# Learning Diffeomorphic Lyapunov Functions from Data

Samuel Tesfazgi<sup>\*1</sup> Leonhard Sprandl<sup>\*1</sup> Sandra Hirche<sup>1</sup>

# Abstract

The practical deployment of learning-based autonomous systems would greatly benefit from tools that flexibly obtain safety guarantees in the form of certificate functions from data. While the geometrical properties of such certificate functions are well understood, synthesizing them using machine learning techniques remains a challenge. To mitigate this issue, we propose a diffeomorphic function learning framework where prior structural knowledge regarding the desired output is encoded in a simple surrogate function, which is subsequently augmented through an expressive, topology-preserving state-space transformation. We demonstrate our approach by learning Lyapunov functions from real-world data and apply the method to different attractor systems.

# 1. Introduction

With recent advances in robotics and machine learning, datadriven autonomous systems are increasingly deployed in safety-critical application scenarios such as autonomous driving (Liu et al., 2024) or robotic rehabilitation (Ai et al., 2023). While learning-based systems are particularly wellsuited for such complex and uncertain environments, a limitation that inhibits their deployment is the lack of formal safety and stability guarantees. A practical method to ascertain the desired safety and stability properties of a dynamical system is through the construction of certificate functions, e.g., a Lyapunov function to show convergence to an equilibrium point (Khalil, 2002). A major strength of these approaches is the existence of converse theorems, i.e., if the desired property holds, the certificate functions are guaranteed to exist (Teel et al., 2014; Liu, 2022) and they may additionally be used for control synthesis (Sontag, 1989; Tesfazgi et al., 2024).

In general, certificate functions express the long-term behavior of a system's trajectory through invariant set constraints. Thereby, the set of states to which the system is bounded or converges to, is geometrically encoded in the level sets of the certificate function. A candidate Lyapunov function for instance, has to be positive definite with a strictly decreasing time-derivative. While these conditions can be resolved efficiently in simple settings, e.g., when the dynamics are known analytically and the hypothesis space is limited to sum-of-squares polynomials (Parrilo, 2000), no constructive approach is known for general, nonlinear systems. Therefore, the need for expressive learning techniques that construct certificate functions directly from data arises.

Recently, the deployment of neural networks (NNs) has been proposed to learn Lyapunov functions from observations (Richards et al., 2018; Ravanbakhsh & Sankaranarayanan, 2019; Chang et al., 2019; Manek & Kolter, 2019). However, even though NNs have the advantage of strong representational capabilities, imposing the necessary constraints efficiently is an open issue. Existing methods either induce the Lyapunov conditions via soft-constraints (Chang et al., 2019), only admitting empirical statements, or strictly by extensively searching for counter-examples (Ravanbakhsh & Sankaranarayanan, 2019), which is computationally demanding. A promising perspective has been to geometrically constrain the output of the NN by using a suitable architecture. However, while such constraints have been shown to be beneficial in the context of partial differential equations (Raissi et al., 2019), when learning Lyapunov functions the imposed output constraints are either not specific enough, only guaranteeing positive definiteness (Richards et al., 2018), or overly conservative (Manek & Kolter, 2019), e.g., using input convex NNs (Amos et al., 2017).

In this work, we follow an alternative approach of encoding structural knowledge and imposing desired geometric properties on the inferred function by deploying invertible models. In particular, instead of constraining the output of a function approximator directly, we specify a simple base function with desired geometric properties and subsequently learn a topology-preserving, state-space transformations under which the augmented base function adheres to the data, thereby indirectly obtaining a Lyapunov function. While the regularity preserving properties of smooth and bijective maps, so called diffeomorphisms (Boumal, 2023), have been

<sup>\*</sup>Equal contribution <sup>1</sup>Chair of Information-oriented Control (ITR), Technical University of Munich (TUM), Munich, Germany. Correspondence to: Samuel Tesfazgi <samuel.tesfazgi@tum.de>.

Accepted as an extended abstract for the Geometry-grounded Representation Learning and Generative Modeling Workshop at the 41<sup>st</sup> International Conference on Machine Learning, ICML 2024, Vienna, Austria. Copyright 2024 by the author(s).

used in the context of imitation learning (Rana et al., 2020) and control (Sun et al., 2023), their utilization for learning certificate functions remains understudied. Beyond the certification of point attractors, we demonstrate the applicability of the proposed approach for more general system classes including multiple equilibria and limit cycles.

