# MODEL REDUCTION ON MANIFOLDS: A DIFFERENTIAL GEOMETRIC FRAMEWORK

PATRICK BUCHFINK[†,‡], SILKE GLAS[‡], BERNARD HAASDONK[†], AND BENJAMIN UNGER[⋆]

Abstract. Using nonlinear projections and preserving structure in model order reduction (MOR) are currently active research fields. In this paper, we provide a novel differential geometric framework for model reduction on smooth manifolds, which emphasizes the geometric nature of the objects involved. The crucial ingredient is the construction of an embedding for the low-dimensional submanifold and a compatible reduction map, for which we discuss several options. Our general framework allows capturing and generalizing several existing MOR techniques, such as structure preservation for Lagrangian- or Hamiltonian dynamics, and using nonlinear projections that are, for instance, relevant in transport-dominated problems. The joint abstraction can be used to derive shared theoretical properties for different methods, such as an exact reproduction result. To connect our framework to existing work in the field, we demonstrate that various techniques for data-driven construction of nonlinear projections can be included in our framework.

Keywords: model reduction on manifolds, nonlinear projection, Petrov–Galerkin, Lagrangian systems, Hamiltonian systems, data-driven projection

AMS subject classification: 34A26, 34C20, 37C05, 37N30, 65P10

## 1. Introduction

To remedy the computational cost associated with repeated solutions of high-dimensional differential equations, *model order reduction* (MOR) has become an established tool over the last three decades. For an overview of MOR we refer to [3, 6, 35, 63, 64]. The essential idea of MOR approaches can be summarized as follows: Given a high-dimensional initial value problem, which we refer to as the *full-order model* (FOM), find a low-dimensional surrogate system, referred to as the *reduced-order model* (ROM), which is computationally efficient to evaluate. A computationally efficient surrogate model can be interesting in various contexts; for instance (i) if the FOM has to be evaluated for many different parameters (e.g., for parameter studies, sampling-based uncertainty quantification, optimization, or inverse problems), (ii) if the FOM has to be evaluated in realtime (e.g., for model-based control), or (iii) if the computational resources are too little to run the FOM (e.g., on embedded devices). To achieve this goal, classical linear-subspace MOR strives to identify a problem-specific low-dimensional linear subspace such that the state of the initial value problem approximately evolves within this subspace. While this is possible in many applications, the existence of a low-dimensional subspace with good approximation properties cannot always be guaranteed. Mathematically, this can be analyzed by studying the Kolmogorov $n$-widths [41, 61] (or equivalently, as shown in [75], the Hankel singular values), associated with the *set of all solutions* (see the forthcoming Section 3 for the precise definition). For wave-like phenomena in solutions as observed in transport-problems (e.g., the wave-equation or advection-equation) it is now well-understood that the $n$-widths decay slowly

*Date*: January 23, 2024.

for certain initial conditions [17, 26], thus requiring a large dimension of the ROM for a good approximation. To resolve this problem, different paths are pursued in the literature, most of which try to replace the linear-subspace assumption with a nonlinear ansatz. We refer to [69, Cha. 1.3.1] for an overview. In more detail, we assume to be given an initial value problem (the FOM) of the form

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}(t;\mu) = \boldsymbol{f}(t, \boldsymbol{x}(t;\mu);\mu), \qquad \boldsymbol{x}(t_0;\mu) = \boldsymbol{x}_0(\mu) \in \mathbb{R}^N, \tag{1.1}$$

with time interval $\mathcal{I} := (t_0, t_{\mathrm{f}})$, $t_0 < t_{\mathrm{f}} < \infty$ parameter set $\mathcal{P} \subseteq \mathbb{R}^p$, corresponding parameter $\mu \in \mathcal{P}$, and right-hand side $\boldsymbol{f}\colon \mathcal{I} \times \mathbb{R}^N \times \mathcal{P} \to \mathbb{R}^N$, which we want to solve for the unknown state $\boldsymbol{x}\colon \mathcal{I} \times \mathcal{P} \to \mathbb{R}^N$. Roughly speaking, the idea of MOR is to construct a projection $\boldsymbol{\varphi} \circ \boldsymbol{\varrho}$ using two mappings $\boldsymbol{\varrho}\colon \mathbb{R}^N \to \mathbb{R}^n$ and $\boldsymbol{\varphi}\colon \mathbb{R}^n \to \mathbb{R}^N$ with $n \ll N$ and to then derive a low-dimensional surrogate model of the FOM with these mappings. In classical linear-subspace MOR, the mappings $\boldsymbol{\varrho}$ and $\boldsymbol{\varphi}$ are linear, i.e., $\boldsymbol{\varrho}(\boldsymbol{x}) = \boldsymbol{W}^\top \boldsymbol{x}$ and $\boldsymbol{\varphi}(\check{\boldsymbol{x}}) = \boldsymbol{V}\check{\boldsymbol{x}}$ with matrices $\boldsymbol{W}, \boldsymbol{V} \in \mathbb{R}^{N \times n}$ satisfying $\boldsymbol{W}^\top \boldsymbol{V} = \boldsymbol{I}_n$. The associated ROM is given by

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\check{\boldsymbol{x}}(t;\mu) = \boldsymbol{W}^\top \boldsymbol{f}(t, \boldsymbol{V}\check{\boldsymbol{x}}(t;\mu);\mu), \qquad \check{\boldsymbol{x}}(t_0;\mu) = \boldsymbol{W}^\top \boldsymbol{x}_0(\mu) \in \mathbb{R}^n, \tag{1.2}$$

which we solve for the reduced state $\check{\boldsymbol{x}}\colon \mathcal{I} \times \mathcal{P} \to \mathbb{R}^n$. In contrast, we allow for nonlinear mappings in this manuscript, which motivates us to study the FOM (1.1) as a differential equation on a manifold.

1.1. **Main contributions.** In the present paper, we introduce a differential geometric framework for MOR as a unifying framework that contains classical linear-subspace approaches, nonlinear projection frameworks (including machine learning approaches such as autoencoders), and structure-preserving MOR. Our main contributions are:

(i) We provide a general differential geometric framework for MOR on manifolds in Section 3.1. Although the geometric elements we introduce in Section 2 are, of course, not novel, to the best of our knowledge, there is no framework unifying this for MOR. Moreover, we inspect recent approaches of MOR on manifolds and show that these fit into this framework (▷ Table 3).

(ii) On top of the general framework for MOR on manifolds, we introduce the *manifold Petrov–Galerkin* (MPG, ▷ Section 3.3) and *generalized manifold Galerkin* (GMG, ▷ Section 3.3) reduction, which generalize the MOR techniques from [46, 57]. Moreover, the GMG reduction forms the basis for novel structure-preserving variants on manifolds for
   (a) Lagrangian systems (▷ Section 5.2), which we denote by *Lagrangian manifold Galerkin* (LMG), thus extending the linear-subspace model reduction methods in [18, 44], and
   (b) Hamiltonian systems (▷ Section 5.3), which we denote by *symplectic manifold Galerkin* (SMG), extending the MOR method in [16].

(iii) For the respective MOR methods, we give an overview of techniques existing in the literature to construct the nonlinear mappings $\boldsymbol{\varrho}$ and $\boldsymbol{\varphi}$ in a data-driven fashion in Section 6.

Moreover, we provide an exact reproduction result for MOR on manifolds (▷ Theorem 3.6) and discuss a relaxation of the point projection property (3.4a) for autoencoders in Theorem 6.4.

We emphasize that we start with the differential geometric perspective already at the level of the FOM. The main reason for this choice is that starting directly with a differential equation on a manifold highlights the different geometric objects that appear. Note that we focus on a general framework and not on an efficient-to-evaluate surrogate model, which calls upon efficient numerical implementations or additional approximation steps commonly referred to as hyper-reduction.

1.2. **State-of-the-art.** In the following we provide an overview of the various aspects of MOR that fit into this geometric framework.

1.2.1. *MOR and manifolds.* Using (smooth) manifolds in the MOR community is a concept that has been introduced previously. For parametric linear models, interpolation of the linear subspaces or the reduced system matrices on certain manifolds is discussed, for instance, in [2, 50, 73, 80] and has recently been extended to a non-intrusive setting in [59]. To reduce a high-dimensional parameter space during the training phase, the concept of *active manifolds* [13] was developed as a generalization of the so-called *active subspace* [20], which can be interpreted as the dual concept of the Kolmogorov $n$-widths [75]. Moreover, *lifting techniques* may be used to obtain a nonlinear projection of the original system, e.g., [30, 43] or for a non-intrusive setting [62].

1.2.2. *Nonlinear mappings and transport MOR.* If a (localized) quantity is transported through the spatial domain of a PDE over time, such as a shock wave, then it is often not possible to construct a low-dimensional linear subspace that well-approximates the solution since the Kolmogorov $n$-widths do not decay exponentially. Examples studied in the literature are, for instance, the advection equation [56, 66], the wave equation [26], Burgers' equation [17], a pulsed detonation combustor [70], a wildland fire model [11], and a rotating detonation engine [52].

Several nonlinear approaches have been presented to overcome the slowly decaying Kolmogorov $n$-widths, many revolving around the symmetry reduction framework [8, 40, 55, 67]. We mention here exemplary the shifted proper orthogonal decomposition [10, 12, 65], the Lagrangian reference frame method [21, 53], a registration method [22, 74], and front transport reduction [42]. The central idea underlying these methods is to either first transform the state with a suitable transformation such that the resulting transformed FOM is easy to approximate or to encode this transformation directly in the MOR ansatz space.

While the previous approaches are all inspired by the underlying physics of the problem at hand by exploiting the symmetries inherent to the initial value problem, the approaches can be generalized by considering arbitrary nonlinear mappings, for instance, obtained via machine learning paradigms. The natural method for dimensionality reduction is a (shallow) autoencoder [33, 38, 39, 46, 57]. In particular, the work [46] uses terminology from differential geometry and has inspired our work to a large extent. Another parameterization of the nonlinear mappings that are currently investigated is given by polynomials; see, for instance, [4, 7, 23, 37, 72].

1.2.3. *Structure-preserving MOR for Lagrangian and Hamiltonian systems.* Classical linear-subspace MOR for Lagrangian systems is discussed in [18, 44]. Notably, the authors of [44] already mention that the same methods can be used with nonlinear embeddings $\boldsymbol{\varphi}$, albeit without explicitly formulating the required differential geometric objects. Moreover, we show how a reduced Lagrangian system can be interpreted as a projection of the Euler–Lagrangian vector field using the GMG reduction (▷ Theorem 5.7).

Structure-preserving MOR for Hamiltonian systems is discussed in [24,47,60,79] using linear subspaces and in [16,72] for manifolds (in coordinates). A Hamiltonian-preserving Neural Galerkin scheme is presented in [71]. Moreover, some of the ideas for structure-preserving MOR for port-Hamiltonian systems [76], a generalization of Hamiltonian systems to open systems, can be used for structure-preserving MOR for Hamiltonian systems. We refer to [51, Rem. 8.2] for an overview.

Besides classical MOR schemes that rely on a given large-scale dynamical system, non-intrusive methods aim to learn a potentially low-dimensional representation from system measurements directly. In the context of learning Hamiltonian systems, we exemplary mention [27,29,32,78].

1.3. **Structure of the paper.** To render the manuscript self-contained, we start our exposition by reviewing all necessary concepts from differential geometry (▷ Section 2). Readers familiar with these concepts might skip this section and directly start with Section 3, where we introduce our general MOR framework for initial value problems on manifolds. Additional structure preservation is detailed in Section 5, which is based on additional geometrical structures (▷ Section 4). A discussion on specific data-driven construction approaches for the required nonlinear mappings is presented in Section 6 and followed by conclusions (▷ Section 7).

1.4. **Notation.** We use the *index notation*, which differentiates between upper indices $v^{\underline{i}}$ and lower indices $\lambda_{\underline{i}}$. Let us emphasize that indices that concern the index notation are underlined. The position of the index indicates the type of the geometric object. Moreover, we utilize the *Einstein summation convention*, which implies the summation over an index if the index appears twice (once as a lower index and once as an upper index). For an $N$-dimensional vector space $\mathcal{V}$, this notation is used to abbreviate (i) the linear combination of a basis $\{E_{\underline{i}}\}_{i=1}^N \subseteq \mathcal{V}$ with coefficients $\{v^{\underline{i}}\}_{i=1}^N \subseteq \mathbb{R}$, (ii) the linear combination of a dual basis $\{F^{\underline{i}}\}_{i=1}^N \subseteq \mathcal{V}^*$ with coefficients $\{\lambda_{\underline{i}}\}_{i=1}^N \subseteq \mathbb{R}$, or (iii) the dual product of the respective coefficients,

$$\text{(i) } v^{\underline{i}} E_{\underline{i}} := \sum_{i=1}^N v^{\underline{i}} E_{\underline{i}} \in \mathcal{V}, \quad \text{(ii) } \lambda_{\underline{i}} F^{\underline{i}} := \sum_{i=1}^N \lambda_{\underline{i}} F^{\underline{i}} \in \mathcal{V}^*, \quad \text{(iii) } v^{\underline{i}} \lambda_{\underline{i}} := \sum_{i=1}^N v^{\underline{i}} \lambda_{\underline{i}} \in \mathbb{R}. \tag{1.3}$$

Moreover, we use $[v^{\underline{i}}]_{1 \leq i \leq N} \in \mathbb{R}^N$ to stack scalars $v^{\underline{i}} \in \mathbb{R}$ as a vector in $\mathbb{R}^N$. Further notation is introduced in Section 2.7.

