Geometry as Inductive Bias in Data-Driven Robotics: Case Studies from the RobotGenSkill Project

Arno Verduyn[*,1,2], Ali Mousavi Mohammadi[1,2], Riccardo Burlizzi[1,2], Wilm Decré[1,2], Erwin Aertbeliën[1,2], Maxim Vochten[3], and Joris De Schutter[1]

***Abstract*— Data-driven machine learning approaches have achieved impressive results. However in robotics data collection is costly and limited. This paper provides a synthetic overview of case studies from the RobotGenSkill research project to show how reliance on data can be reduced and generalization improved in robot learning by exploiting geometry as a representation-level inductive bias. Specifically, we explain how equivariant geometric features and invariant task representations can be used to systematically remove task-irrelevant variability arising from changes in reference frames and time profiles, while also discussing the limitations of imposing inductive biases.**

I. Introduction

Traditional robotics has long relied on explicit task models grounded in kinematics, geometry, and dynamics. While such models offer strong interpretability, they typically require substantial task-specific expert knowledge to construct.

In contrast, a recent trend in machine learning is to reduce the dependency on manual modeling when learning tasks by leveraging *representation learning* methods that operate directly on raw or minimally processed input data [1]. Such systems have demonstrated the ability to jointly learn task-relevant features, internal representations, and mappings to downstream objectives. Sutton's "bitter lesson" [2] captures this trend, observing that methods which leverage computation and data, rather than task-specific expert knowledge, tend to achieve better performance in the long run.

However, for these representation-learning approaches to be both successful and computationally practical, a crucial factor is the incorporation of *inductive biases*. These biases are structures deliberately introduced by experts into the learning architecture. They encode prior assumptions about structural properties in the physical world, such as *symmetry*, *invariance* or *equivariance*, such that these properties do not have to be inferred solely from data. In computer vision, convolutional neural networks [3] provide an example of this principle by embedding inductive biases, such as *locality* and *translation equivariance*, into the learning architecture.

Obtaining data efficiency through inductive biases is especially important in robotics where data acquisition is often costly and time-consuming. This data-acquisition challenge in robotics limits the availability of large-scale datasets for learning. As a result, the careful introduction of inductive bias plays a particularly important role in determining the efficiency and scope of what robotic systems can learn.

[*] Corresponding author (arno.verduyn@kuleuven.be)
[1] Department of Mechanical Engineering, KU Leuven, Leuven, Belgium.
[2] Flanders Make at KU Leuven, Leuven, Belgium.
[3] Department of Mechanics, Royal Military Academy, Brussels, Belgium.
This work was supported by the European Research Council (ERC) through the European Union's Horizon 2020 research and innovation programs under grant agreement No 788298 (RobotGenSkill).

II. Geometry as Inductive Bias

We argue that *geometry* provides a powerful source of inductive bias for data-driven robotics. Specifically, by encoding geometric structure, such as invariance and equivariance, *task-agnostic inductive bias at the level of task representation* can be introduced. In Section III, we show that introducing such bias improves both data efficiency and robot skill generalization by providing a systematic overview of case studies from the *RobotGenSkill* [4] research project.

The primary objective of the RobotGenSkill project was to develop data-efficient methods that enable robots to learn tasks from a small number of human demonstrations and to generalize the learned skills to new situations. To address this objective, the project leveraged geometry to control how task representations change when the *task context* changes. In this paper, the task context encompasses both the *task configuration* (i.e., the choice and placement of world and body reference frames) and the *task execution style* (i.e., choice of motion profile). At a methodological level, the methods developed in RobotGenSkill combine:

1. Equivariant features. These features are constructed from demonstration data and grounded in geometry, such as motion and force screw axes [5, 6], average screw axes [7], moving frames defined along trajectories [8, 9], and data-derived task frames [10]. The resulting features are *equivariant* in the sense that they transform in a predictable and structured manner when the context of the task changes. In robotics, such equivariance is particularly important for action generation and control, where robot actions must remain consistent with changes in the task context.

2. Invariant representations. Building on these equivariant features, representations that are *invariant* to contextual changes are derived, meaning that their values remain unchanged when the task context changes. Such invariance is useful for factoring out task-irrelevant variability. Invariant representations are therefore well suited for task recognition, where task instances performed in different contexts should be treated as equivalent, and for task generalization, where learned tasks should easily transfer across contexts.