## 2. Preliminaries

Lyapunov stability theory. Consider an autonomous system

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}),\tag{1}$$

with continuous state  $x \in \mathbb{R}^n$  and system dynamics  $f : \mathbb{R}^n \to \mathbb{R}^n$ . The problem of certifying stability is concerned with analyzing the behavior of x(t) for  $t \to \infty$ , given some initial state  $x(t_0) = x_0$ . In order to formalize this property, we introduce the following concept of stability.

**Definition 2.1** ((Khalil, 2002)). A system (1) has an asymptotically stable equilibrium  $x^*$  on the set  $\mathcal{X}$  if

- 1. for all d > 0, there exist  $\delta > 0$ ,  $t_0 \ge 0$  such that  $\|\boldsymbol{x}_0 \boldsymbol{x}^*\| < \delta$  implies  $\|\boldsymbol{x}(t) \boldsymbol{x}^*\| < d$ ,  $\forall t \ge t_0$ .
- 2.  $\lim_{t\to\infty} \|\boldsymbol{x}(t) \boldsymbol{x}^*\| = 0$  for all  $\boldsymbol{x}_0 \in \mathcal{X}$ .

If the conditions hold for all states, i.e.,  $x_0 \in \mathbb{R}^n$ , the equilibrium  $x^*$  is globally asymptotically stable. Without loss of generality, we assume  $x^* = 0$  from now on. A practical method to ascertain the convergence property of a system, without solving the underlying dynamics equations, is by means of Lyapunov stability theory.

**Theorem 2.2** (Lyapunov Stability Theorem, (Khalil, 2002)). Let  $\mathbf{x}^* = \mathbf{0}$  be an equilibrium point for (1) and  $\mathcal{X} \subset \mathbb{R}^n$  be the domain of  $f : \mathcal{X} \mapsto \mathbb{R}^n$  with  $\mathbf{x}^* \in \mathcal{X}$ . Let  $V : \mathcal{X} \mapsto \mathbb{R}$  be a continuously differentiable function such that:

$$V(\mathbf{0}) = 0 \tag{2a}$$

$$V(\boldsymbol{x}) > 0 \quad \forall \boldsymbol{x} \in \mathcal{X} \setminus \{\boldsymbol{0}\}$$
 (2b)

$$\dot{V}(\boldsymbol{x}) = \nabla_{\boldsymbol{x}}^{\mathsf{T}} V(\boldsymbol{x}) f(\boldsymbol{x}) < 0 \quad \forall \boldsymbol{x} \in \mathcal{X} \setminus \{\boldsymbol{0}\}$$
(2c)

#### *Then,* $x^*$ *is locally asympt. stable in the sense of Def. 2.1.*

Thus, finding a function  $V(\cdot)$  that satisfies (2a)-(2c) is sufficient to certify stability of  $f(\cdot)$ . Beyond asymptotic stability, a dynamical system may also exhibit other types of attractor landscapes, such as multiple equilibria, where system trajectories converge to different states out of a set  $\mathcal{X}^* := \{ x \in \mathbb{R}^n \mid f(x) = \mathbf{0} \}$  depending on  $x_0$ , or limit cycles, which describe invariant sets  $\mathcal{X}^\circ$  under the dynamics f for some orbital period T. In order to extend the notion of Lyapunov stability analysis to such systems, it is common to introduce a Lyapunov-like function (Patrão, 2011; Björnsson et al., 2015) that satisfies the conditions (2a)-(2c) for the respective sets  $\mathcal{X}^*$  or  $\mathcal{X}^\circ$ , instead of only  $\{\mathbf{0}\}$ .



(a) Original states- (b) Invertible map (c) Non-invertible pace x map

*Figure 1.* Transformations of a 2D space (a) by an invertible (b) and a non-invertible mapping (c). Invertability requires that the mapping for each point is unique, i.e., no crossings.

