mathcal{X}}$ as the standard Lagrangian equations of motion, describing rigid-body dynamics. These equations are by construction $\mathbb{G}$-equivariant (Lanczos, 2012; Ordonez-Apraez et al., 2023).

Inheriting the group structure on the vector space of a numerical model becomes a valuable design principle as this implies the reduction of the search space for the map $\Phi_{\mathcal{X}}$ to the space of $\mathbb{G}$-equivariant maps:

Proposition 2 (Optimal models of $\mathbb{G}$-symmetric systems) *Let* (Ω, Φ_{Ω}) *be a* $\mathbb{G}$*-symmetric dynamical system and* $(\mathcal{X}, \Phi_{\mathcal{X}})$ *its optimal numerical model. If* $\mathcal{X}$ *is a symmetric space,* $\Phi_{\mathcal{X}}$ *is* $\mathbb{G}$*-equivariant.*

Proof [check and correct notation!] Consider *any* model $(\mathcal{A}, \mathbb{T}, \Phi_{\mathcal{A}})$ with $\mathcal{A}$ a symmetric space with group representation $\rho_{\mathcal{A}}$. First, take any orbit of symmetric states $\mathbb{G}_{\Omega}\boldsymbol{a}(\boldsymbol{\omega}_0)$, and identify the state with the lowest prediction error $\hat{g} \diamond \boldsymbol{a}(\boldsymbol{\omega}_0)$ given $\hat{g} = \arg\min_{g \in \mathbb{G}_{\Omega}} \operatorname{err}(\Phi_{\mathcal{A}}, \boldsymbol{\omega}_0)$ (eq. (1)). Then, define a new $\mathbb{G}_{\Omega}$-equivariant map that mimics the best prediction for all symmetric states $\bar{\Phi}_{\mathcal{A}}(\boldsymbol{a}(g \diamond \boldsymbol{\omega}_0), t) \doteq (g \cdot \hat{g}^{-1}) \diamond \Phi_{\mathcal{A}}(\boldsymbol{a}(\hat{g} \diamond \boldsymbol{\omega}_0), t)) \mid \forall\ \boldsymbol{\omega} \in \Omega, g \in \mathbb{G}_{\Omega}$, lowering the overall prediction error. Iteratively repeating this process for all $\boldsymbol{\omega} \in \Omega$ results in $\bar{\Phi}_{\mathcal{A}}$ being $\mathbb{G}_{\Omega}$-equivariant. ■

Now, recalling thm. 1, for a symmetry group $\mathbb{G}$ consisting of k conjugacy classes $\mathbb{G}_i$'s we can decompose the vector space of a $\mathbb{G}$-symmetric model into k isotypic subspaces $\mathcal{X} = \bigoplus_{i=1}^{k} \mathcal{X}_i$. Hence, the states of the $\boldsymbol{x}_t$ can be decomposed, across different times $t \in \mathbb{T}$, into projections on each isotypic subspace $\mathcal{X}_i$ in order to observe $\mathbb{G}_i$-symmetric component of dynamics. While this isotypic decomposition at each time step is [ROBO: definitely useful in robotics, [cite ???] or see [fig ???]], in general it does not allow one to decompose the flow into separable components which would allow one improve on the efficiency and interoperability of the model. Again, having linear models brings forward such benefits, which a consequence of the standard result in harmonic analysis called Schur's lemma (Knapp, 1986, Proposition 1.5). In a nutshell, this lemma is a result on the equivariant linear maps between vector spaces with irreducible representations of the same symmetry group, and it states that if the spaces are not isomorphic (different irreducible representations), then the map is trivial (i.e. equal to zero), while for the finite-dimensional spaces associated with the same irreducible representation nontrivial maps are scalar multiples of identity. Thus, for linear models by projecting the flow on the isotypic subspaces and applying Schur's lemma we obtain the following result, essentially stated in (Golubitsky et al., 2012, Theorem 3.5) for finite dimensional spaces. For simplicity we present it for $\mathbb{T} = \mathbb{N}_0$.