## 2. A primer on differential geometry

We start this work by recalling several important definitions and results from the theory of smooth manifolds to render this manuscript self-contained. Our presentation is largely based on the monograph [45]. In particular, all material within this section that is not explicitly referenced is adopted from [45].

To motivate the forthcoming definitions, we briefly discuss the tools required

(i) to formulate a differential equation on a manifold and
(ii) to define the submanifold and mappings needed

to perform model reduction on manifolds. For the differential geometric formulation of the FOM, we stepwise introduce the structure of a *smooth manifold* starting from *topological spaces*. The topology allows us to characterize *continuous functions* and *smooth manifolds* (▷ Section 2.1). Then, having the needed structure at hand, we continue to define *continuously differentiable functions* on smooth manifolds (▷Section 2.2). Subsequently, we introduce the *tangent space* at a point on the manifold (▷Section 2.3) to be able to formulate the *differential of a function* (▷Section 2.4), which is used to generalize the time-derivative of the state to the manifold setting. In order to describe the evolution of an initial value problem, we set the right-hand side to be a *vector field* (▷Section 2.5). With these preparations, a differential equation on a manifold can be formulated (▷Section 2.6). We refer to Table 1 for a comparison of a dynamical system on a vector space and on a smooth manifold. Furthermore, for the model reduction framework, we discuss embedded submanifolds (▷Section 2.8).

**Table 1** – Formulation of a dynamical system in the time interval $\mathcal{I}$ on a vector space (left) in comparison to a differential geometric formulation (right).

| **dynamical system on a vector space** | | **dynamical system on a manifold** | |
|---|---|---|---|
| $\mathbb{R}^N$ | $N$-dim. vector space | $\mathcal{M}$ | $N$-dim. smooth manifold |
| | | $T_m\mathcal{M}$, $T\mathcal{M}$ | tangent space, tangent bundle |
| $\boldsymbol{f}\colon \mathbb{R}^N \to \mathbb{R}^N$ | right-hand side | $X\colon \mathcal{M} \to T\mathcal{M}$ | vector field |
| $\boldsymbol{x}\colon \mathcal{I} \to \mathbb{R}^N$ | solution curve | $\gamma\colon \mathcal{I} \to \mathcal{M}$ | solution curve |
| $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}(t) \in \mathbb{R}^N$ | time-derivative | $\frac{\mathrm{d}}{\mathrm{d}t}\gamma\big\|_t \in T_{\gamma(t)}\mathcal{M}$ | velocity |
| $\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}(t) = \boldsymbol{f}(\boldsymbol{x}(t)) \in \mathbb{R}^N \\ \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \in \mathbb{R}^N \end{cases}$ | | $\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\gamma\big\|_t = X\|_{\gamma(t)} \in T_{\gamma(t)}\mathcal{M} \\ \gamma(t_0) = \gamma_0 \in \mathcal{M} \end{cases}$ | |

2.1. **Chart and smooth manifold.** For two topological spaces $\mathcal{M}$ and $\mathcal{Q}$ (▷Appendix A.1), a map $F\colon \mathcal{M} \to \mathcal{Q}$ is called a *homeomorphism* if (i) it is bijective (and thus the inverse $F^{-1}\colon \mathcal{Q} \to \mathcal{M}$ exists) and (ii) both, $F$ and $F^{-1}$, are continuous. Moreover, $\mathcal{M}$ is called *locally homeomorphic to* $\mathbb{R}^N$ for $N \in \mathbb{N}$ if for every point $m \in \mathcal{M}$ there exists an open set $U \subseteq \mathcal{M}$ with $m \in U$ and a homeomorphism $F\colon U \to F(U) \subseteq \mathbb{R}^N$. A topological space $\mathcal{M}$ is called a *topological manifold* of *dimension* $N$ if it is locally homeomorphic to $\mathbb{R}^N$ (and additionally Hausdorff and second-countable, see e.g. [45, Cha. 1 and App. A]). We denote the dimension with $\dim(\mathcal{M}) = N$.

Let $\mathcal{M}$ be a topological manifold of dimension $N$. A *chart* is a tuple $(U, x)$ where the *chart domain* $U \subseteq \mathcal{M}$ is an open set and the *chart mapping* $x\colon U \to x(U) \subseteq \mathbb{R}^N$ is a homeomorphism. For two charts $(U, x)$ and $(V, y)$ with $U \cap V \neq \emptyset$, we can define the *transition mappings*

$$x \circ y^{-1}\colon y(U \cap V) \to x(U \cap V) \qquad \text{and} \qquad y \circ x^{-1}\colon x(U \cap V) \to y(U \cap V),$$

which are homeomorphisms as composition of homeomorphisms (▷Figure 1). The charts $(U, x)$ and $(V, y)$ are called $\mathcal{C}^k$*-compatible* for $k \in \mathbb{N}$ or $k = \infty$ if either $U \cap V = \emptyset$ or

$$x \circ y^{-1} \in \mathcal{C}^k(y(U \cap V), x(U \cap V)) \qquad \text{and} \qquad y \circ x^{-1} \in \mathcal{C}^k(x(U \cap V), y(U \cap V)),$$

where differentiability is defined in the classical sense since $x(U \cap V), y(U \cap V) \subseteq \mathbb{R}^N$. A collection of charts $\mathcal{A} = \{(U_i, x_i) \mid i \in I\}$ with some index set $I$ is called an *atlas for* $\mathcal{M}$ if $\mathcal{M} = \bigcup_{i \in I} U_i$. The atlas is called *of class* $\mathcal{C}^k$ (or a $\mathcal{C}^k$*-atlas*) if all charts in $\mathcal{A}$ are mutually $\mathcal{C}^k$-compatible. We call a $\mathcal{C}^k$-atlas $\mathcal{A}$ *maximal* if all charts that are $\mathcal{C}^k$-compatible with any chart in $\mathcal{A}$ are already elements of $\mathcal{A}$. If $\mathcal{A}$ is a maximal $\mathcal{C}^k$ atlas for $\mathcal{M}$, then the tuple $(\mathcal{M}, \mathcal{A})$ is called a $\mathcal{C}^k$*-manifold* and, in particular, a *smooth manifold* if $k = \infty$. As common

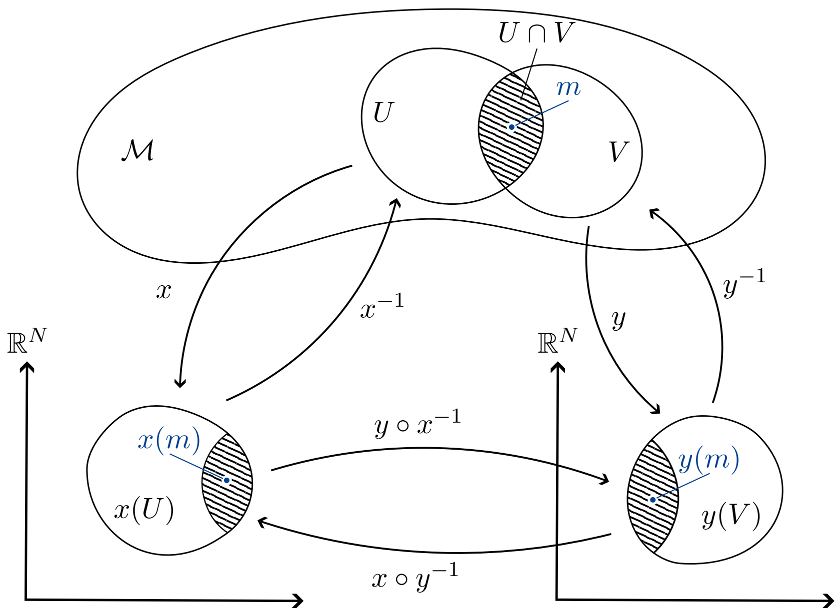



**Figure 1** – Two intersecting chart domains $U, V \subseteq \mathcal{M}$ with respective chart mappings $x, y$ and transition mappings $x \circ y^{-1}, y \circ x^{-1}$ on $x\,(U \cap V) \subseteq \mathbb{R}^N$ and $y\,(U \cap V) \subseteq \mathbb{R}^N$.

in the literature, we omit the explicit mentioning of the maximal atlas whenever possible and say that $\mathcal{M}$ is a $\mathcal{C}^k$-manifold, implicitly assuming a maximal $\mathcal{C}^k$ atlas to be available.

2.2. **Diffeomorphism and partial derivative.** Assume now that we have smooth manifolds $\mathcal{M}$ and $\mathcal{Q}$ of dimension $N$ and $Q$. A mapping $F\colon \mathcal{M} \to \mathcal{Q}$ is called *of class* $\mathcal{C}^k$, or in short notation $F \in \mathcal{C}^k(\mathcal{M}, \mathcal{Q})$, if for every $m \in \mathcal{M}$, there exist charts $(U, x)$ containing $m$ and $(V, y)$ containing $F\,(m)$ such that $y \circ F \circ x^{-1} \in \mathcal{C}^k(x\,(U)\,, y\,(V))$ in the classical sense since $x\,(U) \subseteq \mathbb{R}^N$ and $y\,(V) \subseteq \mathbb{R}^Q$. Note that due to the $\mathcal{C}^k$-compatibility of the charts, this definition of differentiability does not depend on the choice of the chart. Throughout the document, a *smooth* mapping is synonymous with mappings of the class $\mathcal{C}^\infty$. We restrict ourselves in this work to smooth manifolds and smooth mappings to simplify the presentation. A smooth bijective map $F \in \mathcal{C}^\infty(\mathcal{M}, \mathcal{Q})$ which has a smooth inverse is called a *smooth diffeomorphism (from $\mathcal{M}$ to $\mathcal{Q}$)*.

For calculations, we formulate the derivative in the index notation (▷Section 1.4). In more detail, we denote for $1 \le i \le Q$ the *$i$-th component function of the chart mapping* as $x^{\underline{i}}\colon U \to \mathbb{R}$ and the *$i$-th component function (of $F$)* with $F^{\underline{i}} := y^{\underline{i}} \circ F\colon U \to \mathbb{R}$. Then, the *$i$-th partial derivative of the $j$-th component of $F \in \mathcal{C}^1(\mathcal{M}, \mathcal{Q})$ at $m \in \mathcal{M}$ (in $(U, x)$ and $(V, y)$)* is defined by

$$\left.\frac{\partial F^{\underline{j}}}{\partial x^{\underline{i}}}\right|_m := \Big(\partial_i (F^{\underline{j}} \circ x^{-1})\Big)(x\,(m)) \qquad \text{for } 1 \le i \le N, \tag{2.1}$$

where $\partial_i(\cdot)$ describes the $i$-th partial derivative of functions mapping between Euclidean vector spaces. For scalar-valued functions $f \in \mathcal{C}^1(\mathcal{M}, \mathbb{R})$, we omit the index, i.e., $f^{\underline{1}} \equiv f$ and thus $\left.\frac{\partial f^{\underline{1}}}{\partial x^{\underline{i}}}\right|_m = \left.\frac{\partial f}{\partial x^{\underline{i}}}\right|_m$. For the derivative of the chart mapping, we obtain

$$x \in \mathcal{C}^\infty(\mathcal{M}, \mathbb{R}^N) \qquad \text{with} \qquad \left.\frac{\partial x^{\underline{j}}}{\partial x^{\underline{i}}}\right|_m = \partial_i (x^{\underline{j}} \circ x^{-1})\,(x\,(m)) = \delta_{\underline{i}}^{\underline{j}} := \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \tag{2.2}$$

due to $(x^{\underline{j}} \circ x^{-1}) = (x \circ x^{-1})^{\underline{j}} = (\mathrm{id}_{\mathbb{R}^N})^{\underline{j}}$. The function $\delta^{\underline{j}}_{\underline{i}}$ is known as the *Kronecker delta*.

2.3. **Tangent and tangent space.** Consider a smooth manifold $\mathcal{M}$ of dimension $N$. The tangent space of $\mathcal{M}$ can be defined in multiple alternative ways (see, e.g., [1, Sec. 1.6] for an overview). In the present work, we present the *derivation approach* and closely follow [9, Sec. 1.7]. For an arbitrary but fixed point $m \in \mathcal{M}$, we consider the set

$$F^\infty_m := \{f \in \mathcal{C}^\infty(U, \mathbb{R}) \,|\, U \subseteq \mathcal{M} \text{ open with } m \in U\}.$$

Then, a *tangent at* $m \in \mathcal{M}$ is a function on this set $v\colon F^\infty_m \to \mathbb{R}$ which is (i) linear and (ii) fulfills the product rule, i.e., for every $f, g \in F^\infty_m$ and $a, b \in \mathbb{R}$, it holds[1]

$$\text{(i) } v\,(af + bg) = av\,(f) + bv\,(g) \in \mathbb{R}, \quad \text{(ii) } v\,(f \cdot g) = v\,(f) \cdot g\,(m) + f\,(m) \cdot v\,(g) \in \mathbb{R}.$$

In a broader context, the properties (i) and (ii) define a *derivation*. The set of all tangents at $m \in \mathcal{M}$

$$T_m\mathcal{M} := \{v\colon F^\infty_m \to \mathbb{R} \,|\, v \text{ is a tangent at } m\} \tag{2.3}$$

defines the *tangent space at* $m$, which can be shown to be an $N$-dimensional vector space. Thus, we also refer to elements $v \in T_m\mathcal{M}$ as *tangent vectors* at $m$. The $i$-th partial derivative (2.1) of a scalar-valued function $f \in F^\infty_m$ can be used to define elements in $T_m\mathcal{M}$,

$$\left.\frac{\partial}{\partial x^{\underline{i}}}\right|_m \in T_m\mathcal{M} \qquad \text{with} \qquad \left.\frac{\partial}{\partial x^{\underline{i}}}\right|_m\colon F^\infty_m \to \mathbb{R},\ f \mapsto \left.\frac{\partial f}{\partial x^{\underline{i}}}\right|_m.$$

Moreover, $\left(\left.\frac{\partial}{\partial x^{\underline{1}}}\right|_m, \dots, \left.\frac{\partial}{\partial x^{\underline{N}}}\right|_m\right)$ is an ordered basis of $T_m\mathcal{M}$, and, thus, we can represent each tangent vector

$$v \in T_m\mathcal{M} \qquad \text{with} \qquad v = v^{\underline{i}} \left.\frac{\partial}{\partial x^{\underline{i}}}\right|_m,$$

where we refer to $v^{\underline{i}} \in \mathbb{R}$, $1 \leq i \leq N$, as the *components (of* $v$*)* and where we implicitly sum over $1 \leq i \leq N$ by the Einstein summation convention (1.3). Note that for this formalism to work, the index $i$ in the denominator of $\left.\frac{\partial}{\partial x^{\underline{i}}}\right|_m$ counts as a lower index. In the case of a vector space $\mathcal{V}$, the tangent space $T_m\mathcal{V}$ can be identified with the vector space $T_m\mathcal{V} \cong \mathcal{V}$ for all $m \in \mathcal{V}$ [45, p. 59], in particular $T_m\mathbb{R}^k \cong \mathbb{R}^k$ for $k \in \mathbb{N}$.