Using these equivariant features and invariant representations for task representation and learning results in the introduction of geometric inductive bias, which has the following important benefits:

First, it improves data efficiency. By explicitly encoding invariance and equivariance with respect to changes in the task context, learning systems are relieved from rediscovering this structure from data alone. As a result, generic task representations can be learned from fewer demonstrations. This reduction in required data also lowers computational effort during learning, contributing to reduced energy consumption and a smaller environmental footprint.

Second, the resulting invariant representations are, by design, insensitive to reference frame calibration errors, sensor placement, and other sources of contextual variability. When such invariant representations are used as inputs to learning systems, this leads to more robust and consistent task performance across different task configurations.

Third, the resulting equivariant features and invariant representations remain interpretable, as they are grounded in geometry and therefore have direct physical meaning. This stands in contrast to latent vectors learned by typical end-to-end networks. This higher interpretability makes it easier to understand system behavior and to adjust the corresponding task-level control parameters when needed.

However, while these benefits are significant, there is an important limitation. The effectiveness of the proposed inductive bias depends on which invariances and equivariances are chosen, as imposing inappropriate invariances can remove task-relevant information or increase noise sensitivity. For example, removing information about the placement of reference frames can be detrimental in recognition tasks when absolute spatial relationships are discriminative, such as distinguishing tasks performed at different locations. As another example, introducing locality increases sensitivity to small local variations in the input data, such as sensor noise.

Nevertheless, the above limitation (the possibility of introducing an inappropriate bias) reveals a direction for future work: rather than committing to a single bias a priori, multiple representations with different geometric inductive biases could be combined (e.g., through mixture-of-experts approaches [11]) allowing data-driven mechanisms to select or weight the most appropriate bias for a given task.

III. RobotGenSkill Case Studies

This section illustrates the use of geometric inductive bias via case studies from the RobotGenSkill project.

A. Development of Invariant Trajectory Representations

A central line of work within RobotGenSkill focused on the development of invariant representations of both rigid-body motion and contact-wrench trajectories [9, 12] as a front-end for learning from demonstration methods.

Equivariant features: Rather than operating directly on raw trajectory coordinates, we construct geometric features from the demonstration data that transform equivariantly under changes in the context. These features include instantaneous motion and force screw axes, as well as moving reference frames defined functionally and locally along the corresponding motion and contact-wrench trajectories.

Invariant representations: A local invariant trajectory representation is then obtained by expressing the local evolution of the trajectory in its corresponding equivariant frame. The work in [9] introduced such invariant representations for rigid-body motion trajectories and contact-wrench trajectories. These invariant representations were derived from differential-geometric properties, similarly to the derivation of the Frenet-Serret frame [13, 14] and its associated invariants for translation trajectories. More recent work [12] proposes a novel Dual Upper-Triangular Invariant Representation (DUTIR), which exhibits improved robustness to measurement noise and to representation singularities.

Applications and benefits: These invariant representations can be used as inputs to algorithms operating on demonstration data. By removing the dependency on the task context at the representation level, the proposed approach reduces the amount of data required for learning skills that generalize well across task contexts. Another practical benefit is that the obtained invariance eliminates the need for reference frame calibration, and enables learning from demonstrations collected in inconsistent or unknown reference frames.

B. Robust Segmentation of Motion Trajectories

Within RobotGenSkill, geometric structure is also exploited to enable robust segmentation of rigid-body trajectories into meaningful sub-motions [15], as illustrated in Fig. 1.

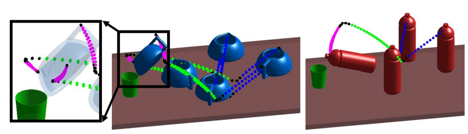

Fig. 1. Robust trajectory segmentation of pouring motions [15]. Points along the trajectories with identical colors (magenta, green, or blue) correspond to the same trajectory segment. Thanks to its invariant properties, the approach has consistent results for different choices of the body reference frame (left) and is transferable to a novel pouring object (right).

Equivariant features: The method is based on constructing the instantaneous motion screw axis (ISA) along the rigid-body motion trajectory (see also Section III-A).

Invariant representations: The rigid body motion is then expressed in terms of rotations and translations along the ISA. Based on this concept, a novel geometric progress parameter is defined that can be used to express the progression along the trajectory in a purely geometric manner, invariantly to changes in the task context.

Applications and benefits: The work in [15] demonstrates improved consistency in trajectory segmentation across different task contexts when using the proposed geometric progress parameter.

C. Robust Recognition of Motion Trajectories

Within RobotGenSkill, geometric structure is further exploited to enable robust motion trajectory recognition across task contexts. The application in [16] focused on the recognition of demonstrated object-manipulation tasks in free space. The application in [17] focused on hand palm motion gesture recognition. A snapshot of the latter is shown in Fig. 2.

