Discovering Bäcklund Transformations with PDE Foundation Models

Anonymous authors
Paper under double-blind review

Abstract

We propose a novel application of Foundation Models trained on multi–Partial-Differential-Equation data. Leveraging the vector embeddings learnt by one such model, we discuss a necessary condition for the existence of a Bäcklund tranformation between any pairs of Partial Differential Equations in its training dataset, which can be used when certain requirements on the dimension of the embedding space M and the size of the training datasets N are satisfied. In this case, the condition assumes a simple linear form and its computation scales no faster than $O((MN)^3)$.

1 Introduction

Partial Differential Equations (PDEs) are the latest addition to a long list of domains that Foundation Models (FMs) have tackled over the past few years. Whilst similar in form to other learning tasks and modalities, data governed by PDEs exhibits much more structured information and, in particular, a degree of redundancy due to the conservation laws and other physical principles that are known to apply. These constraints are not always straightforward to discover or to formulate explicitly, but they are present and can be leveraged to learn more efficiently and accurately from data.

A number of Foundation Models, trained on PDE data from a vast class of physical problems, have appeared recently (Cao, 2021; Li et al., 2023; Herde et al., 2024; Wu et al., 2024; Holzschuh et al., 2025; McCabe et al., 2025; Menon et al., 2025; Nguyen et al., 2025; Park, 2025; Rautela et al., 2025; Wiesner et al., 2025; Zhou et al., 2025). Proposals to constrain what goes under the name of Foundation Model in this specific area have also been put forward (Menon et al., 2025; Choi et al., 2025). Following the usual FM framework, these models have then been proved to generalise to other PDE-based problems, so long as they are fine-tuned to a minimal amount of data from such cases. The applications of such tools are broad: from general-purpose emulation frameworks, capable of generating solutions of arbitrary PDEs with only a few known data points, to forecast tools for systems whose governing PDEs are unknown.

In this paper, we suggest an additional use of PDE FMs. As these models encode data from all the problems contained in their training datasets onto a unified embedding space, they may be ideally positioned to provide information about the overlap of information from different PDEs and their families of solutions. In other words, it may be possible to use the data representation in this space to discover any Bäcklund transformations between pairs of PDEs in the training dataset.

In Section 2, we introduce Bäcklund transformations and their possible uses. In Section 3, we provide a necessary condition for the existence of a BT between a pair of PDEs. Finally, we summarize the results (and mention a few caveats) in Section 4.

2 Bäcklund Tranformations

Given two PDEs:

$$P(u) = 0 \tag{1}$$
$$Q(v) = 0 \tag{2}$$

where $P(\cdot)$ and $Q(\cdot)$ are two differential operators acting on the two fields u and v, respectively, a BT is defined as a pair of relationships:

$$R_1(u, v) = 0 \qquad R_2(u, v) = 0 \tag{3}$$

that can be solved for v if (1) holds, such that v satisfies (2). $R_i(\cdot, \cdot)$ will, in general, also be differential operators, but they may be simpler, exactly integrable, and/or linear, so that solving them is preferrable to solving (2) directly. Furthermore, if $P = Q$, the transformation is referred to as an auto-Bäcklund Tranformation (aBT), and can be used to iteratively generate solutions to a PDE using a single seed solution (Rogers & Shadwick, 1982).

A number of BTs are known to date, such as the Cole-Hopf transformation that connects the solution spaces of Burgers' equation and the heat equation (Chen et al., 2016), or the aBT that leads to the nonlinear superposition principle of the sine-Gordon equation (Miura, 2006). A surprisingly broad class of non-linear PDEs is known to admit such transformations; discovering such relationships can be a practical advantage when one of the PDEs is simpler to solve, better numerical tools are available to generate its solutions, or larger datasets exist. BTs for the Schrödinger equation, in particular, could enable the integration of arbitrary PDEs on quantum hardware via Hamiltonian simulation (Feynman, 1982). In the most general sense, BTs are powerful conceptual tools for modelling nonlinearities in mathematical physics.

3 A CONDITION FOR THE EXISTENCE OF BTS BETWEEN PDE PAIRS

A key feature of Foundation Models is the ability to map disparate information into a single vector space, through vector embeddings that promote semantic clustering. The existence of a shared vector space, where all known information can be gathered and organized, is used, for instance, in domains where different data modalities are available.