**Diffeomorphism.** A mapping  $\phi : \mathbb{R}^n \mapsto \mathbb{R}^n$  is bijective, if it's inverse  $\phi^{-1}$  is guaranteed to exist. If the mapping  $\phi$  and its inverse  $\phi^{-1}$  are further smooth, it is referred to as a diffeomorphism (Boumal, 2023). We denote the set of diffeomorphic maps with  $\phi \in \mathcal{D}$ . The requirement of  $\phi$  being smooth allows the mapping between two differentiable manifolds. Since a differentiable manifold is additionally equipped with a differential structure (Lee, 2012), it gives rise to the tangent space required to define gradients, which are necessary for any gradient-based analysis framework, such as Lyapunov stability analysis. Conveniently, diffeomorphic maps preserve the topology of objects, such as functions or differential equations. Intuitively, two sets  $U \subset \mathbb{R}^n$  and  $W \subset \mathbb{R}^n$  are topologically equivalent, if a mapping between the two can be established, with the map and its inverse being continuous (Lee, 2000). Figure 1 illustratively depicts the difference between a topologypreserving and a non-topology-preserving transformation.

### 3. Diffeomorphic Lyapunov Functions

Directly searching for a Lyapunov function is difficult, since constraints (2b) and (2c) need to hold for an uncountable, infinite set of states. To overcome this, we propose to exploit the topological-equivalence of Lyapunov functions (Grüne et al., 1999) to reformulate the function approximation problem to an optimization over state-space transformations.

#### 3.1. Lyapunov function hypothesis space

The primary challenge in synthesizing a Lyapunov function is guaranteeing that the gradient  $\nabla_x V$  fulfills the descent condition (2c). Typically, the dynamics f are not known analytically. Thus, in our considered scenario we only have access to trajectory samples  $\{x_i, \dot{x}_i\}_{i=1}^N$ . However, we may still derive shape-constraints that any potential Lyapunov function candidate has to adhere to locally, since

$$\nabla_x^{\mathsf{T}} V(\boldsymbol{x}) f(\boldsymbol{x}) < 0 \Longrightarrow \nabla_x V(\boldsymbol{x}) \neq \boldsymbol{0}, \ \forall \boldsymbol{x} \in \mathcal{X} \setminus \{\boldsymbol{0}\}.$$
 (3)

Consequently, a positive definite function V with nonvanishing gradient  $\nabla_x V(x) \neq 0$  is always a valid Lyapunov function for *some* system. Therefore, a diffeomorphic transformation that preserves these topological properties is guaranteed to generate an output that remains in the space of Lyapunov functions, which we demonstrate in the following:

**Proposition 3.1.** Consider a smooth function  $V : \mathbb{R}^n \mapsto \mathbb{R}$ and let  $N_V$  denote the number of unique gradient roots of V

$$N_V = |\mathcal{S}_V|, \quad \text{with } \mathcal{S}_V = \{ \boldsymbol{x} | \nabla_{\boldsymbol{x}} V(\boldsymbol{x}) = \boldsymbol{0} \}$$
 (4)

where |S| denotes the cardinality of the set S. Next, let  $\phi : \mathbb{R}^n \mapsto \mathbb{R}^n$  be an orientation-preserving diffeomorphism, *i.e.*, its Jacobian  $J_{\phi} \in \mathbb{R}^{n \times n}$  satisfies

$$\det(\boldsymbol{J}_{\phi}(\boldsymbol{x})) > 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^{n}.$$
(5)

Then, the number of gradient roots remains unchanged by the map  $\phi$  and we have  $N_V = N_W$ , where  $W := V \circ \phi$ .

*Proof.* Due to (5),  $J_{\phi}$  is full rank  $\forall x \in \mathbb{R}^n$ , and consequently, the nullspace null( $J_{\phi}(x)$ ) only contains the trivial solution (Strang, 2019). The same holds for the transpose, since det( $J_{\phi}(x)$ ) = det( $J_{\phi}(x)^{\top}$ ) > 0. Thus, the *left* nullspace (Strang, 2019) also only contains the trivial solution. Applying the chain rule, the gradient of W yields

$$\nabla_{\boldsymbol{x}} W(\boldsymbol{x}) = \frac{\partial}{\partial \boldsymbol{x}} V(\phi(\boldsymbol{x})) = \boldsymbol{J}_{\phi}(\boldsymbol{x})^{\top} \nabla_{\boldsymbol{x}} V(\boldsymbol{x}). \quad (6)$$

From (5) and (6), it trivially follows that  $|S_V| = |S_W|$ , which concludes the proof.