Theorem 3 (Isotypic decomposition of symmetric linear models) *Let $\mathbb{G}$ be a finite group, let $(\mathcal{X}, \mathsf{A})$ be $\mathbb{G}$-symmetric linear model, and let $\mathcal{X} = \bigoplus_{i=1}^{k} \mathcal{X}_i$ be isotypic decomposition. Then A is block-diagonal $\mathsf{A} = \bigoplus_{i=1}^{k} \mathsf{A}_i$, where diagonal blocks $\mathsf{A}_i \colon \mathcal{X}_i \to \mathcal{X}_i$ are $\mathbb{G}_i$-equivariant projections onto the isotypic subspaces represented by $m_i \times m_i$ matrices $\mathsf{A}_i =$. Moreover, the $\mathbb{G}$-symmetric model $(\mathcal{X}, A)$ is separable into sub-models $(\mathcal{X}_i, \mathsf{A}_i)$ that are $\mathbb{G}_i$-symmetric, $i = 1, \dots, k$ and* $\mathrm{err}(\mathsf{A}) = \sum_{\mathrm{i=1}}^{\mathrm{k}} \mathrm{err}(\mathsf{A}_\mathrm{i})$ *holds.*

[ROBO: @CARLOS @Giulio -¿¿ review] [rewrite this saying why decomposable models are useful] I think we should focus only on koopman for control or forecasting -¿ we can discard some state on the original nonlinear form, hence fit using koopman a better model. We show where this useless state is in the experiment of the Quadruped. The other parts should be future works/conclusion.. We believe there are two main applications of thm. 3. The first is the decomposition of the local linearizations of non-linear dynamical systems, used in iLQR (**?**) and DDP (**?**), into k independent linear models, each evolving the isotypic subspaces. This decomposition enables the parallel optimization of the local $\mathbb{G}_i$-equivariant linearizations, which is a promising direction for future work. The second application is the exploitation of the isotypic decomposition for the block diagonalization of the Koopman operator. This is the focus of the next section.

Dynamic Harmonic Analysis. When applied to Koopman model thm. 3 becomes particularly useful. Namely,...[rewrite this] As stated in thm. 3, for a dynamical system $(\Omega, \mathbb{T}, \Phi_\Omega)$ with a state symmetry group $\mathbb{G}$ and $\mathbb{G}$-equivariant dynamics Φ_Ω, any linear model can be decomposed into k submodels. Although in practice, most systems of interest have non-linear dynamics Φ_Ω, it is important to note, that all dynamical system possess a linear model if the state vector space is chosen as the entire space of observable functions $\mathcal{L}^2_\mu(\Omega)$. [...] The operator theoretic paradidm for modeling dynamical system's relies in the study of the evolution of the system dynamics on an infinite-dimensional function space. Using the nomenclature of sec 3, this translates to defining a model $(\Phi_\mathcal{X}, \mathcal{X})$ in which the state space is the entire space of observable functions $\mathcal{X} \doteq \mathcal{L}^2_\mu(\Omega)$, and the dynamics of any observable function becomes linear $\chi_{t+\Delta t} = \Phi_\mathcal{X}(\chi(\boldsymbol{\omega}_t), \Delta t) \doteq \mathsf{U}^{\Delta t} \diamond \chi(\boldsymbol{\omega}_t) = \chi(\boldsymbol{\omega}_t; \mathsf{U}^{\Delta t} \curvearrowleft)$

4. Symmetric Data-driven Koopman models

In practice, from data, one can only approximate the linear dynamics of the infinite-dimensional model with an empirical finite-dimensional one ($\mathcal{Z} \subseteq \mathbb{C}^{|\mathcal{Z}|}, \mathsf{Z}_{\Delta t} \in \mathbb{C}^{|\mathcal{Z}| \times |\mathcal{Z}|}$). This optimization process involves finding a finite set of relevant observable functions that span the state vector space $\mathcal{Z} = \mathrm{span}\{z_i(\boldsymbol{\omega})\}_{j=1}^{|\mathcal{Z}|} \mid z_j \in \mathcal{L}^2_\mu(\Omega), j \in [1, |\mathcal{Z}|]$, together with an approximation of the empirical Koopman operator on this space $\mathsf{Z}_{\Delta t}$.