2.4. **Differential and chain rule.** Consider smooth manifolds $\mathcal{M}$, $\mathcal{Q}$, and $\mathcal{L}$ of dimension $N$, $Q$, and $L$ with charts $(U, x)$, $(V, y)$, and $(W, z)$ and a point $m \in U$. The *differential of a smooth map* $F \in \mathcal{C}^\infty(\mathcal{M}, \mathcal{Q})$ *at* $m$ is a linear map

$$\mathrm{d}F|_m \in \mathcal{C}^\infty(T_m\mathcal{M}, T_{F(m)}\mathcal{Q}), \quad v^{\underline{i}} \left.\frac{\partial}{\partial x^{\underline{i}}}\right|_m \mapsto v^{\underline{i}} \left.\frac{\partial F^{\underline{j}}}{\partial x^{\underline{i}}}\right|_m \left.\frac{\partial}{\partial y^{\underline{j}}}\right|_{F(m)}, \tag{2.4}$$

which maps between the respective tangent spaces using the $i$-th partial derivative (2.1) of the $j$-th component function $F^{\underline{j}}$ of $F$, where we sum over $1 \leq i \leq N$ and $1 \leq j \leq Q$ by the Einstein summation convention (1.3). The *chain rule* is an important property of the differential: For two smooth mappings $F \in \mathcal{C}^\infty(\mathcal{M}, \mathcal{Q})$, $G \in \mathcal{C}^\infty(\mathcal{Q}, \mathcal{L})$, it holds

$$\mathrm{d}(G \circ F)|_m = \mathrm{d}G|_{F(m)} \circ \mathrm{d}F|_m\colon T_m\mathcal{M} \to T_{(G \circ F)(m)}\mathcal{L}. \tag{2.5}$$

[1]Note that for the sum/product of functions $f : U_f \to \mathbb{R}$ and $g : U_g \to R$ from $F^\infty_m$, the domain of the products/sums is shrinked to the intersection $U_f \cap U_g$, which is still open and $m \in U_f \cap U_g$ and thus $(f + g), (f \cdot g) \in F^\infty_m$.

In respective charts $(U, x)$, $(V, y)$, $(W, z)$ with $m \in U$, $F(m) \in V$ and $(G \circ F)(m) \in W$, the chain rule reads

$$\left.\frac{\partial (G \circ F)^{\underline{i}}}{\partial x^{\underline{j}}}\right|_m = \left.\frac{\partial G^{\underline{i}}}{\partial y^{\underline{k}}}\right|_{F(m)} \left.\frac{\partial F^{\underline{k}}}{\partial x^{\underline{j}}}\right|_m \qquad \text{for all} \quad \begin{cases} 1 \leq j \leq N, \\ 1 \leq i \leq L, \end{cases}$$

where the right-hand side sums over $1 \leq k \leq Q$ by the Einstein summation convention (1.3).

2.5. **Tangent bundle and vector field.** The *tangent bundle* is the disjoint union of all tangent spaces

$$T\mathcal{M} := \dot{\bigcup}_{m \in \mathcal{M}} T_m\mathcal{M} := \{(m, v) \mid m \in \mathcal{M},\ v \in T_m\mathcal{M}\}, \tag{2.6}$$

which bundles all points $m \in \mathcal{M}$ and corresponding tangent vectors $v \in T_m\mathcal{M}$ in one set. The *tangent bundle* itself is a smooth manifold of dimension $2N$. In the scope of the present work, we typically use $(m, v) \in T\mathcal{M}$ to denote elements in $T\mathcal{M}$. Whenever we have a mapping into a tangent bundle, then we use the notation $\left.(\cdot)\right|_m$ to denote the second part of the mapping. For a given smooth mapping $F \in \mathcal{C}^\infty(\mathcal{M}, \mathcal{Q})$, the *differential (on the tangent bundle)*

$$\mathrm{d}F \in \mathcal{C}^\infty(T\mathcal{M}, T\mathcal{Q}),\ (m, v) \mapsto \left(F(m), \left.\mathrm{d}F\right|_m(v)\right) \tag{2.7}$$

collects the differentials $\left.\mathrm{d}F\right|_m \in \mathcal{C}^\infty(T_m\mathcal{M}, T_{F(m)}\mathcal{Q})$ at $m$ for all points $m \in \mathcal{M}$. For a given chart $(U, x)$ of $\mathcal{M}$, the differential of the chart mapping $\mathrm{d}x \in \mathcal{C}^\infty(TU, T\mathbb{R}^N)$ defines a *natural chart* of $TU$ by identifying $T\mathbb{R}^N$ with $\mathbb{R}^{2N}$. It maps

$$\mathrm{d}x\left(\left(m, v^{\underline{i}} \left.\frac{\partial}{\partial x^{\underline{i}}}\right|_m\right)\right) = \left(x(m), \left[v^{\underline{i}}\right]_{1 \leq i \leq N}\right) \in \mathbb{R}^{2N} \tag{2.8}$$

since for a chart mapping[2] $y \in \mathcal{C}^\infty(\mathbb{R}^N, \mathbb{R}^N)$ of $\mathbb{R}^N$, it holds with (2.2) and (2.4) that

$$\left.\mathrm{d}x\right|_m\left(v^{\underline{i}} \left.\frac{\partial}{\partial x^{\underline{i}}}\right|_m\right) = v^{\underline{i}} \left.\frac{\partial x^{\underline{j}}}{\partial x^{\underline{i}}}\right|_m \left.\frac{\partial}{\partial y^{\underline{j}}}\right|_{x(m)} = v^{\underline{i}} \delta_{\underline{i}}^{\underline{j}} \left.\frac{\partial}{\partial y^{\underline{j}}}\right|_{x(m)} = v^{\underline{i}} \left.\frac{\partial}{\partial y^{\underline{i}}}\right|_{x(m)} \in T_{x(m)}\mathbb{R}^N,$$

which we identify with $\left[v^{\underline{i}}\right]_{1 \leq i \leq N} \in \mathbb{R}^N$.

Since $\mathcal{M}$ and $T\mathcal{M}$ are both smooth manifolds, we are able to define smooth mappings from $\mathcal{M}$ to $T\mathcal{M}$ based on Section 2.2. A *smooth vector field* is a mapping $X \in \mathcal{C}^\infty(\mathcal{M}, T\mathcal{M})$ with $\pi \circ X = \mathrm{id}_\mathcal{M}$ with $\pi \colon T\mathcal{M} \to \mathcal{M}$, $(m, v) \mapsto m$. It assigns each point $m \in \mathcal{M}$ an element $X(m) := (m, \left.X\right|_m) \in T\mathcal{M}$ in the tangent bundle, where we denote the *vector field at* $m \in \mathcal{M}$ with $\left.X\right|_m \in T_m\mathcal{M}$. The *set of all smooth vector fields on* $\mathcal{M}$ is denoted with $\mathfrak{X}_\mathcal{M}$.

2.6. **Curve and initial value problem.** For a given smooth manifold $\mathcal{M}$ and an interval $\mathcal{I} := (t_0, t_\mathrm{f})$ with $t_0 < t_\mathrm{f} < \infty$, we call the mapping $\gamma \in \mathcal{C}^\infty(\mathcal{I}, \mathcal{M})$ a *smooth curve.* We refer to elements $t \in \mathcal{I}$ as *time points.* By custom, we use for the derivative w.r.t. time the notation $\frac{\mathrm{d}}{\mathrm{d}t}(\cdot)$. The *velocity of a curve* $\gamma$ *at* $t \in \mathcal{I}$ is

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\gamma\right|_t := \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\gamma^{\underline{i}}\right|_t\right) \left.\frac{\partial}{\partial x^{\underline{i}}}\right|_{\gamma(t)} \in T_{\gamma(t)}\mathcal{M},$$

[2]This chart mapping may seem redundant as $y \equiv \mathrm{id}_{\mathbb{R}^N}$. However, we use it to illustrate how $T_{x(m)}\mathbb{R}^N$ is identified with $\mathbb{R}^N$ by using $y$ to denote the basis vectors $\left.\frac{\partial}{\partial y^{\underline{j}}}\right|_{x(m)} \in T_{x(m)}\mathbb{R}^N$.

i.e., an element in the tangent space based on the (classical) derivative of the component functions $\gamma^{\underline{i}}\colon \mathbb{R} \supseteq \mathcal{I} \to \mathbb{R}$ of the curve.[3]

For a smooth vector field $X \in \mathfrak{X}_{\mathcal{M}}$, we call $\gamma \in \mathcal{C}^\infty(\mathcal{I}, \mathcal{M})$ an *integral curve of* $X$ with *initial value* $\gamma_0 \in \mathcal{M}$, if

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\gamma\Big|_t = X|_{\gamma(t)} \in T_{\gamma(t)}\mathcal{M}, & t \in \mathcal{I} \\ \gamma(t_0) = \gamma_0 \in \mathcal{M}. \end{cases} \tag{2.9}$$

We refer to (2.9) as an *initial value problem (on $\mathcal{M}$)*. For a given chart $(U, x)$, the system (2.9) can be solved via an $N$-dimensional initial value problem on $\mathbb{R}^N$

$$\frac{\mathrm{d}}{\mathrm{d}t}\gamma^{\underline{i}}\Big|_t = \left(X|_{\gamma(t)}\right)^{\underline{i}} \in \mathbb{R} \qquad \text{for } 1 \leq i \leq N, \qquad \gamma^{\underline{i}}(t_0) = x^{\underline{i}}(\gamma_0) \in \mathbb{R}. \tag{2.10}$$

Due to the assumption of a smooth vector field, we know that there exists a unique integral curve by the fundamental theorem on flows [45, Thm. 9.12], if the final time $t_{\mathrm{f}}$ is small enough. If we assume that there exists a time interval such that all integral curves exist for a set $M_0 \subseteq \mathcal{M}$ with the starting points $\gamma_0 \in M_0$, the *flow of* $X$ can be defined as

$$\theta_t\colon M_0 \to \mathcal{M}, \quad \gamma_0 \mapsto \gamma(t; \gamma_0).$$

2.7. **Bold notation.** We introduce a notation that collects all previously introduced types of differential geometric objects (like points, functions, tangent vectors) in a fixed chart and thereby reduces to computations in $\mathbb{R}$-vector spaces. We refer to this notation as the *bold notation*[4]. For a given smooth manifold $\mathcal{M}$ with a chart $(U, x)$, we use

$$x \in \mathcal{C}^\infty(U, \mathbb{R}^N), \qquad \mathrm{d}x|_m \in \mathcal{C}^\infty(T_mU, \mathbb{R}^N) \quad \text{with } m \in U, \qquad \mathrm{d}x \in \mathcal{C}^\infty(TU, \mathbb{R}^{2N})$$

to map the different types of objects accordingly, where we identify $T_{x(m)}\mathbb{R}^N$ with $\mathbb{R}^N$ for $\mathrm{d}x|_m$ and $T\mathbb{R}^N$ with $\mathbb{R}^{2N}$ for $\mathrm{d}x$. Let us state clearly that (i) this formulation loses geometrical information (as it treats different types of objects as a vector in $\mathbb{R}^N$) and (ii) it only works for one fixed chart (since the explicit dependence on the chart is neglected). However, this formulation can be helpful for readers new to the field of differential geometry with more background in classical numerical analysis and engineering. The notation for the different types of differential geometric objects for two smooth manifolds $\mathcal{M}, \mathcal{Q}$ with charts $(U, x)$, $(V, y)$, respectively, are summarized in Table 2.