Thus, it follows that candidate Lyapunov functions are diffeomorphic to one another and consequently that any Lyapunov functions can be transformed into a simple  $\mathcal{K}_{\infty}$  function under a change of coordinates, as proposed in (Grüne et al., 1999). This can be seen intuitively in Figure 2. For a single point attractor system, the time derivative of a valid Lyapunov function V has to decrease along the trajectories, thereby necessitating non-vanishing gradients outside of the equilibrium (3). Therefore, each contour line of any V is topologically equivalent to a sphere, hence, admitting a diffeomorphic transformation to one another.

#### 3.2. Reformulation as diffeomorphic learning problem

Based on the previous section, we search over the space of Lyapunov functions by finding an appropriate diffeomorphic transformation without the need to explicitly incorporate shape constraints. The data-driven, diffeomorphic Lyapunov learning problem is formalized as follows:

**Definition 3.2.** Given a dataset  $D = \{x_i, \dot{x}_i\}_{i=1}^N$  generated by an unknown stable system  $\dot{x} = f(x)$  with f(0) = 0 and any initial Lyapunov-like function  $V_b$  with

$$V_b(\boldsymbol{x}) > 0 \land \nabla_x V_b(\boldsymbol{x}) \neq \boldsymbol{0}, \quad \forall \boldsymbol{x} \in \mathcal{X} \setminus \{\boldsymbol{0}\}, \quad (7)$$



Figure 2. Contour plots for different Lyapunov function candidates  $(V_1, V_2 \text{ and } V_3)$  for single attractor systems. Highlighted in red are the contour lines for a specific value of V. Note that all three red contour lines can be continuously deformed into each other.

find a diffeomorphism

$$\phi^* = \operatorname*{argmin}_{\phi \in \mathcal{D}} l(V_\phi) \tag{8a}$$

s.t. 
$$\nabla_{\boldsymbol{x}}^{\mathsf{T}} V_{\phi}(\boldsymbol{x}_i) \dot{\boldsymbol{x}}_i < 0 \quad \forall i \in [1 \dots N]$$
 (8b)

$$\phi(\mathbf{0}) = \mathbf{0} \tag{8c}$$

where  $V_{\phi}(\boldsymbol{x}) = V_b \circ \phi(\boldsymbol{x})$  and l is a loss function of choice.

Intuitively, with (8) we reformulate the search for a Lyapunov function, which is a functional optimization problem, as a diffeomorphic optimization problem. First, a base function  $V_b$  is specified that adheres to the topological properties of a Lyapunov candidate function (7). Then a diffeomorphism is found such that the function under the diffeomorphic transformation  $V_{\phi}$  satisfies the Lyapunov conditions on the samples (8b). This is convenient, since the geometric properties of a Lyapunov function are well known, and therefore, the surrogate function can be trivially specified, e.g., to  $V_b(x) = x^{\mathsf{T}} x$ . The distinct advantage of encoding geometric knowledge through a base function  $V_b$  becomes even more apparent when considering more general system classes with different attractor landscapes. Typical function approximation approaches, do not readily extend to more involved attractor landscapes, since the new Lyapunov-like function requires different geometric constraints. On the other hand, our proposed diffeomorphic learning framework merely requires an appropriate base function  $V_b$ , that encodes the topology of the desired attractor landscapes.

### 4. Evaluation

To evaluate the diffeomorphic Lyapunov function, we apply it to systems with different attractor landscapes. While many diffeomorphism constructions rely on fixed dimension splitting to guarantee bijectivity and obtain an analytical inverse (Dinh et al., 2016; Kingma & Dhariwal, 2018), the resulting triangular Jacobian structure is less flexible, which can be unfavorable in a constrained, data-driven setting. Therefore, we deploy a diffeomorphism construction that admits a freeform Jacobian (Chen et al., 2018; Behrmann et al., 2019)



*Figure 3.* Diffeomorphically learned Lyapunov functions for 3 exemplary shapes from the LASA handwriting dataset. The demonstrated trajectories are shown in black and the samples on which the conditions (8b) and (8c) are satisfied are shown in green

in the following evaluation. In particular, a custom kernelbased approach inspired by a residual expansion structure similar to (Perrin & Schlehuber-Caissier, 2016) is used.