TODO: talk about training trajectories, samples, etc. Potentially pseudo-code. Reserve a symbol n for sample size. Diff from dimensionality of the states of the system To approach this problem we use the Dynamics Auto-Encoder (DAE) deep learning architecture (Lusch et al., 2018), which defines an encoder-decoder architecture that optimizes for $\mathcal{Z}$ and $\mathsf{Z}_{\Delta t}$ jointly. In the DAE architecture, the observable functions (encoder) $\boldsymbol{z} : \Omega \rightarrow \mathcal{Z}$ and their inverse (decoder) $\boldsymbol{z}^{-1} : \mathcal{Z} \rightarrow \Omega$ are parameterized by trainable neural networks. Likewise, the empirical operator $\mathsf{Z}_{\Delta t}$ is parameterized as a trainable matrix. The cost function for DAE is composed of an autoencoder reconstruction loss, a prediction error in $\operatorname{err}(\Phi_\Omega)$, and a linear Koopman prediction error $\operatorname{err}(\Phi_{\mathcal{Z}})$:

$$\mathcal{C}_{DAE}(\boldsymbol{\omega}_0) = \underbrace{||\boldsymbol{z}^{-1}(\boldsymbol{z}(\boldsymbol{\omega}_0)) - \boldsymbol{\omega}_0||^2}_{\text{Reconstruction}} + \textstyle\sum_{t=1}^{H} [\underbrace{||\boldsymbol{z}^{-1}(\mathsf{Z}_{\Delta t}^{t}\boldsymbol{z}(\boldsymbol{\omega}_0)) - \boldsymbol{\omega}_t||^2}_{\operatorname{err}_{\boldsymbol{\omega}_0}(\Phi_\Omega)\text{: Prediction}} + \lambda_{\mathcal{Z}} \underbrace{||\mathsf{Z}_{\Delta t}^{t}\boldsymbol{z}(\boldsymbol{\omega}_0)) - \boldsymbol{z}(\boldsymbol{\omega}_t)||^2}_{\operatorname{err}_{\boldsymbol{\omega}_0}(\Phi_{\mathcal{Z}})\text{: Linear prediction}}] \quad (3)$$

Where H defines the prediction horizon's steps, and λ_{Z}, set to $\lambda_{\mathsf{Z}} = \sqrt{|\mathcal{Z}|/|\Omega|}$, balances the error in a dimension of $\mathcal{Z}$ with that in a dimension of Ω, a strategy we found effective in practice Any references on this? Or you are the first doing it?. I think first.

The equivariant DAE (eDAE): By proposition 2, if $\mathcal{Z}$ is a symmetric space, the optimal operator $\mathsf{Z}_{\Delta t}$ is $\mathbb{G}$-equivariant, and consequently by thm. 3, it can be decomposed into several linear systems. To ensure $\mathcal{Z}$ is a symmetric space the learned observable functions $\boldsymbol{z} : \Omega \rightarrow \mathcal{Z}$ need be $\mathbb{G}$-equivariant $g \diamond \boldsymbol{z}(\boldsymbol{\omega}) = \boldsymbol{z}(g \diamond \boldsymbol{\omega}) \mid \forall\, g \in \mathbb{G}, \boldsymbol{\omega} \in \Omega$. This enforces the symmetry of the space by defining the group representation $\rho_{\mathcal{Z}} : \mathbb{G} \rightarrow \mathbb{U}(\mathcal{Z})$. Then, we can obtain a symmetric model (sec 3) by constraining the trainable Koopman operator to be $\mathbb{G}$-equivariant.

Furthermore, by changing basis the basis of $\mathcal{Z}$ to the isotypic basis (thm. 1), we expose the isotypic subspaces $\mathcal{Z} = \bigoplus_{i=1}^{k} \mathcal{Z}_i$, and their associated group representations $\rho_{\mathcal{Z}} = \bigoplus_{i=1}^{k} \rho_{\mathcal{Z}_i}$. In this basis, a $\mathbb{G}$-equivariant Koopman operator will factorize into a block-diagonal sum of Koopman operators acting on each isotypic subspace: $\mathsf{Z}_{\Delta t} \doteq \bigoplus_{i=1}^{k} \mathsf{Z}_{i,\Delta t}$ (thm. 3). Each of these is required to be equivariant to a unique subgroup of the state symmetry group $\rho_{\mathcal{Z}_i}(g)\mathsf{Z}_{i,\Delta t} = \mathsf{Z}_{i,\Delta t}\rho_{\mathcal{Z}_i}(g) \mid \forall\, g \in \mathbb{G}^{(i)} \subseteq \mathbb{G}, i \in [1,k]$. This decomposition, on top of providing relevant numerical implications, represent a significant gain in the interpretability of the learned model, since the observable functions spanning each isotypic subspace are constrained to measure relevant quantities of the state that transform only under the symmetry subgroup $\mathbb{G}^{(i)}$.