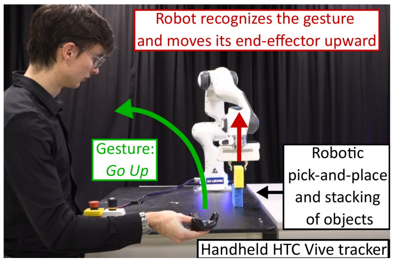

Fig. 2. Robot setup for hand palm motion trajectory recognition [17].

Equivariant features and invariant representations: In both [16] and [17], the invariant representation of the performed motion is obtained by first expressing the motion trajectory as a function of a geometric progress parameter (see Section III-B), and subsequently expressing the local evolution of the trajectory in the corresponding equivariant reference frame (see Section III-A).

Applications and benefits: The resulting invariant representation, together with a novel invariant trajectory similarity measure [16], enable the robust comparison and recognition of motion trajectories across task contexts.

D. Efficient Data Augmentation for Trajectory Learning

Within RobotGenSkill, geometric structure is also exploited to improve data efficiency in robot learning from demonstration. This is done specifically by augmenting the available dataset for training with new synthetic training instances [18], as illustrated in Fig. 3.

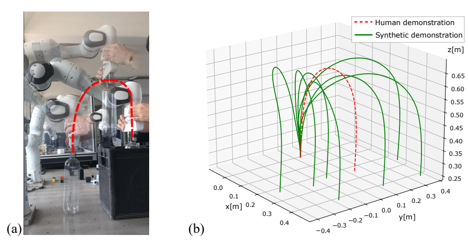

Fig. 3. Data augmentation framework [18] applied to a crate filling task: (a) a single human demonstration (in red); (b) synthetic demonstrations (in green) generated towards new targets.

Invariant representations: Synthetic demonstrations are generated using invariant representations of the demonstrated tasks (see Section III-A). These invariant representations are used to define equivalence classes of motions corresponding to the same underlying skill. Synthetic demonstrations are then generated by instantiating new task executions in different contexts while minimizing deviations from the task's invariant representation.

Applications and benefits: By sampling synthetic demonstrations from these equivalence classes, augmented datasets can be constructed that extend well beyond the contextual variability present in the original demonstrations. Probabilistic motion models trained on these augmented datasets exhibit improved extrapolation capabilities, while significantly reducing the number of required human demonstrations [18].

E. Geometric Derivation of Optimal Task Frames for Contact-Rich Manipulation Tasks

Lastly, geometric structure is also exploited to support learning, control, and generalization of contact-rich manipulation tasks. This is done specifically through the data-driven derivation of task-relevant reference frames [10], with representative tasks shown in Fig. 4.

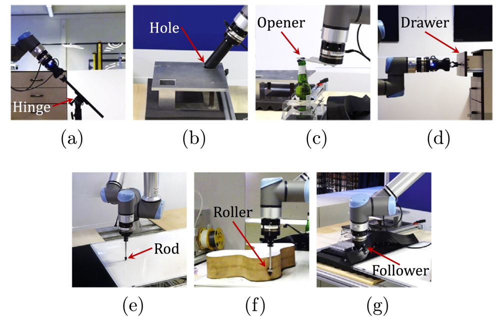

Fig. 4. UR10e robot executing a range of contact-rich manipulation tasks learned from human demonstrations [10]: (a) revolute joint, (b) peg-in-hole alignment, (c) bottle opening, (d) prismatic joint, (e) drawing, (f) 2-D contour following, and (g) 3-D contour following.

Equivariant features: For each task, a suitable task frame is automatically derived from demonstrated motion and force data using a generic method grounded in screw theory. This frame maximizes the decoupling between the motion and force components in the task, enabling the design of a decoupled task-space controller.

Invariant representations: By expressing control objectives in the learned task frame, the resulting representation of the task becomes invariant to changes in reference frames, sensor placement, etc. This representation hence allows learning and control policies to focus on task-relevant structure rather than incidental contextual variability.

Applications and benefits: Controllers designed in these data-derived task frames become simple, with the decoupling between motion and force control maximized. These controllers contain only a few, easily tunable parameters which easily generalize across different task contexts. Compared to optimization-based frame-selection methods [19], the proposed analytical approach is significantly faster and more robust, while also quantifying uncertainty, thereby enabling uncertainty-aware task control.