2.8. **Embedding and embedded submanifold.** Consider two smooth manifolds $\check{\mathcal{M}}$ and $\mathcal{M}$ of dimension $n$ and $N$, respectively. A smooth mapping $F \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$ is called an *immersion* if the respective differential $\mathrm{d}F|_{\check{m}}\colon T_{\check{m}}\check{\mathcal{M}} \to T_{F(\check{m})}\mathcal{M}$ is injective at each point $\check{m} \in \check{\mathcal{M}}$. Moreover, $F$ is called a *smooth embedding* if it is a smooth immersion and a homeomorphism onto its image $F(\check{\mathcal{M}}) \subseteq \mathcal{M}$. For a given smooth embedding $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$, the image $\varphi(\check{\mathcal{M}})$ is an $n$-dimensional smooth manifold, which is called an *embedded (or regular) submanifold* of $\mathcal{M}$. We denote the tangent space of $\varphi(\check{\mathcal{M}})$ at $\varphi(\check{m})$

[3]Alternatively, the velocity of a curve can be understood in terms of the derivative introduced in Section 2.4 with $\frac{\mathrm{d}}{\mathrm{d}t}\gamma\big|_t = \mathrm{d}\gamma|_t$. In the presented notation, this would require to understand $\mathcal{I}$ as a smooth manifold with the chart $(\mathcal{I}, x_{\mathcal{I}})$, chart mapping $x_{\mathcal{I}} \equiv \mathrm{id}_{\mathcal{I}}\colon \mathcal{I} \to \mathbb{R}$ and the derivative $\frac{\mathrm{d}}{\mathrm{d}t}\gamma^{\underline{i}}\big|_t = \frac{\partial\gamma^{\underline{i}}}{\partial x_{\mathcal{I}}^{\underline{1}}}\Big|_t$.

[4]Be aware that bold symbols may be used for other purposes in other scripts on differential geometry.

[5]To be more precise, the Jacobian matrix is the coordinate matrix of the linear mapping described by the differential in coordinates $\mathrm{d}y|_{F(m)} \circ \mathrm{d}F|_m \circ \mathrm{d}x|_m^{-1}\colon \mathbb{R}^N \to \mathbb{R}^Q$. Moreover, we use in the last column of this row a notation for stacking scalars as matrices similarly to stacking scalars as vectors from Section 1.4.

**Table 2** – Bold notation for two smooth manifolds $\mathcal{M}$, $\mathcal{Q}$ with charts $(U, x)$, $(V, y)$, respectively.

| **type** | **element** | **bold notation** |
|---|---|---|
| point | $m \in U \subseteq \mathcal{M}$ | $\boldsymbol{m} := x\,(m) \in \mathbb{R}^N$ |
| mapping | $F \in \mathcal{C}^\infty(U, V)$ | $\boldsymbol{F} := y \circ F \circ x^{-1} : \mathbb{R}^N \supseteq x(U) \to y(V) \subseteq \mathbb{R}^Q$ |
| tangent vector | $v = v^{\underline{i}} \frac{\partial}{\partial x^{\underline{i}}}\Big\|_m \in T_m U$ | $\boldsymbol{v} := \left[v^{\underline{i}}\right]_{1 \leq i \leq N} \in \mathbb{R}^N$ |
| Jacobian matrix[5] | $\mathrm{d}F\|_m \in \mathcal{C}^\infty(T_m U, T_{F(m)} V)$ | $\boldsymbol{DF}\|_{\boldsymbol{m}} := \left[\frac{\partial F^{\underline{i}}}{\partial x^{\underline{j}}}\Big\|_m\right]_{\substack{1 \leq i \leq Q, \\ 1 \leq j \leq N}} \in \mathbb{R}^{Q \times N}$ |
| dynamical system | $\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\gamma\big\|_t = X\|_{\gamma(t)} \in T_{\gamma(t)} U, \\ \gamma(t_0) = \gamma_0 \in \mathcal{M} \end{cases}$ | $\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\gamma}\big\|_t = \boldsymbol{X}\|_{\boldsymbol{\gamma}(t)} \in \mathbb{R}^N, \\ \boldsymbol{\gamma}(t_0) = \boldsymbol{\gamma}_0 \in \mathbb{R}^N \end{cases}$ |

with $T_{\varphi(\check{m})}(\varphi(\check{\mathcal{M}})) := \mathrm{d}\varphi|_{\check{m}}\,(T_{\check{m}}\check{\mathcal{M}})$. From the assumptions, it follows automatically that the embedding $\varphi$ is a smooth diffeomorphism onto its image [45, Prop. 5.2].

**Lemma 2.1.** *Consider smooth manifolds $\check{\mathcal{M}}, \mathcal{M}$ and smooth mappings $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$ and $\varrho \in \mathcal{C}^\infty(\mathcal{M}, \check{\mathcal{M}})$ with $\varrho \circ \varphi \equiv \mathrm{id}_{\check{\mathcal{M}}}$. Then, $\varphi$ is a smooth embedding and $\varphi(\check{\mathcal{M}}) \subseteq \mathcal{M}$ is an embedded submanifold.*

*Proof.* (▷ Appendix A.2). □

## 3. Model order reduction on manifolds

With the geometric sundries at hand, we can now introduce *model order reduction on manifolds*. We start with the general framework for model order reduction (▷ Section 3.1). Then, we detail conditions such that exact reproduction can be achieved (▷ Section 3.2). Finally, we present an example fitting this framework, the so-called manifold Petrov–Galerkin (▷ Section 3.3).

### 3.1. **General framework.**

This section sits at the heart of this paper and introduces the general framework upon which the remainder is built. We start this section by defining the FOM on manifolds (▷ Section 3.1.1). We then focus on the goal that MOR strives to achieve and what assumptions are required to reach this goal (▷ Section 3.1.2). Subsequently, we define the reduction map, which is needed to define the reduced-order model (▷ Section 3.1.3). We conclude the general framework with a workflow for MOR on manifolds (▷ Section 3.1.4).

3.1.1. *Full-order model.* In the scope of the present work, we consider high-dimensional parametric initial value problems. More precisely, assume that we are given a time interval $\mathcal{I} := (t_0, t_\mathrm{f})$ with initial time $t_0$ and final time $t_\mathrm{f} > t_0$, a parameter set $\mathcal{P} \subseteq \mathbb{R}^p$, an $N$-dimensional smooth manifold $\mathcal{M}$ with large $N$, a (possibly parametric) smooth vector field $X\colon \mathcal{P} \to \mathfrak{X}_\mathcal{M}$, and a (possibly parametric) initial value $\gamma_0\colon \mathcal{P} \to \mathcal{M}$. We consider for $\mu \in \mathcal{P}$ the initial value problem

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\gamma\Big|_{t;\mu} = X(\mu)|_{\gamma(t;\mu)} \in T_{\gamma(t;\mu)}\mathcal{M}, \quad t \in \mathcal{I} \\ \gamma\,(t_0;\mu) = \gamma_0(\mu) \in \mathcal{M}, \end{cases} \tag{3.1}$$

which we want to solve for the integral curve $\gamma\,(\cdot;\mu) \in \mathcal{C}^\infty(\mathcal{I}, \mathcal{M})$. We refer to (3.1) as the FOM and to $X\,(\mu)$ as the *FOM vector field*.

**Remark 3.1** (Parameter dependency)**.** *In the following, we may suppress the explicit notation of the parameter dependence for the sake of brevity. This is possible since the parameter is fixed for each* FOM *evaluation. We indicate the parameter dependence only if it is relevant in a specific context.*

3.1.2. *Goal of model order reduction.* The goal of MOR can be formulated as to be able to well-approximate the *set of all solutions*

$$S := \{\gamma\,(t;\mu) \in \mathcal{M} \mid (t,\mu) \in \mathcal{I} \times \mathcal{P}\} \subseteq \mathcal{M} \tag{3.2}$$

computationally efficiently. Sometimes, the set of all solutions is referred to as the *solution manifold*. However, this set does not necessarily have the structure of a manifold. For example, [31, Ex. 2.9] describes a case where the solution might be arbitrarily complex in the parameter (including discontinuous behavior). The crucial assumption for MOR to be reasonable is the following.

**Assumption 3.2.** *Given a metric $d_{\mathcal{M}} \colon \mathcal{M} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$, we assume that there exists a low-dimensional embedded submanifold $\varphi(\check{\mathcal{M}}) \subseteq \mathcal{M}$ defined by an $n$-dimensional manifold $\check{\mathcal{M}}$ and a smooth embedding $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$ with $\dim(\check{\mathcal{M}}) = n \ll N = \dim(\mathcal{M})$ such that the set of solutions $S$ can be approximated well, i.e.,*

$$d_{\mathcal{M}}(\varphi(\check{\mathcal{M}}), S) := \sup_{m \in S} \inf_{\check{m} \in \check{\mathcal{M}}} d_{\mathcal{M}}(m, \varphi(\check{m}))$$

*is small.*

We refer to $\mathcal{M}$ as the *full(-order) manifold* and to $\check{\mathcal{M}}$ as the *reduced(-order) manifold*. Let us emphasize that the goal is to approximate the set $S \subseteq \mathcal{M}$, not the full manifold $\mathcal{M}$. We refer to Figure 2 for a schematic illustration of the relation between the full manifold $\mathcal{M}$, the set of all solutions $S$, and the approximating embedded submanifold $\varphi(\check{\mathcal{M}})$.

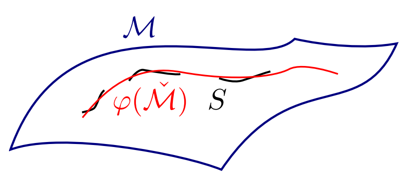



**Figure 2** – Schematic illustration of the full manifold $\mathcal{M}$ (dark blue), the set of all solutions $S$ (black), and the approximating embedded submanifold $\varphi(\check{\mathcal{M}})$ (red). The set of solutions is schematically depicted as three separate trajectories that may occur due to a possible discontinuous behavior in the parameter $\mu$.

3.1.3. *Reduction map and reduced-order model.* Assume that we have identified an $n$-dimensional embedded submanifold $\varphi(\check{\mathcal{M}}) \subseteq \mathcal{M}$ with smooth embedding $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$ and that Assumption 3.2 is satisfied. To find a ROM, we want to replace $\gamma\,(t)$ in (3.1) with the approximation $\varphi\,(\check{\gamma}\,(t))$ based on a reduced integral curve $\check{\gamma} \in \mathcal{C}^\infty(\mathcal{I}, \check{\mathcal{M}})$. Note that even if we would have an exact reproduction, i.e., $\gamma\,(t) = \varphi\,(\check{\gamma}\,(t))$ for all $t \in \mathcal{I}$, the initial value problem (3.1) in the reduced integral curve $\check{\gamma}$ would be overdetermined, in the sense that we have (locally in each chart) $N$ equations for $n$ unknowns. Thus, we must also reduce the initial value problem and give the following definition.

**Definition 3.3** (Reduction map)**.** *A map* $R \in \mathcal{C}^\infty(T\mathcal{M}, T\check{\mathcal{M}})$ *is called* reduction map *for a smooth embedding* $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$ *if it satisfies the* projection property

$$R \circ \mathrm{d}\varphi = \mathrm{id}_{T\check{\mathcal{M}}}. \tag{3.3}$$

*As in Section 2.5, we split the reduction map*

$$R \in \mathcal{C}^\infty(T\mathcal{M}, T\check{\mathcal{M}}), \qquad (m, v) \mapsto (\varrho(m), R|_m(v))$$

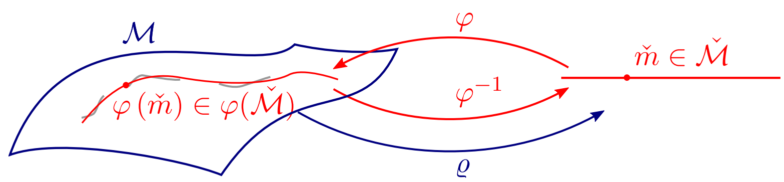
*with* $\varrho \in \mathcal{C}^\infty(\mathcal{M}, \check{\mathcal{M}})$ *and* $R|_m \in \mathcal{C}^\infty(T_m\mathcal{M}, T_{\varrho(m)}\check{\mathcal{M}})$ *for* $m \in \mathcal{M}$*. We refer to* $\varrho$ *as a* point reduction *for* $\varphi$ *and to* $R|_m$ *as a* tangent reduction *for* $\varphi$*.*

Note that (3.3) immediately implies that $\mathrm{d}\varphi \circ R \in \mathcal{C}^\infty(T\mathcal{M}, T\varphi(\check{\mathcal{M}}))$ is idempotent and thus a projection. Moreover, (3.3) implies that a point reduction and a tangent reduction for $\varphi$ satisfy

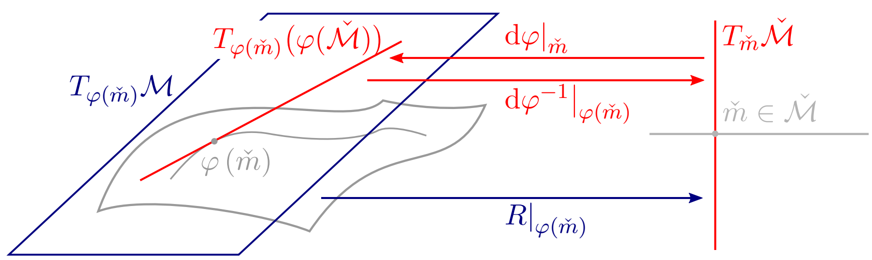
$$\varrho \circ \varphi = \mathrm{id}_{\check{\mathcal{M}}}, \tag{3.4a}$$

$$R|_{\varphi(\check{m})} \circ \mathrm{d}\varphi|_{\check{m}} = \mathrm{id}_{T_{\check{m}}\check{\mathcal{M}}} \qquad \text{for all } \check{m} \in \check{\mathcal{M}}, \tag{3.4b}$$

which we refer to as the *point projection property* and the *tangent projection property*, respectively. The relation between the embedding $\varphi$ and the reduction map $R$ is illustrated in Figure 3.




**(a)** Schematic illustration of the relation of the embedding $\varphi$ and the point reduction $\varrho$.




**(b)** Schematic illustration of the relation of the tangent spaces involved in MOR on manifolds. The reduced tangent space $T_{\check{m}}\check{\mathcal{M}}$ is displayed orthogonally to $\check{\mathcal{M}}$ for a better visualization.