To ground some intuition, recall that we have already studied examples of such symmetric sets of observable functions when we analyzed the space of generalized position coordinates $\mathcal{Q}$ of the quadruped robot Solo in the isotypic basis $\mathcal{Q} = \bigoplus_{i=1}^{k} \mathcal{Q}_i$ (see sec 2 and fig. 1). In this scenario, each isotypic observable function measured a type of harmonic motion of the system's degrees of freedom, that respected only one of the reflection subgroups $\mathbb{C}_2^{(i)} \subset \mathbb{G}$ of the robots symmetry group $\mathbb{G} = \mathbb{K}_4$.

In practice, to achieve this structure, we parameterize the observable function (encoder) and its inverse (decoder) as $\mathbb{G}$-equivariant neural networks, and each isotypic subspace' Koopman operator as a linear $\mathbb{G}^{(i)}$-equivariant map, using the ESCNN library (Cesa et al., 2022).

5. Experiments and results

We present two experiments comparing the performance between equivariant and non-equivariant Koopman models using the DAE and eDAE architectures outlined in the last section.

Synthetic symmetric dynamical systems with finite state symmetry groups In this experiment, we model synthetic non-linear symmetric dynamical systems with arbitrary state symmetry groups $\mathbb{G}$. The systems are designed as constrained stable linear stochastic systems $\boldsymbol{a}_{t+\Delta t} = \mathsf{A}\boldsymbol{x}_t + \boldsymbol{\epsilon} \mid \mathsf{C}\boldsymbol{a} \geq \boldsymbol{c}$. Where $\boldsymbol{a} \in \mathcal{A} \subseteq \mathbb{R}^n$ represent the system state, $\mathsf{A} \in \mathbb{R}^{n \times n}$ the linear dynamics matrix, $\boldsymbol{\epsilon} \in \mathbb{R}^n$ is a white-noise stochastic process with standard deviation σ, and $\mathsf{C} \in \mathbb{R}^{n_c \times n}$, $\boldsymbol{c} \in \mathbb{R}^{n_c}$ the parameters describing n_c inequality constraints on the dynamics, that make the overall dynamics

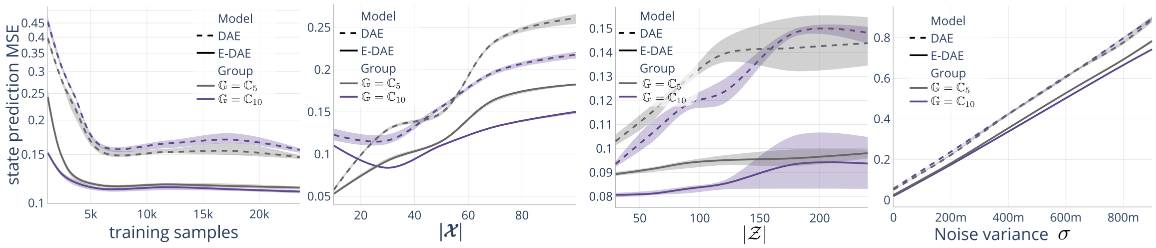

Figure 3: Performance of learned Koopman models using DAE and E-DAE algorithms for modeling synthetic systems with symmetry groups $\mathbb{C}_5$ and $\mathbb{C}_{10}$, different state dimensions $|\mathcal{X}|$, modeling state dimension $|\mathcal{Z}|$, and noise variance σ. Solid lines and shaded areas represent the mean, maximum and minimum prediction error among 4 training seeds. **Left**: Test prediction MSE vs. samples in training dataset. **Left-Middle:** Test prediction MSE for systems with varying state dimension $|\mathcal{X}|$. **Right-Middle:** Test prediction MSE for Koopman models with varying state dimension $|\mathcal{Z}|$. **Right:** Test prediction MSE for systems increasing noise variance σ.

non-linear. This systems are $\mathbb{G}$-symmetric if A is $\mathbb{G}$-equivariant $\rho_{\mathbb{R}^n}(g)\mathsf{A} = \mathsf{A}\rho_{\mathbb{R}^n}(g) \mid \forall\, g \in \mathbb{G}$ and any constraint is also enforced for all symmetric states $\boldsymbol{C}_{k,:}\ g \diamond \boldsymbol{a} \geq \boldsymbol{c}_k \mid \forall\, k \in [1, n_c], g \in \mathbb{G}$.