**Single Equilibrium System.** For this evaluation we use the LASA handwriting dataset (Khansari-Zadeh & Billard, 2011) which is a popular benchmark in the learning stable dynamical system literature (Perrin & Schlehuber-Caissier, 2016; Rana et al., 2020; Zhang et al., 2023). Figure 3 shows the result of applying the diffeomorphic learning approach to 3 exemplary shapes with an initial guess  $V_b(x) = 0.1x^{\top}x$ . It can be seen that the proposed approach successfully learns a diffeomorphism such that the transformed function  $V_{\phi}$ constitutes a valid Lyapunov function for the trajectory data.

**Two Attractor System.** Additionally, we deploy the approach on a dynamical system with two stable equilibria and one unstable equilibrium as depicted by the vector field in Figure 4 (left). For the training data, the system is initialized at 6 different positions and simulated to convergence. The resulting  $V_{\phi}$  after applying the learned diffeomorphic transformation is depicted in Figure 4 (bottom). It is apparent that the identified diffeomorphic Lyapunov function is consistent with all demonstrations included in the dataset and successfully identified the position of the two stable equilibria.

**Limit Cycle System.** Finally, we evaluate the proposed approach by finding a Lyapunov-like function for a system with a stable limit cycle. To this end, we consider the well-known Van der Pol oscillator depicted in Figure 5 (left) with the stable limit cycle highlighted in red. For training, we sample 20 approximately equally spaced data points along the limit cycle and simulate 4 trajectories starting in the corners of the state space. The sampled data and the resulting diffeomorphic function  $V_{\phi}$  are shown in Figure 5 (right). It can be seen that the zero gradient contour line of  $V_{\phi}$  (marked in red) aligns well with the data sampled along the limit cycle (blue dots). Additionally, the Lyapunov constraints are fulfilled along the trajectories sampled outside



*Figure 4.* Diffeomorphically learned Lyapunov-like functions for a system with multiple equilibria. (Left) Vector field of the dynamical system in black and data samples in blue. (Bottom) Learned diffeomorphic Lyapunov-like function.



*Figure 5.* (Left) Vector field of the dynamical system with limit cycle marked in red. (Right) Contour lines of the learned diffeomorphic Lyapunov-like function together with sampled data.

the limit cycle as indicated by the green data points.

# 5. Conclusion

In this work, we demonstrate the use of diffeomorphism to learn Lyapunov functions from data. By selecting an appropriate base function and optimizing over topologypreserving maps, prior geometrical knowledge is encoded. We evaluate the approach on system with different attractor landscapes including multiple equilibria and limit cycles.

### References

- Ai, Q., Liu, Z., Meng, W., Liu, Q., and Xie, S. Q. Machine learning in robot-assisted upper limb rehabilitation: A focused review. *IEEE Transactions on Cognitive and Developmental Systems*, 15(4):2053–2063, 2023.
- Amos, B., Xu, L., and Kolter, J. Z. Input convex neural networks. In *Proceedings of the 34th International Conference on Machine Learning*, pp. 146–155, 2017.
- Behrmann, J., Grathwohl, W., Chen, R. T., Duvenaud, D., and Jacobsen, J.-H. Invertible residual networks. *Proceed*ings of the 36th International Conference on Machine Learning, pp. 573–582, 2019.
- Björnsson, J., Giesl, P., Hafstein, S. F., and Kellett, C. M. Computation of lyapunov functions for systems with multiple local attractors. *Discrete and Continuous Dynamical Systems*, 35, 2015.
- Boumal, N. An introduction to optimization on smooth manifolds. Cambridge University Press, 2023.
- Chang, Y.-C., Roohi, N., and Gao, S. Neural lyapunov control. In *Advances in Neural Information Processing Systems*, 2019.
- Chen, R. T., Rubanova, Y., Bettencourt, J., and Duvenaud, D. K. Neural ordinary differential equations. *Advances* in Neural Information Processing Systems, 31, 2018.
- Dinh, L., Sohl-Dickstein, J., and Bengio, S. Density estimation using real nvp. *International Conference on Learning Representations*, 2016.
- Grüne, L., Sontag, E. D., and Wirth, F. R. Asymptotic stability equals exponential stability, and iss equals finite energy gain—if you twist your eyes. *Systems & Control Letters*, 38(2):127–134, 1999.
- Khalil, H. K. *Nonlinear systems; 3rd ed.* Prentice-Hall, Upper Saddle River, NJ, 2002.
- Khansari-Zadeh, S. M. and Billard, A. Learning stable nonlinear dynamical systems with gaussian mixture models. *IEEE Transactions on Robotics*, 27(5):943–957, 2011.
- Kingma, D. P. and Dhariwal, P. Glow: Generative flow with invertible 1x1 convolutions. In *Advances in Neural Information Processing Systems*, 2018.
- Lee, J. M. *Introduction to Topological Manifolds*, volume 2. Springer, 2000. ISBN 0387987592.
- Lee, J. M. Smooth manifolds. Springer, 2012.