In this spirit, a PDE FM might map spatiotemporal sections of individual solutions, represented analytically by multivariate fields $u(x_i)$ and numerically by their discretization into tensors $u^{(N)}_{\{i_n\}} \in \mathbb{R}^{I_1} \times \cdots \times \mathbb{R}^{I_n}$, into vectors $h^{(N)}_m$ in $H \subseteq \mathbb{R}^M$. Model pretraining strives to learn the embeddings that organize the solutions of different PDEs in such a way that the model accuracy is maximised. The specific map details are unimportant for the derivation below; only its existence is assumed.

Let us assume a model is trained on two datasets of N solutions each, corresponding to two distinct PDEs, and that the embedding space is M-dimensional, with $M < N$. A necessary condition for the existence of a BT between the two PDEs is the existence of a linear map $\boldsymbol{T}$ such that

$$\boldsymbol{H}^{(Q)} = \boldsymbol{T}\boldsymbol{H}^{(P)} \tag{4}$$

where $\boldsymbol{H}^{(P)}$ and $\boldsymbol{H}^{(Q)}$ are $M \times N$ matrices whose columns are the M-dimensional embeddings of N solutions of $P(\cdot)$ and $Q(\cdot)$, respectively. This system is overdetermined unless

$$\boxed{\operatorname{rank}(\boldsymbol{C}|\boldsymbol{H}^{(Q)}) \leq \operatorname{rank}\boldsymbol{C}} \tag{5}$$

where $\boldsymbol{C}$ is the $MN \times MN$ coefficient matrix whose non-zero elements are given by:

$$C_{i,\, j+M(\mathrm{mod}(i,N)-1)} = H^P_{i,j} \tag{6}$$

and $(\boldsymbol{C}|\boldsymbol{H}^{(Q)})$ is the corresponding augmented matrix.

It is not surprising that the existence of a BT relies on a compatibility condition – this can be seen as the discrete analog of the integrability conditions that need to be satisfied for the overdetermined system (1)-(3) to have solutions (for a discussion of this aspect in the context of BTs of Euler-Lagrange equations, see Sokalski et al. (2001) and Sokalski (2008)).

4 CONCLUSIONS

We have established a criterion for the existence of BTs between pairs of PDEs; this condition is predicated on the existence of a FM trained on solutions of the two different PDEs, and of a shared embedding space where these solutions are mapped; it permits the formulation of the integrability condition as a linear mapping between families of vectors. The criterion requires the computation of two matrix ranks, which scales, at most, as the cube of the matrix dimension via Gaussian Elimination. For embedding and training dimensions of $O(10^2) - O(10^3)$, this is achievable if sparsity is leveraged (Gould et al., 2007).

The present formulation opens an interesting route for the data-driven discovery of BTs between arbitrary pairs of PDEs; it is, however, only the first step in a more complex algorithmic prescription, which will need to consider the following aspects:

- Condition (5) is only necessary. Even when it is satisfied on the discretized embeddings $h_m^{(N)}$, it may not hold on the continuum solutions $u(x_i)$, or it may only provide a BT over the local solution-space patch spanned by the training data and not globally.
- The criterion does not take into account any ordering among the solutions within each family; as the rank is preserved by row permutation, the ordering does not affect condition (5), but it may affect the form of the reconstructed BT.
- A reconstruction algorithm will need to be deployed, once condition (5) is proven to hold. This may involve, for instance, a regression of the matrix $\boldsymbol{T}$, its composition with the FM encoder map, combined with symbolic regression of the resulting map. This closed-form expression should then be tested by demonstrating analytically that it maps solutions to one PDE problem to the other.

We will explore these extensions in future work.

REFERENCES

Shuhao Cao. Choose a Transformer: Fourier or Galerkin. In *Advances in Neural Information Processing Systems*, volume 34, pp. 24924–24940. Curran Associates, Inc., 2021.

Yang Chen, Engui Fan, and Manwai Yuen. The Hopf–Cole transformation, topological solitons and multiple fusion solutions for the n-dimensional Burgers system. *Physics Letters A*, 380(1-2): 9–14, January 2016. ISSN 03759601. doi: 10.1016/j.physleta.2015.09.033.