The parametric construction of these systems enables us to test the impact of symmetry exploitation in learning Koopman models for arbitrary groups $\mathbb{G}$, system's dimensionality $|\mathcal{A}|$, dimensionality of the Koopman model state $|\mathcal{Z}|$, and standard deviation σ of the noise. The results of the experiment, summarized in fig. 3, show that the eDAE architecture leads to overall better forecasting models featuring enhanced sample efficiency and generalization (fig. 3-left), reduced sensitivity to the system's dimensionality $|\mathcal{A}|$ (fig. 3-middle-left), reduced sensitivity to the dimension of the Koopman model state $|\mathcal{Z}|$ (fig. 3-middle-right), and mildly better robustness to noise fig. 3-right.

Modeling quadruped $\mathbb{G}_\Omega$-equivariant controlled dynamics In this experiment, we explore the potential use of a Koopman model as a surrogate rigid-body dynamics linear model, while quantifying the role of symmetry exploitation in this objective. The objective is to model the closed-loop dynamics of locomotion of the mini-cheetah quadruped robot on a mildly uneven terrain. The robot is controlled with a model predictive controller tracking a desired target base velocity with a prefixed trotting periodic gait, by jointly optimizing ground reaction forces and feet locations (**?**) TODO: @GIULIO PLEASE FIX REF. Thus, the closed-loop dynamics are stable with a limit-cycle trajectory describing the periodic gait and transient dynamics modeling how the controller corrects for tracking errors. The state of the system is defined as $\boldsymbol{\omega}_t = [\boldsymbol{q}_t, \dot{\boldsymbol{q}}_t, z_t, \boldsymbol{o}_t, \boldsymbol{v}_{\text{err},t}, \boldsymbol{w}_{\text{err},t}] \in \Omega \subseteq \mathbb{R}^{46}$. Composed of the joint's generalized position coordinates $\boldsymbol{q} \in \mathcal{Q} \subseteq \mathbb{R}^{24}$, and velocity coordinates $\dot{\boldsymbol{q}} \in \mathcal{T}_{\boldsymbol{q}}\mathcal{Q} \subseteq \mathbb{R}^{12}$, base height $z_t \in \mathbb{R}^1$, base orientation quaternion $\boldsymbol{o} \in \mathbb{R}^4$, and the error in the desired linear and angular base velocities $\boldsymbol{v}_{\text{err},t} \in \mathbb{R}^3$ and $\boldsymbol{w}_{\text{err},t} \in \mathbb{R}^3$, respectively.

This robot features a state symmetry group of order 8, $\mathbb{G}_\Omega = \mathbb{K}_4 \times \mathbb{C}_2$. The equivariance of the controller ensures that the closed-loop dynamics Φ_Ω remain $\mathbb{G}_\Omega$-equivariant, while the prefixed gait pattern ensures the periodicity of the dynamics.

Given the system's symmetry group, the training dataset, collected from a few trajectories of motion, may not adequately sample all 8 copies of the state quotient set $\Omega/\mathbb{G}_\Omega$ (fig. 2-left). Therefore, for this experiment, exploiting symmetry becomes crucial to mitigate the effects of the curse of dimensionality and the bias of the training to a reduced volume of the entire state space. The results, summarized in fig. 4, confirm this, showing that the DAE algorithm is unable to learn a model that generalizes to the entire state space. The use of data-augmentation (DAE$_{\text{aug}}$), exploiting the entire group $\mathbb{K}_4 \times \mathbb{C}_2$ and subgroups $\mathbb{K}_4$ and $\mathbb{C}_2$, alleviate the bias of the training data, resulting in improved models. There is a clear correlation between the number of symmetries exploited and the performance of the model. The best results are obtained when enforcing equivariance (eDAE), leading to a model that generalizes to the entire state space.