**Figure 3** – Schematic illustration of the relation between the embedding $\varphi$ and the reduction map $R(m, v) = (\varrho(m), R|_m(v))$ with $m \in \mathcal{M}$.

**Example 3.4** (Linear-subspace MOR)**.** *Projection-based linear-subspace MOR with a reduced-basis matrix* $\boldsymbol{V} \in \mathbb{R}^{N \times n}$ *and a projection matrix* $\boldsymbol{W} \in \mathbb{R}^{N \times n}$ *is contained as a special case of the presented formulation with* $\mathcal{M} = U = \mathbb{R}^N$*,* $\check{\mathcal{M}} = \check{U} = \mathbb{R}^n$*,* $x = \mathrm{id}_{\mathbb{R}^N}$*,* $\check{x} = \mathrm{id}_{\mathbb{R}^n}$ *and*

$$\boldsymbol{\varrho}(\boldsymbol{m}) := \boldsymbol{W}^\top \boldsymbol{m}, \qquad \boldsymbol{R}|_{\boldsymbol{m}}(\boldsymbol{v}) := \boldsymbol{W}^\top \boldsymbol{v}, \qquad \boldsymbol{\varphi}(\check{\boldsymbol{m}}) := \boldsymbol{V}\check{\boldsymbol{m}}.$$

*This exactly covers the case where* $\boldsymbol{\varphi}$ *and* $\boldsymbol{\varrho}$ *are linear. The projection property (3.3) then relates to the* biorthogonality *of* $\boldsymbol{W}$ *and* $\boldsymbol{V}$

$$\boldsymbol{\varrho} \circ \boldsymbol{\varphi} \equiv \mathrm{id}_{\mathbb{R}^n} \qquad \iff \qquad \boldsymbol{W}^\top \boldsymbol{V} = \boldsymbol{I}_n \in \mathbb{R}^{n \times n},$$

$$\boldsymbol{R}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})} \circ \mathrm{d}\boldsymbol{\varphi}|_{\check{\boldsymbol{m}}} \equiv \mathrm{id}_{\mathbb{R}^n} \qquad\qquad \iff \qquad\qquad \boldsymbol{W}^\top \boldsymbol{V} = \boldsymbol{I}_n \in \mathbb{R}^{n\times n},$$

*which is often assumed in linear-subspace* MOR.

**Definition 3.5** (Reduced-order model)**.** *Consider the* FOM (3.1)*, a smooth embedding* $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$*, and a reduction map* $R \in \mathcal{C}^\infty(T\mathcal{M}, T\check{\mathcal{M}})$ *for* $\varphi$ *with point and tangent reduction for* $\varphi$ *given by* $R\,(m, v) = (\varrho\,(m)\,, R|_m\,(v))$*. We define the* ROM vector field *as* $\check{X} \colon \mathcal{P} \to \mathfrak{X}_{\check{\mathcal{M}}}$ *via*

$$\check{X}(\mu)\Big|_{\check{m}} := R|_{\varphi(\check{m})} \left( X(\mu)|_{\varphi(\check{m})} \right) \in T_{\check{m}}\check{\mathcal{M}}.$$

*Then, for* $\mu \in \mathcal{P}$*, we call the initial value problem on* $\check{\mathcal{M}}$

$$\begin{cases} \quad \frac{\mathrm{d}}{\mathrm{d}t}\check{\gamma}\Big|_{t;\mu} = \check{X}(\mu)\Big|_{\check{\gamma}(t;\mu)} \in T_{\check{\gamma}(t;\mu)}\check{\mathcal{M}} \\ \quad \check{\gamma}\,(t_0;\mu) = \check{\gamma}_0(\mu) := \varrho\,(\gamma_0\,(\mu)) \in \check{\mathcal{M}} \end{cases} \tag{3.5}$$

*the* ROM *for* (3.1) *under the reduction map* $R$ *with solution* $\check{\gamma}(\cdot;\mu) \in \mathcal{C}^\infty(\mathcal{I}, \check{\mathcal{M}})$*.*

We emphasize that both, the point and the tangent reduction, are relevant for the ROM, since the point reduction is used to map the initial value $\gamma_0$, while the tangent reduction maps the FOM vector field to the tangent space of the reduced manifold $\check{\mathcal{M}}$. Moreover, we see that it is not sufficient to define $\varrho$ and $R|_{\varphi(\check{m})}$ only in the image of $\varphi$ and $\mathrm{d}\varphi|_{\check{m}}$, respectively, since the initial value and the evaluated FOM vector field may be elements of $\mathcal{M} \setminus \varphi\,(\check{\mathcal{M}})$ and $T_{\varphi(\check{m})}\mathcal{M} \setminus T_{\varphi(\check{m})}(\varphi(\check{\mathcal{M}}))$, respectively.

3.1.4. *MOR workflow.* With Assumption 3.2 at hand, MOR (in the scope of this work) can be summarized in three steps:

- (i) Approximation: Given the FOM (3.1), find a reduced manifold $\check{\mathcal{M}}$ and a smooth embedding $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$ such that $d_{\mathcal{M}}(\varphi(\check{\mathcal{M}}), S)$ is small.
- (ii) Reduction: Identify a reduction map $R \in \mathcal{C}^\infty(T\mathcal{M}, T\check{\mathcal{M}})$ for $\varphi$ and construct the ROM (3.5).
- (iii) Reconstruction: Solve the ROM (3.5) for $\check{\gamma}$ and approximate the FOM solution curve $\gamma$ with

$$\gamma\,(t;\mu) \approx \varphi\,(\check{\gamma}\,(t;\mu)) \qquad \text{for } (t,\mu) \in \mathcal{I} \times \mathcal{P}. \tag{3.6}$$

In the remainder of the manuscript, we discuss all three steps, starting with the Reconstruction step in the subsequent subsection. Possible constructions of the reduction map in the Reduction step are discussed in Sections 3.3 and 5. The construction of the embedding $\varphi$ in the Approximation step is analyzed in a data-driven framework in Section 6.

3.2. **Exact reproduction.** A desirable property in the Reconstruction step is to answer the question when the approximation in (3.6) is exact, which we refer to as *exact reproduction*. Clearly, if for a given parameter $\mu \in \mathcal{P}$, the FOM solution $\gamma$ evolves on $\varphi(\check{\mathcal{M}})$, i.e., $\gamma\,(t;\mu) \in \varphi(\check{\mathcal{M}})$ for all $t \in \mathcal{I}$, then we can define the smooth curve

$$\check{\beta} := \varphi^{-1}(\gamma\,(\cdot;\mu)) \in \mathcal{C}^\infty(\mathcal{I}, \check{\mathcal{M}}), \tag{3.7}$$

since, by assumption, $\varphi$ is a diffeomorphism onto its image. With this choice, we immediately obtain

$$X(\mu)|_{\varphi(\check{\beta}(t;\mu))} = X(\mu)|_{\gamma(t;\mu)} = \tfrac{\mathrm{d}}{\mathrm{d}t}\gamma\Big|_{t;\mu} = \tfrac{\mathrm{d}}{\mathrm{d}t}(\varphi \circ \beta)\Big|_{t;\mu} = \mathrm{d}\varphi|_{\check{\beta}(t;\mu)}\left(\tfrac{\mathrm{d}}{\mathrm{d}t}\check{\beta}\Big|_{t;\mu}\right), \tag{3.8}$$

where the last equality follows from the chain rule (2.5). It remains to prove that the ROM (3.5) is able to recover the reduced curve $\check{\beta}$, which we show in the following.

**Theorem 3.6** (Exact reproduction of a solution)**.** *Assume that the FOM (3.1) is uniquely solvable and consider a reduction map $R \in \mathcal{C}^\infty(T\mathcal{M}, T\check{\mathcal{M}})$ for the smooth embedding $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$ and a parameter $\mu \in \mathcal{P}$. Assume that the ROM (3.5) is uniquely solvable and $\gamma(t;\mu) \in \varphi(\check{\mathcal{M}})$ for all $t \in \mathcal{I}$. Then the ROM solution $\check{\gamma}(\cdot;\mu)$ exactly recovers the solution $\gamma(\cdot;\mu)$ of the FOM (3.1) for this parameter, i.e.,*

$$\varphi(\check{\gamma}(t;\mu)) = \gamma(t;\mu) \qquad\qquad \textit{for all } t \in \mathcal{I}. \tag{3.9}$$

*Proof.* Since $\gamma(t;\mu) \in \varphi(\check{\mathcal{M}})$ for all $t \in \mathcal{I}$, we can construct $\check{\beta}$ as in (3.7). It remains to show that $\check{\beta}$ satisfies the ROM (3.5). First, we obtain

$$\check{\gamma}_0(\mu) = \varrho(\gamma_0(\mu)) = \varrho(\gamma(t_0;\mu)) = (\varrho \circ \varphi)\left(\check{\beta}(t_0;\mu)\right) = \check{\beta}(t_0;\mu),$$

where the last equality is due to the projection property (3.4) for the point reduction. Second, $\check{\beta}$ suffices the initial value problem of the ROM since the tangent projection property (3.4b) implies with (3.8)

$$R|_{\varphi(\check{\beta}(t;\mu))}\left(X(\mu)|_{\varphi(\check{\beta}(t;\mu))}\right) = \left(R|_{\varphi(\check{\beta}(t;\mu))} \circ \mathrm{d}\varphi|_{\check{\beta}(t;\mu)}\right)\left(\tfrac{\mathrm{d}}{\mathrm{d}t}\check{\beta}\Big|_{t;\mu}\right) = \tfrac{\mathrm{d}}{\mathrm{d}t}\check{\beta}\Big|_{t;\mu}. \qquad \square$$

In the following, we give an example of for which the exact reproduction can be achieved for a specific choice of $\check{\mathcal{M}}$ and $\varphi$.

**Corollary 3.7** (Special case $\varphi \equiv \gamma$)**.** *For a given FOM (3.1) on $\mathcal{M}$, assume that $\check{\mathcal{M}} = \mathcal{I} \times \mathcal{P}$ is an $(n_p+1)$-dimensional smooth manifold, that the FOM is uniquely solvable, that the FOM solution $\gamma\colon \mathcal{I} \times \mathcal{P} =: \check{\mathcal{M}} \to \mathcal{M}$ is a smooth embedding, and that there exists a reduction map $R \in \mathcal{C}^\infty(T\mathcal{M}, T\check{\mathcal{M}})$. Then, the ROM (3.5) reproduces the FOM solution exactly with the reduced integral curve $\check{\gamma}(t;\mu) = (t,\mu)$ such that the flow of the ROM is $\check{\theta}_t \equiv \mathrm{id}_{\check{\mathcal{M}}}$. Moreover, the ROM in bold notation reads*

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\check{\boldsymbol{\gamma}}\Big|_{t;\mu} = \boldsymbol{e}_1 \in \mathbb{R}^{n_p+1}, \qquad\qquad \check{\boldsymbol{\gamma}}_0(\mu) = (t_0, \mu),$$

*where $\boldsymbol{e}_1 \in \mathbb{R}^{n_p+1}$ denotes the first unit vector.*

*Proof.* With the assumptions of Corollary 3.7, the choice $\varphi \equiv \gamma$ guarantees that the assumptions of Theorem 3.6 are fulfilled and $\check{\beta}(t;\mu) = (t,\mu)$ is a valid choice for the curve in (3.7), which was used in the proof of Theorem 3.6 as the ROM solution candidate. For the remaining statement, we observe

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\check{\beta}^1\Big|_{t;\mu} = 1, \qquad\qquad \tfrac{\mathrm{d}}{\mathrm{d}t}\check{\beta}^i\Big|_{t;\mu} = 0 \qquad \text{for } 1 < i \leq n_p + 1,$$

and $\check{\gamma}_0(\mu) = \check{\beta}(t_0;\mu) = (t_0,\mu)$, which completes the proof. $\square$

**Example 3.8.** *A particular example describing the situation from Corollary 3.7 is given by the linear advection equation with constant coefficients and periodic boundary conditions, see for instance [10, Ex. 5.12].*

3.3. **Manifold Petrov–Galerkin (MPG).** Now we want to address one example of how to construct a reduction map (▷ Definition 3.3), i.e., how to do the REDUCTION step from the general MOR workflow described in Section 3.1.4. Note that this specific choice of reduction map has been independently developed in [57]. Moreover, we emphasize that the specific choice of the reduction map is crucial for the approximation quality of the ROM; see, for instance, [58]. Nevertheless, our goal here is not to present an optimal choice but rather an example of leveraging the smooth embedding $\varphi$ to construct a reduction map using the previously introduced framework in Section 3.1.

Assume that we have completed the APPROXIMATION step from the general MOR workflow, i.e., we have already identified a reduced manifold $\check{\mathcal{M}}$ together with a smooth embedding $\varphi$. Since $\varphi$ is a homeomorphism onto its image, we know that $\varphi^{-1}\colon \varphi(\check{\mathcal{M}}) \to \check{\mathcal{M}}$ exists. Under for MOR reasonable assumptions, the extension lemma for smooth functions (see for instance [45, Lem. 2.26]) guarantees that we can find a smooth extension $\varrho$ of $\varphi^{-1}$, which by construction satisfies the point projection property (3.4a). We refer to Figure 3a for an illustration of the relation between $\varphi^{-1}$ and $\varrho$. Differentiating the point projection property (3.4a) with the chain rule (2.5) implies

$$\mathrm{d}\varrho|_{\varphi(\check{m})} \circ \mathrm{d}\varphi|_{\check{m}} = \mathrm{d}(\mathrm{id}_{\check{\mathcal{M}}})|_{\check{m}} = \mathrm{id}_{T_{\check{m}}\check{\mathcal{M}}}\colon T_{\check{m}}\check{\mathcal{M}} \to T_{\check{m}}\check{\mathcal{M}}, \tag{3.10}$$

i.e., $\mathrm{d}\varrho|_{\varphi(\check{m})}$ is a left-inverse to $\mathrm{d}\varphi|_{\check{m}}$. In particular, we have proven the following duality result.