Youngsoo Choi, Siu Wun Cheung, Youngkyu Kim, Ping-Hsuan Tsai, Alejandro N. Diaz, Ivan Zanardi, Seung Whan Chung, Dylan Matthew Copeland, Coleman Kendrick, William Anderson, Traian Iliescu, and Matthias Heinkenschloss. Defining Foundation Models for Computational Science: A Call for Clarity and Rigor, May 2025.

Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6):467–488, June 1982. ISSN 1572-9575. doi: 10.1007/BF02650179.

Nicholas I. M. Gould, Jennifer A. Scott, and Yifan Hu. A numerical evaluation of sparse direct solvers for the solution of large sparse symmetric linear systems of equations. *ACM Transactions on Mathematical Software*, 33(2):10, June 2007. ISSN 0098-3500, 1557-7295. doi: 10.1145/1236463.1236465.

Maximilian Herde, Bogdan Raonić, Tobias Rohner, Roger Käppeli, Roberto Molinaro, Emmanuel de Bézenac, and Siddhartha Mishra. Poseidon: Efficient Foundation Models for PDEs, November 2024.

Benjamin Holzschuh, Georg Kohl, Florian Redinger, and Nils Thuerey. P3D: Scalable Neural Surrogates for High-Resolution 3D Physics Simulations with Global Context, September 2025.

Zijie Li, Dule Shu, and Amir Barati Farimani. Scalable Transformer for PDE Surrogate Modeling, November 2023.

Michael McCabe, Payel Mukhopadhyay, Tanya Marwah, Bruno Regaldo-Saint Blancard, Francois Rozet, Cristiana Diaconu, Lucas Meyer, Kaze W. K. Wong, Hadi Sotoudeh, Alberto Bietti, Irina Espejo, Rio Fear, Siavash Golkar, Tom Hehir, Keiya Hirashima, Geraud Krawezik, Francois Lanusse, Rudy Morel, Ruben Ohana, Liam Parker, Mariel Pettee, Jeff Shen, Kyunghyun Cho, Miles Cranmer, and Shirley Ho. Walrus: A Cross-Domain Foundation Model for Continuum Dynamics, November 2025.

Sidharth S. Menon, Trishit Mondal, Shuvayan Brahmachary, Aniruddha Panda, Subodh M. Joshi, Kaushic Kalyanaraman, and Ameya D. Jagtap. On Scientific Foundation Models:Rigorous Definitions, Key Applications, and a Survey, August 2025.

Robert M. Miura. *Bäcklund Transformations, the Inverse Scattering Method, Solitons, and Their Applications: Proceedings of the NSF Research Workshop on Contact Transformations, held in Nashville, Tennessee, 1974*. Springer, November 2006. ISBN 978-3-540-38220-1.

Tung Nguyen, Arsh Koneru, Shufan Li, and Aditya Grover. PhysiX: A Foundation Model for Physics Simulations, July 2025.

Keon Vin Park. Towards a Foundation Model for Physics-Informed Neural Networks: Multi-PDE Learning with Active Sampling, February 2025.

Mahindra Singh Rautela, Alexander Most, Siddharth Mansingh, Bradley C. Love, Ayan Biswas, Diane Oyen, and Earl Lawrence. MORPH: Shape-agnostic PDE Foundation Models, October 2025.

C. Rogers and W. F. Shadwick. *Bäcklund Transformations and Their Applications*. Academic Press, 1982. ISBN 978-0-12-592850-2.

K. Sokalski, T. Wietecha, and Z. Lisowski. A concept of strong necessary condition in nonlinear field theory. *Acta Phys. Polon. B*, 32:2771–2792, 2001.

Krzysztof Sokalski. Exact Analytic Solutions of Nonlinear Partial Differential Equations Supported by Symbolic Programming. September 2008.

Florian Wiesner, Matthias Wessling, and Stephen Baek. Towards a Physics Foundation Model, September 2025.

Haixu Wu, Huakun Luo, Haowen Wang, Jianmin Wang, and Mingsheng Long. Transolver: A Fast Transformer Solver for PDEs on General Geometries, June 2024.

Hang Zhou, Yuezhou Ma, Haixu Wu, Haowen Wang, and Mingsheng Long. Unisolver: PDE-Conditional Transformers Towards Universal Neural PDE Solvers, September 2025.