TODO: Discuss reduction of variance across training seeds, and robustness to the selecion of $|\mathcal{X}|$

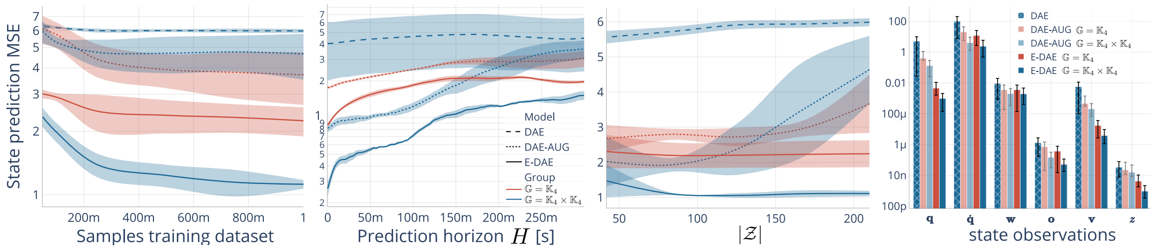

Figure 4: Test set MSE for learned Koopman models using DAE, DAE$_{\text{aug}}$ and E-DAE architectures for the mini-cheetah robot, exploiting the system's entire symmetry group $\mathbb{G} = \mathbb{K}_4 \times \mathbb{C}_2$ and only a subgroup $\mathbb{G} = \mathbb{K}_4$. Solid lines and shaded areas represent the mean, maximum and minimum prediction error among 4 training seeds. **Left**: Sample efficiency exploiting the entire symmetry group. **Left-Middle:** Prediction error for systems with varying state dimension $|\mathcal{X}|$. **Right-Middle:** Prediction error for systems with state dimension $|\mathcal{X}| = 50$ and varying modeling state dimension $|\mathcal{Z}|$. **Right:** Prediction error for systems varying noise variance σ.

6. Conclusions

Acknowledgments

We thank a bunch of people.

References

Gabriele Cesa, Leon Lang, and Maurice Weiler. A program to build E(N)-equivariant steerable CNNs. In International Conference on Learning Representations, 2022. URL https://openreview.net/forum?id=WE4qe9xlnQw.

Gregory S Chirikjian and Alexander B Kyatkin. Engineering applications of noncommutative harmonic analysis: with emphasis on rotation and motion groups. CRC press, 2000.

Traiko Dinev, Carlos Mastalli, Vladimir Ivan, Steve Tonneau, and Sethu Vijayakumar. A versatile co-design approach for dynamic legged robots. In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2022.

R. Featherstone. Rigid Body Dynamics Algorithms. Springer-Verlag, Berlin, Heidelberg, 2007.

D. Gabay. Minimizing a differentiable function over a differential manifold. Journal of Optimization Theory and Applications, 37, 1982.

Martin Golubitsky, Ian Stewart, and David G Schaeffer. Singularities and Groups in Bifurcation Theory: Volume II, volume 69. Springer Science & Business Media, 2012.

Anthony W. Knapp. Representation Theory of Semisimple Groups. Princeton University Press, Princeton, 1986. ISBN 9781400883974. doi: doi:10.1515/9781400883974. URL https://doi.org/10.1515/9781400883974.

Marin Kobilarov, Duy-Nguyen Ta, and Frank Dellaert. Differential dynamic programming for optimal estimation. In IEEE International Conference on Robotics and Automation (ICRA), 2015.

Cornelius Lanczos. The variational principles of mechanics. Courier Corporation, 2012.

Bethany Lusch, J Nathan Kutz, and Steven L Brunton. Deep learning for universal linear embeddings of nonlinear dynamics. Nature communications, 9(1):4950, 2018.

Carlos Mastalli, Rohan Budhiraja, Wolfgang Merkt, Guilhem Saurel, Bilal Hammoud, Maximilien Naveau, Justin Carpentier, Ludovic Righetti, Sethu Vijayakumar, and Nicolas Mansard. Crocoddyl: An efficient and versatile framework for multi-contact optimal control. In IEEE International Conference on Robotics and Automation (ICRA), 2020.