**Theorem 3.9** (MPG reduction map)**.** *Consider a smooth embedding $\varphi$ and a point reduction $\varrho$ for $\varphi$. Then, the differential of the point reduction $\varrho$ is a left inverse to the differential of the embedding $\varphi$. Consequently,*

$$R_{\mathsf{MPG}}\colon T\mathcal{M} \to T\check{\mathcal{M}} \qquad (m, v) \mapsto (\varrho(m), \mathrm{d}\varrho|_m(v)) \tag{3.11}$$

*is a smooth reduction map for $\varphi$, which we call the* MPG reduction map *for $(\varrho, \varphi)$.*

We refer to the ROM (3.5) obtained with the MPG reduction map from Theorem 3.9 as the *MPG-ROM* for $(\varrho, \varphi)$. In index and bold notation, the tangent projection property (3.10) reads

$$\frac{\partial \varrho^{i}}{\partial x^{\underline{k}}}\Big|_{\varphi(\check{m})} \frac{\partial \varphi^{\underline{k}}}{\partial \check{x}^{\underline{j}}}\Big|_{\check{m}} = \delta^{i}_{\underline{j}}, \qquad\qquad \boldsymbol{D\varrho}|_{\varphi(\check{\boldsymbol{m}})} \boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}} = \boldsymbol{I}_n \in \mathbb{R}^{n\times n}. \tag{3.12}$$

It can be interpreted as that the columns of $\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}$ span an $n$-dimensional reduced vector space that changes with the reduced coordinates $\check{\boldsymbol{m}} \in \mathbb{R}^n$, whereas the rows of $\boldsymbol{D\varrho}|_{\varphi(\check{\boldsymbol{m}})}$ span an $n$-dimensional vector space dual to the reduced vector space.

**Example 3.10** (Linear-subspace MOR)**.** *If $\varphi$ and $\varrho$ are linear as in Example 3.4, then the MPG-ROM (3.5) with the MPG reduction map from Theorem 3.9 is the ROM obtained in classical linear-subspace MOR via Petrov–Galerkin projection*

$$\boldsymbol{R}_{\mathsf{MPG}}|_{\varphi(\check{m})} = \boldsymbol{D\varrho}|_{\varphi(\check{m})} = \boldsymbol{W}^{\top}, \qquad\qquad \frac{\mathrm{d}}{\mathrm{d}t}\check{\boldsymbol{\gamma}}\Big|_{t} = \boldsymbol{W}^{\top}\boldsymbol{X}|_{\check{\boldsymbol{\gamma}}(t)},$$

*which is the motivation for the terminology MPG.*

## 4. MANIFOLDS WITH STRUCTURE

As a next step, we want to discuss structure-preserving MOR on manifolds (▷ Section 5). Beforehand, we specify the relevant structures on the FOM level in the present section. The idea is to equip the underlying full manifold $\mathcal{M}$ with additional structure to formulate a FOM vector field $X$, which guarantees physical properties, e.g., that the FOM solutions preserve energy over time. We introduce additional structure on $\mathcal{M}$ (▷Section 4.1), which allows us to formulate Lagrangian systems (▷Section 4.2) and Hamiltonian systems (▷Section 4.3) on manifolds. Both systems admit a FOM vector field, which guarantees that the FOM solutions preserve the corresponding energy over time.

4.1. **Additional structure on $\mathcal{M}$.** To keep this work self-contained, we proceed by detailing more concepts of differential geometry. We discuss the cotangent space and covectors (▷Section 4.1.1), tensors (▷Section 4.1.2), tensor fields (▷Section 4.1.3), structured tensor fields (▷Section 4.1.4), and pullbacks of covectors, tensor fields, and functions (▷Section 4.1.5).

4.1.1. *Cotangent space, covectors, and cotangent bundle.* The dual of the tangent space at $m \in \mathcal{M}$ (2.3) is the *cotangent space at $m \in \mathcal{M}$*

$$T_m^*\mathcal{M} := \{\lambda \mid \lambda\colon T_m\mathcal{M} \to \mathbb{R} \text{ linear}\},$$

which is again an $N$-dimensional vector space. Elements in the cotangent space are called *cotangent vectors* or simply *covectors*. Covectors can be constructed from scalar-valued functions: For each scalar-valued function $f \in \mathcal{C}^\infty(\mathcal{M}, \mathbb{R})$, its differential at $m$, $\mathrm{d}f|_m \in \mathcal{C}^\infty(T_m\mathcal{M}, T_{f(m)}\mathbb{R})$, defines a linear functional on $T_m\mathcal{M}$ if we identify $T_{f(m)}\mathbb{R}$ with $\mathbb{R}$. Thus, the differential at $m$ of a scalar-valued function is a covector $\mathrm{d}f|_m \in T_m^*\mathcal{M}$. For a given chart $(U, x)$ of $\mathcal{M}$, this construction can be used to define a basis of $T_m^*\mathcal{M}$: For each $i \in \{1, \dots, N\}$, the $i$-th component function of the chart mapping $x^{\underline{i}} \in \mathcal{C}^\infty(U, \mathbb{R})$ is a scalar-valued function and thus $\mathrm{d}x^{\underline{i}}|_m \in T_m^*\mathcal{M}$. Moreover, with (2.2), (2.4), and identifying $T_{x(m)}\mathbb{R} \cong \mathbb{R}$, it holds for all basis vectors of the tangent space $\frac{\partial}{\partial x^{\underline{j}}}\Big|_m \in T_m\mathcal{M}$, $1 \le j \le N$ the dual relationship

$$\mathrm{d}x^{\underline{i}}\Big|_m \left(\frac{\partial}{\partial x^{\underline{j}}}\Big|_m\right) = \frac{\partial x^{\underline{i}}}{\partial x^{\underline{j}}}\Big|_m = \delta^{\underline{i}}_{\underline{j}} \in \mathbb{R} \cong T_{x(m)}\mathbb{R}$$

The differentials $\{\mathrm{d}x^{\underline{i}}|_m\}_{1 \le i \le N}$ define a basis of $T_m^*\mathcal{M}$ and we can represent each covector $\lambda \in T_m^*\mathcal{M}$ as

$$\lambda = \lambda_{\underline{i}}\, \mathrm{d}x^{\underline{i}}\Big|_m \in T_m^*\mathcal{M},$$

with *(covector) components* $\lambda_{\underline{i}} \in \mathbb{R}$, where the right-hand side sums over $1 \le i \le N$ by Einstein summation convention (1.3). By the duality of the bases of $T_m\mathcal{M}$ and $T_m^*\mathcal{M}$, it holds for each covector $\lambda \in T_m^*\mathcal{M}$ and vector $v \in T_m\mathcal{M}$ that

$$\lambda(v) = \left(\lambda_{\underline{j}}\, \mathrm{d}x^{\underline{j}}\right)\left(v^{\underline{i}} \frac{\partial}{\partial x^{\underline{i}}}\Big|_m\right) = \lambda_{\underline{j}}\, v^{\underline{i}}\, \mathrm{d}x^{\underline{j}}\Big|_m \left(\frac{\partial}{\partial x^{\underline{i}}}\Big|_m\right) = \lambda_{\underline{i}}\, v^{\underline{i}} \in \mathbb{R}.$$

Analogously to the tangent bundle (2.6), a *cotangent bundle* $T^*\mathcal{M}$ can be formulated as the disjoint union of $T_m^*\mathcal{M}$, which can be shown to be a smooth manifold of dimension $2N$.

4.1.2. *Tensors.* A generalization of vectors and covectors are the so-called *tensors*. For a vector space $\mathcal{V}$ and its dual $\mathcal{V}^*$, the *space of $(r, s)$-tensors* given by

$$T^{(r,s)}(\mathcal{V}) := \underbrace{\mathcal{V} \otimes \cdots \otimes \mathcal{V}}_{r \text{ times}} \otimes \underbrace{\mathcal{V}^* \otimes \cdots \otimes \mathcal{V}^*}_{s \text{ times}}.$$

In the present work we consider tensors on the tangent and cotangent space, i.e., $\mathcal{V} = T_m\mathcal{M}$ and $\mathcal{V}^* = T_m^*\mathcal{M}$. Special cases are $T^{(1,0)}(T_m\mathcal{M}) = T_m\mathcal{M}$ and $T^{(0,1)}(T_m\mathcal{M}) = T_m^*\mathcal{M}$. An element $\sigma \in T^{(r,s)}(T_m\mathcal{M})$ of a general $(r,s)$-tensor space is called an *$r$-times contravariant $s$-times covariant tensor*. This element can be represented by

$$\sigma = \sigma^{i_1 \dots i_r}_{\underline{j_1 \dots j_s}} \left.\frac{\partial}{\partial x^{\underline{i_1}}}\right|_m \otimes \dots \otimes \left.\frac{\partial}{\partial x^{\underline{i_r}}}\right|_m \otimes \left.\mathrm{d}x^{\underline{j_1}}\right|_m \otimes \dots \otimes \left.\mathrm{d}x^{\underline{j_s}}\right|_m$$

with *components* $\sigma^{i_1 \dots i_r}_{\underline{j_1 \dots j_s}} \in \mathbb{R}$ for $1 \leq i_1, \dots, i_r, j_1, \dots, j_s \leq N$, where the right-hand side sums over each index $1 \leq i_1, \dots, i_r, j_1, \dots, j_s \leq N$ by the Einstein summation convention (1.3). The position of the index (upper or lower index) indicates which type (co- or contravariant) the respective index belongs to. To extend the bold notation from Section 2.7 for tensors, we stack the components with

$$\boldsymbol{\sigma} := \left[\sigma^{i_1 \dots i_r}_{\underline{j_1 \dots j_s}}\right]_{1 \leq i_1, \dots, i_r, j_1, \dots, j_s \leq N} \in \mathbb{R}^{\overbrace{N \times N \times \dots \times N}^{r+s \text{ times}}}.$$

4.1.3. *Tensor field and bundle of $(r,s)$-tensors.* A so-called *tensor field* is a mapping which assigns each point $m \in \mathcal{M}$ a tensor in the corresponding $(r,s)$-tensor space $T^{(r,s)}(T_m\mathcal{M})$ analogous to the smooth vector field introduced in Section 2.5. To this end we define the *bundle of $(r,s)$-tensors* as the disjoint union of all $(r,s)$-tensor spaces

$$T^{(r,s)}(T\mathcal{M}) := \dot{\bigcup}_{m \in \mathcal{M}} T^{(r,s)}(T_m\mathcal{M}) = \{(m, \sigma) \mid m \in \mathcal{M},\ \sigma \in T^{(r,s)}(T_m\mathcal{M})\}.$$

Similarly as before, we obtain the special cases $T^{(1,0)}(T\mathcal{M}) = T\mathcal{M}$ and $T^{(0,1)}(T\mathcal{M}) = T^*\mathcal{M}$. An $(r,s)$-*tensor field* is defined as a map

$$\tau \colon \mathcal{M} \to T^{(r,s)}(T\mathcal{M}),\ m \mapsto (m, \tau|_m) \qquad \text{such that } \tau|_m \in T^{(r,s)}(T_m\mathcal{M}).$$

For a given chart $(U,x)$ of $\mathcal{M}$, we denote the $((r+s) \cdot N)$ functions $\tau^{i_1 \dots i_r}_{\underline{j_1 \dots j_s}} \colon U \to \mathbb{R}$, $1 \leq i_1, \dots, i_r, j_1, \dots, j_s \leq N$, with $\tau^{i_1 \dots i_r}_{\underline{j_1 \dots j_s}}(m) := (\tau|_m)^{i_1 \dots i_r}_{\underline{j_1 \dots j_s}}$ as the *component functions* to stress the dependence on the point $m$. To extend the bold notation from Section 2.7 for tensor fields, we stack the component functions with

$$\boldsymbol{\tau} := \left[\tau^{i_1 \dots i_r}_{\underline{j_1 \dots j_s}} \circ x^{-1}\right]_{1 \leq i_1, \dots, i_r, j_1, \dots, j_s \leq N} : \mathbb{R}^N \supseteq x(U) \to \mathbb{R}^{\overbrace{N \times N \times \dots \times N}^{r+s \text{ times}}}.$$

An $(r,s)$-tensor field $\tau$ is called *smooth* if all of its component functions are smooth, i.e., $\tau^{i_1 \dots i_r}_{\underline{j_1 \dots j_s}} \in \mathcal{C}^\infty(U, \mathbb{R})$. The set of all smooth $(r,s)$-tensor fields is the so-called *smooth section of the $(r,s)$-tensor bundle* $\Gamma(T^{(r,s)}(T\mathcal{M}))$. A special case are the smooth vector fields $\mathfrak{X}_\mathcal{M} = \Gamma(T^{(1,0)}(T\mathcal{M}))$.

4.1.4. *Structured tensor fields and musical isomorphisms.* Tensor fields may possess additional properties, which we refer to as *structure*. In the following, we introduce two important examples of tensor fields with special structures, namely Riemannian metrics and symplectic forms.