- Liu, H., Huang, Z., Mo, X., and Lv, C. Augmenting reinforcement learning with transformer-based scene representation learning for decision-making of autonomous driving. *IEEE Transactions on Intelligent Vehicles*, 2024.
- Liu, J. Converse barrier functions via lyapunov functions. *IEEE Transactions on Automatic Control*, 67(1):497–503, 2022. doi: 10.1109/TAC.2021.3085419.
- Manek, G. and Kolter, J. Z. Learning stable deep dynamics models. In *Advances in Neural Information Processing Systems*, 2019.
- Parrilo, P. A. Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization. California Institute of Technology, 2000.
- Patrão, M. Existence of complete lyapunov functions for semiflows on separable metric spaces. *Far East Journal of Dynamical Systems*, 2011.
- Perrin, N. and Schlehuber-Caissier, P. Fast diffeomorphic matching to learn globally asymptotically stable nonlinear dynamical systems. *Systems & Control Letters*, 96:51–59, 2016.
- Raissi, M., Perdikaris, P., and Karniadakis, G. Physicsinformed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 2019.
- Rana, M. A., Li, A., Fox, D., Boots, B., Ramos, F., and Ratliff, N. Euclideanizing flows: Diffeomorphic reduction for learning stable dynamical systems. In *Proceedings of the 2nd Conference on Learning for Dynamics and Control*, volume 120, pp. 630–639. PMLR, 2020.
- Ravanbakhsh, H. and Sankaranarayanan, S. Learning control lyapunov functions from counterexamples and demonstrations. *Auton. Robots*, 2019.
- Richards, S. M., Berkenkamp, F., and Krause, A. The Lyapunov Neural Network: Adaptive Stability Certification for Safe Learning of Dynamical Systems. In *Conference* on Robot Learning, 2018.
- Sontag, E. D. A 'universal' construction of artstein's theorem on nonlinear stabilization. *Systems Control Letters*, 13(2):117–123, 1989.
- Strang, G. *Linear algebra and learning from data*. SIAM, 2019.
- Sun, Q., Wang, X., Yang, G., Chen, Y.-H., and Ma, F. Adaptive robust formation control of connected and autonomous vehicle swarm system based on constraint following. *IEEE Transactions on Cybernetics*, 53(7):4189– 4203, 2023.

- Teel, A. R., Hespanha, J. P., and Subbaraman, A. A converse lyapunov theorem and robustness for asymptotic stability in probability. *IEEE Transactions on Automatic Control*, 59(9):2426–2441, 2014.
- Tesfazgi, S., Sprandl, L., Lederer, A., and Hirche, S. Stable inverse reinforcement learning: Policies from control lyapunov landscapes, 2024. URL https://arxiv. org/abs/2405.08756.
- Zhang, J., Beik-Mohammadi, H., and Rozo, L. Learning riemannian stable dynamical systems via diffeomorphisms. In 6th Annual Conference on Robot Learning, volume 205, pp. 1211 – 1221, 2023.