Carlos Mastalli, Saroj Prasad Chhatoi, Thomas Corbéres, Steve Tonneau, and Sethu Vijayakumar. Inverse-dynamics mpc via nullspace resolution. IEEE Transactions on Robotics, 39:3222–3241, 2023.

Igor Mezić. Koopman operator, geometry, and learning of dynamical systems. Not. Am. Math. Soc., 68(7):1087–1105, 2021.

Daniel F Ordonez-Apraez, Mario Martin, Antonio Agudo, and Francesc Moreno. On discrete symmetries of robotics systems: A group-theoretic and data-driven analysis. In Proceedings of Robotics: Science and Systems, Daegu, Republic of Korea, July 2023. doi: 10.15607/RSS.2023.XIX.053.

Samuel E Otto and Clarence W Rowley. Koopman operators for estimation and control of dynamical systems. Annual Review of Control, Robotics, and Autonomous Systems, 4:59–87, 2021.

Anastasiya Salova, Jeffrey Emenheiser, Adam Rupe, James P Crutchfield, and Raissa M D'Souza. Koopman operator and its approximations for systems with symmetries. Chaos: An Interdisciplinary Journal of Nonlinear Science, 29(9), 2019.

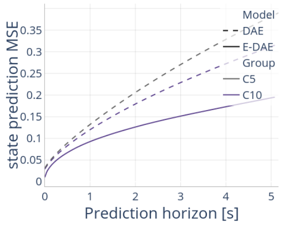

Figure 5: Linear exp prediction horizon error

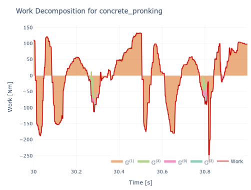

Figure 6: Pronking Work decomposition

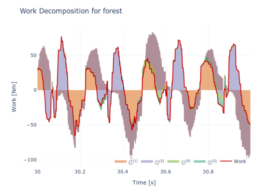

Figure 7: Trot on Forest Work decomposition

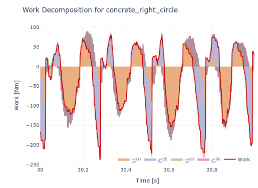

Figure 8: Trot walk on circles Work decomposition

[font in the last 3 figures is too small and the last fig is no vertically aligned to the others]

7. Default Notation

In an attempt to encourage standardized notation, we have included the notation file from the textbook, *Deep Learning* available at https://github.com/goodfeli/dlbook_notation/. Use of this style is not required and can be disabled by cosmmenting out `math_commands.tex`.

Numbers and Arrays

a	A scalar (integer or real)
$\boldsymbol{A}$	A matrix
A	A tensor
$\boldsymbol{I}_n$	Identity matrix with n rows and n columns
$\boldsymbol{I}$	Identity matrix with dimensionality implied by context
$\boldsymbol{e}^{(i)}$	Standard basis vector $[0, \dots, 0, 1, 0, \dots, 0]$ with a 1 at position i
a	A scalar random variable
$\mathbf{a}$	A vector-valued random variable
$\mathbf{A}$	A matrix-valued random variable

Sets and Graphs

$\mathbb{A}$	A set
$\mathbb{R}$	The set of real numbers
$\{0, 1\}$	The set containing 0 and 1
$\{0, 1, \dots, n\}$	The set of all integers between 0 and n
$[a, b]$	The real interval including a and b
$(a, b]$	The real interval excluding a but including b
$\mathbb{A} \backslash \mathbb{B}$	Set subtraction, i.e., the set containing the elements of $\mathbb{A}$ that are not in $\mathbb{B}$
$\mathcal{G}$	A graph
$Pa_{\mathcal{G}}(\mathrm{x}_i)$	The parents of x_i in $\mathcal{G}$

Indexing

$A_{i,j}$	Element i, j of matrix $\boldsymbol{A}$
$\boldsymbol{A}_{i,:}$	Row i of matrix $\boldsymbol{A}$
$\boldsymbol{A}_{:,i}$	Column i of matrix $\boldsymbol{A}$
$\mathsf{A}_{i,j,k}$	Element (i, j, k) of a 3-D tensor $\mathbf{\mathsf{A}}$
$\mathbf{\mathsf{A}}_{:,:,i}$	2-D slice of a 3-D tensor
a_i	Element i of the random vector $\mathbf{a}$