For a smooth $(0,2)$-tensor field $\tau \in \Gamma(T^{(0,2)}(T\mathcal{M}))$ with its component functions $\tau_{\underline{ij}} \in \mathcal{C}^\infty(U, \mathbb{R})$ for $1 \leq i,j \leq N$ in a given chart $(U,x)$, the tensor field $\tau$ is called

- *symmetric*, if $(\tau|_m)_{\underline{ij}} = (\tau|_m)_{\underline{ji}}$ for each $m \in U$ and for all $1 \leq i,j \leq N$;
- *skew-symmetric* or *2-form*, if $(\tau|_m)_{\underline{ij}} = -(\tau|_m)_{\underline{ji}}$ for each $m \in U$ and for all $1 \leq i,j \leq N$;
- *nondegenerate*, if $[(\tau|_m)_{\underline{ij}}]_{1 \leq i,j \leq N} \in \mathbb{R}^{N \times N}$ is nondegenerate for each $m \in U$;

- *positive definite*, if $[(\tau|_m)_{\underline{ij}}]_{1 \leq i,j \leq N} \in \mathbb{R}^{N \times N}$ is positive definite for all $m \in U$;
- a *closed* 2*-form*, if $\tau$ is a 2-form and for each $m \in U$

$$\left.\frac{\partial \tau_{\underline{jk}}}{\partial x^{\underline{i}}}\right|_m + \left.\frac{\partial \tau_{\underline{ki}}}{\partial x^{\underline{j}}}\right|_m + \left.\frac{\partial \tau_{\underline{ij}}}{\partial x^{\underline{k}}}\right|_m = 0 \qquad \text{for all } 1 \leq i \leq j \leq k \leq N. \tag{4.1}$$

Combining some of the previous properties, we obtain the following concepts. A smooth $(0,2)$-tensor field $\tau \in \Gamma(T^{(0,2)}(T\mathcal{M}))$ on $\mathcal{M}$ is called

- a *Riemannian metric on* $\mathcal{M}$ if $\tau$ is symmetric and positive definite;
- a *symplectic form on* $\mathcal{M}$ if $\tau$ is skew-symmetric, nondegenerate, and closed.

If $\tau, \omega \in \Gamma(T^{(0,2)}(T\mathcal{M}))$ are a Riemannian metric and a symplectic form on $\mathcal{M}$, respectively, then we call $(\mathcal{M}, \tau)$ and $(\mathcal{M}, \omega)$ a *Riemannian manifold* and *symplectic manifold*, respectively. Note that the nondegeneracy of a symplectic form implies that a symplectic manifold has even dimension.

Both the Riemannian metric and the symplectic form are nondegenerate tensor fields. This allows to formulate the inverse $(2,0)$-tensor field $\tau^{-1} \in \Gamma(T^{(2,0)}(T\mathcal{M}))$ such that $(\tau|_m^{-1})^{\underline{ik}}(\tau|_m)_{\underline{kj}} = \delta^{\underline{i}}_{\underline{j}}$, where the components are typically denoted with $(\tau|_m)^{\underline{ik}} := (\tau|_m^{-1})^{\underline{ik}}$ for the sake of brevity. Moreover, the nondegeneracy allows to formulate an isomorphism between the tangent and the cotangent bundle. Loosely speaking, this means that the indices in the index notation can be switched from covariant (superindices) to contravariant (subindices) and vice versa. This is typically referred to as *musical isomorphisms*

$$\flat_\tau \in \mathcal{C}^\infty(T\mathcal{M}, T^*\mathcal{M}), \quad \left(m,\, v^{\underline{i}} \left.\tfrac{\partial}{\partial x^{\underline{i}}}\right|_m\right) \mapsto \left(m,\, (\tau|_m)_{\underline{ij}}\, v^{\underline{j}}\, \left.\mathrm{d}x^{\underline{i}}\right|_m\right), \tag{4.2}$$

$$\sharp_\tau \in \mathcal{C}^\infty(T^*\mathcal{M}, T\mathcal{M}), \quad \left(m,\, \lambda_{\underline{i}}\, \left.\mathrm{d}x^{\underline{i}}\right|_m\right) \mapsto \left(m,\, (\tau|_m)^{\underline{ij}}\, \lambda_{\underline{j}}\, \left.\tfrac{\partial}{\partial x^{\underline{i}}}\right|_m\right). \tag{4.3}$$

Due to the nondegeneracy of $\tau$, the two mappings are inverses of each other, i.e.,

$$\sharp_\tau \circ \flat_\tau \equiv \mathrm{id}_{T\mathcal{M}}. \tag{4.4}$$

By a slight abuse of notation, we use the same symbols from (4.2) and (4.3) also to map between (co)tangent spaces $\flat_\tau \colon T_m\mathcal{M} \to T_m^*\mathcal{M}$ and $\sharp_\tau \colon T_m^*\mathcal{M} \to T_m\mathcal{M}$ (instead of the respective bundles).

4.1.5. *Pullback of covectors, tensor fields, and functions.* Consider two smooth manifolds $\mathcal{M}$, $\mathcal{Q}$ and a smooth map $F \in \mathcal{C}^\infty(\mathcal{M}, \mathcal{Q})$. Let $(U, x)$ and $(V, y)$ be charts of $\mathcal{M}$ and $\mathcal{Q}$ respectively such that $m \in U$ and $F(m) \in V$. The differential (2.4) of $F$ can be used to define the *pointwise pullback (of covectors) by* $F$ *at* $m$ via

$$\mathrm{d}F^*|_m \in \mathcal{C}^\infty\left(T^*_{F(m)}\mathcal{Q}, T^*_m\mathcal{M}\right), \quad \lambda_{\underline{i}}\, \left.\mathrm{d}y^{\underline{i}}\right|_{F(m)} \mapsto \left.\tfrac{\partial F^{\underline{i}}}{\partial x^{\underline{j}}}\right|_m \lambda_{\underline{i}}\, \left.\mathrm{d}x^{\underline{j}}\right|_m. \tag{4.5}$$

For a smooth $(0,s)$-tensor field $\tau \in \Gamma(T^{(0,s)}(T\mathcal{Q}))$, the *pullback of* $\tau$ *by* $F$, denoted by $F^*\tau \in \Gamma(T^{(0,s)}(T\mathcal{M}))$, is a smooth tensor field (see [45, Prop. 11.26]) with component functions[6]

$$(F^*\tau|_m)_{\underline{j_1 \ldots j_s}} := (\tau|_{F(m)})_{\underline{\ell_1 \ldots \ell_s}} \cdot \left.\frac{\partial F^{\underline{\ell_1}}}{\partial x^{\underline{j_1}}}\right|_m \cdots \left.\frac{\partial F^{\underline{\ell_s}}}{\partial x^{\underline{j_s}}}\right|_m. \tag{4.6}$$

[6] The pullbacks from (4.5) and (4.6) can be related in the case of smooth covector fields $\alpha \in \Gamma(T^{(0,1)}(T\mathcal{Q}))$, i.e., $s = 1$, with $(F^*\alpha)|_m = \mathrm{d}F^*|_m \alpha|_{F(m)} \in T_m^*\mathcal{M}$.

A scalar-valued smooth function $h \in \mathcal{C}^\infty(\mathcal{Q}, \mathbb{R})$ can be interpreted as a $(0,0)$-tensor field. Then, as a special case of (4.6), the *pullback of (a function) $h$ by $F$* is a smooth function $F^* h \in \mathcal{C}^\infty(\mathcal{M}, \mathbb{R})$ with

$$(F^* h)(m) = h(F(m)) = (h \circ F)(m). \tag{4.7}$$

By Section 4.1.1, the differential of a smooth scalar-valued function $G \in \mathcal{C}^\infty(\mathcal{Q}, \mathbb{R})$ defines a covector $\mathrm{d}G|_{F(m)} \in T_m^* \mathcal{Q}$. Then an analogue to the chain rule (2.5) is

$$\mathrm{d}(F^* G)|_m = \mathrm{d}F^*|_m \mathrm{d}G|_{F(m)} \in T_m^* \mathcal{M}, \tag{4.8}$$

which uses the pullback of a function (4.7) on the left-hand side and applies the pointwise pullback $\mathrm{d}F^*|_m \in \mathcal{C}^\infty(T_m^* \mathcal{Q}, T_m^* \mathcal{M})$ to the covector $\mathrm{d}G|_{F(m)} \in T_m^* \mathcal{Q}$ on the right-hand side of the equation.

4.2. **Lagrangian systems.** This subsection defines Lagrangian systems formulated on a manifold and additionally introduces further structure required for the MOR part discussed in the forthcoming Section 5.2. Consider a $Q$-dimensional smooth manifold $\mathcal{Q}$ with chart $(V, y)$. As mentioned in Section 2.5, the tangent bundle $T\mathcal{Q}$ is a $2Q$-dimensional smooth manifold and the differential $\mathrm{d}y \in \mathcal{C}^\infty(TV, \mathbb{R}^{2Q})$ defines a natural chart (2.8). We abbreviate this chart with $\xi := \mathrm{d}y$ for brevity. By (2.8), it holds

$$\xi \colon TV \to \mathbb{R}^{2Q}, \ \left( q, v^{\underline{i}} \left. \tfrac{\partial}{\partial y^{\underline{i}}} \right|_q \right) \mapsto \left( y(q), \left[ v^{\underline{i}} \right]_{1 \le i \le Q} \right).$$

It will be relevant to differentiate between the first $Q$ and the latter $Q$ entries of $\xi$ for a point $Y_{\mathcal{Q}} = (q, v) \in T\mathcal{Q}$, which will be denoted with

$$\xi^{\underline{i}}(Y_{\mathcal{Q}}) = y^{\underline{i}}(q), \qquad \xi^{\underline{Q+i}}(Y_{\mathcal{Q}}) = v^{\underline{i}}, \qquad \text{for } 1 \le i \le Q.$$

To lift a smooth curve $\gamma_{\mathcal{Q}} \in \mathcal{C}^\infty(\mathcal{I}, \mathcal{Q})$ to its tangent bundle, we define

$$\Gamma_{\gamma_{\mathcal{Q}}} \in \mathcal{C}^\infty(\mathcal{I}, T\mathcal{Q}), \ t \mapsto \left( \gamma_{\mathcal{Q}}(t), \left. \tfrac{\mathrm{d}}{\mathrm{d}t} \gamma_{\mathcal{Q}} \right|_t \right).$$

We denote a *Lagrangian system* as the tuple $(\mathcal{Q}, \mathscr{L})$ of a smooth manifold $\mathcal{Q}$ and a smooth function $\mathscr{L} \in \mathcal{C}^\infty(T\mathcal{Q}, \mathbb{R})$, which we refer to as the *Lagrangian function*. The associated second-order differential equation on the manifold is given by the *Euler–Lagrange equation*

$$\left. \tfrac{\partial \mathscr{L}}{\partial \xi^{\underline{i}}} \right|_{\Gamma_{\gamma_{\mathcal{Q}}}(t)} - \tfrac{\mathrm{d}}{\mathrm{d}t} \left( \left. \tfrac{\partial \mathscr{L}}{\partial \xi^{\underline{Q+i}}} \right|_{\Gamma_{\gamma_{\mathcal{Q}}}(\cdot)} \right) \Bigg|_t = 0 \qquad \text{for } 1 \le i \le Q, \qquad \Gamma_{\gamma_{\mathcal{Q}}}(t_0) = \begin{pmatrix} q_0 \\ v_0 \end{pmatrix}, \tag{4.9}$$

with *initial value* $(q_0, v_0) \in T\mathcal{Q}$, which has to be solved for $\gamma_{\mathcal{Q}} \in \mathcal{C}^\infty(\mathcal{I}, \mathcal{Q})$. In bold notation, the equation in coordinates reads for the Lagrangian $\boldsymbol{\mathscr{L}} := \mathscr{L} \circ \xi^{-1} \colon \mathbb{R}^{2Q} \supseteq \xi(TV) \to \mathbb{R}$, $(\boldsymbol{q}, \boldsymbol{v}) \mapsto \boldsymbol{\mathscr{L}}(\boldsymbol{q}, \boldsymbol{v})$

$$\boldsymbol{D_q} \boldsymbol{\mathscr{L}}|_{\left(\gamma_{\mathcal{Q}}(t), \frac{\mathrm{d}}{\mathrm{d}t} \gamma_{\mathcal{Q}}(t)\right)} - \tfrac{\mathrm{d}}{\mathrm{d}t} \left( \boldsymbol{D_v} \boldsymbol{\mathscr{L}}|_{\left(\gamma_{\mathcal{Q}}(\cdot), \frac{\mathrm{d}}{\mathrm{d}t} \gamma_{\mathcal{Q}}(\cdot)\right)} \right) \Bigg|_t = \boldsymbol{0} \in \mathbb{R}^Q,$$

where $\boldsymbol{D_q}(\cdot)$ denotes the derivative with respect to the first $Q$ coordinates (named $\boldsymbol{q}$ here) and $\boldsymbol{D_v}(\cdot)$ the derivative for the last $Q$ coordinates (named $\boldsymbol{v}$ here).