IV. Conclusion

This paper positions geometry as a powerful source of representation-level inductive bias, enabling data-efficient and generalizable data-driven robotics. Through case studies from the RobotGenSkill project, we show how geometric inductive bias (encoded via equivariance and invariance to changes in task context) can be effectively introduced. These case studies demonstrate data-efficient and robust learning, segmentation, recognition, control, and generalization of robot skills across task contexts.

References

[1] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[2] R. S. Sutton, "The bitter lesson," http://www.incompleteideas.net/IncIdeas/BitterLesson.html, 2019, accessed: 2026-04-07.

[3] P. Wang, W. Li, Z. Gao, J. Zhang, C. Tang, and P. O. Ogunbona, "Action recognition from depth maps using deep convolutional neural networks," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 4, pp. 498–509, 2016.

[4] J. De Schutter *et al.*, "RobotGenSkill: Generalizing Human-Demonstrated Robot Skills," https://robotgenskill.pages.gitlab.kuleuven.be.

[5] S. Stramigioli and H. Bruyninckx, "Geometry and screw theory for robotics," 2001, tutorial presented at the IEEE International Conference on Robotics and Automation (ICRA 2001).

[6] J. De Schutter, "Invariant description of rigid body motion trajectories," *ASME Journal of Mechanisms and Robotics*, vol. 2, no. 1, 2010.

[7] A. Ancillao, M. Vochten, A. Verduyn, J. De Schutter, and E. Aertbeliën, "An optimal method for calculating an average screw axis for a joint, with improved sensitivity to noise and providing an analysis of the dispersion of the instantaneous axes," *PLOS ONE*, vol. 17, no. 10, pp. 1–22, 2022.

[8] E. Cartan, *La Théorie Des Groupes Finis Et Continus Et La Géométrie Différentielle, Traitées Par La Méthode Du Repère Mobile*. Paris, France: Gauthier-Villars, 1937.

[9] M. Vochten, A. Mousavi Mohammadi, A. Verduyn, T. De Laet, E. Aertbeliën, and J. De Schutter, "Invariant descriptors of motion and force trajectories for interpreting object manipulation tasks in contact," *IEEE Transactions on Robotics*, pp. 1–21, 2023.

[10] A. Mousavi Mohammadi, M. Vochten, E. Aertbeliën, and J. De Schutter, "A generic task model and control strategy to support learning, robust control, and generalization of contact-rich manipulation tasks," *Robotics and Autonomous Systems*, vol. 197, p. 105270, 2026.

[11] S. Masoudnia and R. Ebrahimpour, "Mixture of experts: a literature survey," *Artificial Intelligence Review*, vol. 42, no. 2, pp. 275–293, 2014.

[12] A. Verduyn, E. Aertbeliën, M. Vochten, and J. D. Schutter, "A coordinate-invariant local representation of motion and force trajectories for identification and generalization across coordinate systems," 2026. [Online]. Available: https://arxiv.org/abs/2604.10241

[13] F. Frenet, "Sur les courbes à double courbure," *Journal de Mathématiques Pures et Appliquées*, vol. 17, no. 1, pp. 437–447, 1852.

[14] J. Serret, "Mémoire sur une classe d'équations différentielles simultanées qui se rattachent à la théorie des courbes à double courbure," *Journal de Mathématiques Pures et Appliquées*, vol. 18, no. 1, pp. 1–40, 1853.

[15] A. Verduyn, M. Vochten, and J. De Schutter, "Enhancing motion trajectory segmentation of rigid bodies using a novel screw-based trajectory-shape representation," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 7179–7185.

[16] A. Verduyn, E. Aertbeliën, G. Maes, J. D. Schutter, and M. Vochten, "Bilts: A bi-invariant similarity measure for robust object trajectory recognition under reference frame variations," 2025. [Online]. Available: https://arxiv.org/abs/2405.04392

[17] A. Verduyn, M. Vochten, and J. De Schutter, "Enhancing hand palm motion gesture recognition by eliminating reference frame bias via frame-invariant similarity measures," in *2025 IEEE 21st International Conference on Automation Science and Engineering (CASE)*, 2025, pp. 866–873.

[18] R. Burlizzi, M. Vochten, J. De Schutter, and E. Aertbeliën, "Extending extrapolation capabilities of probabilistic motion models learned from human demonstrations using shape-preserving virtual demonstrations," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 10 772–10 779.

[19] A. H. G. Overbeek, D. Dresscher, H. van der Kooij, and M. Vlutters, "Versatile kinematics-based constraint identification applied to robot task reproduction," *Frontiers in Robotics and AI*, vol. 12, p. 1574110, 2025.