Calculus

$\dfrac{dy}{dx}$	Derivative of y with respect to x
$\dfrac{\partial y}{\partial x}$	Partial derivative of y with respect to x
$\nabla_{\boldsymbol{x}} y$	Gradient of y with respect to $\boldsymbol{x}$
$\nabla_{\boldsymbol{X}} y$	Matrix derivatives of y with respect to $\boldsymbol{X}$
$\nabla_{\mathbf{\mathsf{X}}} y$	Tensor containing derivatives of y with respect to $\mathbf{\mathsf{X}}$
$\dfrac{\partial f}{\partial \boldsymbol{x}}$	Jacobian matrix $\boldsymbol{J} \in \mathbb{R}^{m \times n}$ of $f : \mathbb{R}^n \to \mathbb{R}^m$
$\nabla_{\boldsymbol{x}}^2 f(\boldsymbol{x})$ or $\boldsymbol{H}(f)(\boldsymbol{x})$	The Hessian matrix of f at input point $\boldsymbol{x}$
$\int f(\boldsymbol{x}) d\boldsymbol{x}$	Definite integral over the entire domain of $\boldsymbol{x}$
$\int_{\mathbb{S}} f(\boldsymbol{x}) d\boldsymbol{x}$	Definite integral with respect to $\boldsymbol{x}$ over the set $\mathbb{S}$

Probability and Information Theory

$P(\mathrm{a})$	A probability distribution over a discrete variable
$p(\mathrm{a})$	A probability distribution over a continuous variable, or over a variable whose type has not been specified
$\mathrm{a} \sim P$	Random variable a has distribution P
$\mathbb{E}_{\mathrm{x}\sim P}[f(x)]$ or $\mathbb{E}f(x)$	Expectation of $f(x)$ with respect to $P(\mathrm{x})$
$\mathrm{Var}(f(x))$	Variance of $f(x)$ under $P(\mathrm{x})$
$\mathrm{Cov}(f(x), g(x))$	Covariance of $f(x)$ and $g(x)$ under $P(\mathrm{x})$
$H(\mathrm{x})$	Shannon entropy of the random variable x
$D_{\mathrm{KL}}(P\|Q)$	Kullback-Leibler divergence of P and Q
$\mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$	Gaussian distribution over $\boldsymbol{x}$ with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$

Functions

$f : \mathbb{A} \to \mathbb{B}$	The function f with domain $\mathbb{A}$ and range $\mathbb{B}$
$f \circ g$	Composition of the functions f and g
$f(\boldsymbol{x}; \boldsymbol{\theta})$	A function of $\boldsymbol{x}$ parametrized by $\boldsymbol{\theta}$. (Sometimes we write $f(\boldsymbol{x})$ and omit the argument $\boldsymbol{\theta}$ to lighten notation)
$\log x$	Natural logarithm of x
$\sigma(x)$	Logistic sigmoid, $\dfrac{1}{1+\exp(-x)}$
$\zeta(x)$	Softplus, $\log(1+\exp(x))$
$\|\|\boldsymbol{x}\|\|_p$	L^p norm of $\boldsymbol{x}$
$\|\|\boldsymbol{x}\|\|$	L^2 norm of $\boldsymbol{x}$
x^+	Positive part of x, i.e., $\max(0, x)$
$frm[o]--_{\text{condition}}$	is 1 if the condition is true, 0 otherwise

Group and Representation Theory

$\mathbb{G}$	A symmetry group
$\mathbb{U}(\mathcal{X})$	Unitary group on the vector space $\mathcal{X}$
$\mathbb{GL}(\mathcal{X})$	General Linear group on the vector space $\mathcal{X}$
$\mathbb{O}(\mathcal{X})$	Orthogonal group on the vector space $\mathcal{X}$
g	A symmetry group action. The symbol g is **reserved** for elements of symmetry groups. Different actions are denoted by g_1, g_n, g_*
h	A symmetry group action generator. The symbol h is reserved alone for generator actions. Different generator actions are denoted by h_1, h_n
$g_a \cdot g_b$	Binary composition operator of two group elements
$\rho_{\mathcal{X}}(g_a)$	Representation of the group element g_a on the vector space $\mathcal{X}$