Since the Euler–Lagrange equations are obtained from a variation of an action functional, it is well-known that the solution curve is guaranteed to conserve a scalar-valued function (see, e.g., [1, Sec. 3.5] and [49, Prop. 7.3.1]):

**Theorem 4.1.** *The* energy

$$\mathcal{E}\colon T\mathcal{Q} \to \mathbb{R}, \qquad Y_{\mathcal{Q}} = \left( q, v^{\underline{i}} \frac{\partial}{\partial y^{\underline{i}}}\Big|_q \right) \mapsto v^{\underline{j}} \frac{\partial \mathscr{L}}{\partial \xi^{\underline{Q+j}}}\bigg|_{Y_{\mathcal{Q}}} - \mathscr{L}\left(Y_{\mathcal{Q}}\right)$$

*is conserved along the lift of the solution curve* $\gamma_{\mathcal{Q}}$ *of the Euler–Lagrange equations, i.e.,*

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}\left(\Gamma_{\gamma_{\mathcal{Q}}}(\cdot)\right)\Big|_t = 0 \qquad \textit{for all } t \in \mathcal{I}.$$

The Lagrangian is called *regular* if the smooth $(0,2)$-tensor field defined by the second-order derivative of the Lagrangian w.r.t. the velocity

$$\tau_v|_{Y_{\mathcal{Q}}} := \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{Q+j}} \partial \xi^{\underline{Q+i}}}\bigg|_{Y_{\mathcal{Q}}} \mathrm{d}\xi^{\underline{Q+i}}\Big|_{Y_{\mathcal{Q}}} \otimes \mathrm{d}\xi^{\underline{Q+j}}\Big|_{Y_{\mathcal{Q}}} \tag{4.10}$$

at each point $Y_{\mathcal{Q}} \in T\mathcal{Q}$ is nondegenerate ($\triangleright$ Section 4.1.4). In this case, we can formulate the *Euler–Lagrangian vector field* $X_{\mathscr{L}} \in \mathfrak{X}_{T\mathcal{Q}}$ such that at a point $Y_{\mathcal{Q}} = \left( q, v^{\underline{i}} \frac{\partial}{\partial x^{\underline{i}}}\Big|_q \right) \in T\mathcal{Q}$, it holds

$$X_{\mathscr{L}}|_{Y_{\mathcal{Q}}} := v^{\underline{i}} \frac{\partial}{\partial \xi^{\underline{i}}}\Big|_{Y_{\mathcal{Q}}} + \left(\tau_v|_{Y_{\mathcal{Q}}}\right)^{\underline{(Q+i)(Q+j)}} \left( \frac{\partial \mathscr{L}}{\partial \xi^{\underline{j}}}\bigg|_{Y_{\mathcal{Q}}} - \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{k}} \partial \xi^{\underline{Q+j}}}\bigg|_{Y_{\mathcal{Q}}} v^{\underline{k}} \right) \frac{\partial}{\partial \xi^{\underline{Q+i}}}\bigg|_{Y_{\mathcal{Q}}}, \tag{4.11}$$

where we use the convention from Section 4.1.4 to use upper indices to denote the corresponding inverse tensor field. This vector field can be used to formulate the Lagrangian system: Let $\gamma \in \mathcal{C}^\infty(\mathcal{I}, T\mathcal{Q})$ be an integral curve of $X_{\mathscr{L}}$ with starting point $(q_0, v_0) \in T\mathcal{Q}$. Then, solving the Euler–Lagrange equations (4.9) for $\gamma_{\mathcal{Q}}$ is equivalent to finding the integral curve $\gamma$ of $X_{\mathscr{L}}$ with $\gamma(t) = \Gamma_{\gamma_{\mathcal{Q}}}(t)$. In bold notation, the system for $\gamma$ reads

$$\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\gamma}\Big|_t = \begin{pmatrix} \boldsymbol{\gamma}_v(t) \\ \boldsymbol{\tau}_v|_{\boldsymbol{\gamma}(t)}^{-1} \left( \boldsymbol{D}_{\boldsymbol{q}} \boldsymbol{\mathscr{L}}|_{\boldsymbol{\gamma}(t)} - \boldsymbol{D}^2_{\boldsymbol{vq}} \boldsymbol{\mathscr{L}}\Big|_{\boldsymbol{\gamma}(t)} \boldsymbol{\gamma}_v(t) \right) \end{pmatrix} \in \xi(TV) \subseteq \mathbb{R}^{2Q}. \tag{4.12}$$

Here we denote by $\boldsymbol{D}^2_{\boldsymbol{vq}} \boldsymbol{\mathscr{L}}\Big|_{\boldsymbol{Y}_{\mathcal{Q}}} \in \mathbb{R}^{Q \times Q}$ the mixed derivative w.r.t. $\boldsymbol{v}$ and $\boldsymbol{q}$ and the solution curve is split $\boldsymbol{\gamma}(t) = (\boldsymbol{\gamma}_q(t), \boldsymbol{\gamma}_v(t)) \in \xi(TV) \subseteq \mathbb{R}^{2Q}$ in a part for $\boldsymbol{q}$ and a part for $\boldsymbol{v}$. The system (4.12) is typically referred to as the *first-order formulation* for the Lagrangian system.

4.3. **Hamiltonian systems.** In this subsection, we derive a formulation of Hamiltonian systems on a manifold,[7] providing the structure to perform MOR in the forthcoming Section 5.3. Let us recall from Section 4.1.1 that the differential of a smooth function $G \in \mathcal{C}^\infty(\mathcal{M}, \mathbb{R})$ at a point $m \in \mathcal{M}$ defines a covector $\mathrm{d}G|_m \in T^*_m\mathcal{M}$. Extending this idea, the differential $\mathrm{d}G \in \mathcal{C}^\infty(T\mathcal{M}, T\mathbb{R})$ defines a smooth covector field $\mathrm{d}G \in \Gamma\left(T^{(0,1)}(T\mathcal{M})\right)$ with component functions $(\mathrm{d}G|_m)_{\underline{i}} = \frac{\partial G}{\partial x^{\underline{i}}}\Big|_m$.

For a given symplectic manifold $(\mathcal{M}, \omega)$ and a smooth function $\mathscr{H} \in \mathcal{C}^\infty(\mathcal{M}, \mathbb{R})$ referred to as the *Hamiltonian (function)*, the *Hamiltonian vector field*

$$X_{\mathscr{H}} := \sharp_\omega(\mathrm{d}\mathscr{H}) \in \Gamma\left(T^{(1,0)}(T\mathcal{M})\right), \quad \text{or in index notation: } (X_{\mathscr{H}}|_m)^{\underline{i}} = (\omega|_m)^{\underline{ij}} \left(\mathrm{d}\mathscr{H}|_m\right)_{\underline{j}}$$

[7] Hamiltonian systems may result from Lagrangian systems via a Legendre transformation, but this is not the subject of the current work, so we refer to [1, Sec. 3.6].

is uniquely defined due to the nondegeneracy of $\omega$. A *Hamiltonian system* $(\mathcal{M}, \omega, \mathscr{H})$ is an initial value problem (2.9) with an integral curve $\gamma \in \mathcal{C}^\infty(\mathcal{I}, \mathcal{M})$ of $X_{\mathscr{H}}$ with starting point $\gamma_0 \in \mathcal{M}$, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}t}\gamma\Big|_t = X_{\mathscr{H}}|_{\gamma(t)} \in T_{\gamma(t)}\mathcal{M} \qquad \text{and} \qquad \gamma(t_0) = \gamma_0 \in \mathcal{M}. \tag{4.13}$$

We denote a Hamiltonian system in bold notation[8] with

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\gamma}\Big|_t = \left(\boldsymbol{\omega}|_{\gamma(t)}\right)^{-1} \boldsymbol{D\mathscr{H}}|_{\gamma(t)}^\top \in \mathbb{R}^N, \qquad \boldsymbol{\gamma}(t_0) = \boldsymbol{\gamma}_0 \in \mathbb{R}^N. \tag{4.14}$$

This special construction of the vector field guarantees that the Hamiltonian is conserved along the solution curve, since

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathscr{H} \circ \gamma)\Big|_t \overset{(2.5)}{=} \left(\mathrm{d}\mathscr{H}|_{\gamma(t)}\right)_{\underline{i}} \left(\frac{\mathrm{d}}{\mathrm{d}t}\gamma\Big|_t\right)^{\underline{i}} \overset{(4.13)}{=} \left(\mathrm{d}\mathscr{H}|_{\gamma(t)}\right)_{\underline{i}} \left(\omega|_{\gamma(t)}\right)^{\underline{ij}} \left(\mathrm{d}\mathscr{H}|_{\gamma(t)}\right)_{\underline{j}} = 0,$$

where the last step uses that for skew-symmetric tensors $\sigma \in T^{(2,0)}(T_m\mathcal{M})$, it holds $\lambda_{\underline{i}}\, \sigma^{\underline{ij}}\, \lambda_{\underline{j}} = -\lambda_{\underline{i}}\, \sigma^{\underline{ij}}\, \lambda_{\underline{j}} = 0$ for all covectors $\lambda \in T_m^*\mathcal{M}$.

For two given symplectic manifolds $(\mathcal{M}, \omega)$ and $(\mathcal{Q}, \eta)$, we call a smooth diffeomorphism $F \in \mathcal{C}^\infty(\mathcal{Q}, \mathcal{M})$ a *symplectomorphism* if $F^*\omega = \eta$. It can be shown that the flow of a Hamiltonian system $\theta_t : \mathcal{M} \to \mathcal{M}$ is a symplectomorphism.

The theorem of Darboux (see e.g. [1, Thm. 3.2.2]) guarantees that for each point $m \in \mathcal{M}$, there exists a chart $(U, x)$ with $m \in U$ which is *canonical*, i.e., the symplectic form in these coordinates can be represented with $\boldsymbol{\omega} \equiv \mathbb{J}_{2N}^\top$ by the *canonical Poisson tensor*[9]

$$\mathbb{J}_{2N} = \begin{pmatrix} \mathbf{0}_N & \boldsymbol{I}_N \\ -\boldsymbol{I}_N & \mathbf{0}_N \end{pmatrix} \in \mathbb{R}^{2N \times 2N} \qquad \text{for which} \qquad \mathbb{J}_{2N}^\top = -\mathbb{J}_{2N} = \mathbb{J}_{2N}^{-1}, \tag{4.15}$$

where $\boldsymbol{I}_N, \mathbf{0}_N \in \mathbb{R}^{N \times N}$ are the identity matrix and matrix of all zeros, respectively. In the case of $\mathcal{M} = \mathbb{R}^{2N}$ with $\boldsymbol{\omega}|_{\boldsymbol{m}} = \mathbb{J}_{2N}^\top$ for all $m \in \mathcal{M}$, we call $(\mathbb{R}^{2N}, \mathbb{J}_{2N}^\top, \boldsymbol{\mathscr{H}})$ a *canonical Hamiltonian system.*

## 5. Structure-preserving MOR on manifolds

With the general model reduction framework presented in Section 3 at hand, we now discuss how the general framework can be specialized to preserve important features of the initial value problem on the manifold. In more detail, we first introduce the *generalized manifold Galerkin* (GMG) reduction map in Section 5.1 and then use it to discuss the structure-preserving MOR of

- Lagrangian systems in Section 5.2, and
- Hamiltonian systems in Section 5.3.

[8] As the Jacobian $\boldsymbol{D\mathscr{H}}|_{\boldsymbol{m}} \in \mathbb{R}^{1 \times N}$ is a row vector, we need to transpose it for the multiplication to match dimensions.

[9] Note that in contrast to existing work in the field of structure-preserving MOR of Hamiltonian systems, we speak of the symplectic form $\boldsymbol{\omega} = \mathbb{J}_{2N}^\top$ instead of $\mathbb{J}_{2N}$. This yields the same Hamiltonian vector field $\boldsymbol{X}_{\mathscr{H}}|_{\boldsymbol{m}} = \mathbb{J}_{2N}\boldsymbol{D\mathscr{H}}|_{\boldsymbol{m}}^\top$ and does not change the reduction formulas later, but it helps to understand the more general case of noncanonical coordinates $\boldsymbol{\omega} \neq \mathbb{J}_{2N}^\top$.

5.1. **Generalized manifold Galerkin.** Assume that the manifold $\mathcal{M}$ of dimension $N$ is endowed with a nondegenerate $(0,2)$-tensor field $\tau \in \Gamma\,(T^{(0,2)}(T\mathcal{M}))$, as defined in Section 4.1.4. As in Section 3.3, we assume that we have already constructed an embedded submanifold $\varphi(\check{\mathcal{M}}) \subseteq \mathcal{M}$ defined by a smooth embedding $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$, i.e., we have completed the Approximation step from the general MOR workflow in Section 3.1.4. The straightforward way to define a reduced tensor field is to use the pullback from Section 4.1.5. Hence, we make the following assumption.

**Assumption 5.1.** *Given the nondegenerate $(0,2)$-tensor field $\tau \in \Gamma\,(T^{(0,2)}(T\mathcal{M}))$, the smooth embedding $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$ is such that the reduced tensor field*

$$\check{\tau} := \varphi^* \tau \in \Gamma\,(T^{(0,2)}(T\check{\mathcal{M}}))\,,$$

*is nondegenerate.*

Note that the reduced tensor field in bold notation reads

$$\check{\boldsymbol{\tau}}|_{\check{\boldsymbol{m}}} = \boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}^\top \boldsymbol{\tau}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})} \boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}} \in \mathbb{R}^{n \times n}, \tag{5.1}$$

which immediately illustrates that Assumption 5.1 may not be satisfied, in general. For instance, if we take $\mathcal{M} = \mathbb{R}^2$ and $\check{\mathcal{M}} = \mathbb{R}$, the tensor field to be a constant skew-symmetric matrix and a linear embedding, then Assumption 5.1 is violated. See also the forthcoming Example 5.14. On the other hand, if the tensor field is a Riemannian metric on $\mathcal{M}$, i.e., symmetric and positive definite, then the reduced tensor field is also a Riemannian metric.

We immediately obtain the following relation between the full and reduced musical isomorphisms discussed in Section 4.1.4.

**Lemma 5.2.** *Under Assumption 5.1, it holds*

$$\mathrm{d}\varphi^*|_{\check{m}} \circ \flat_\tau \circ \mathrm{d}\varphi|_{\check{m}} = \flat_{\check{\tau}} \in \mathcal{C}^\infty(T_{\check{m}}\check{\mathcal{M}}, T^*_{\check{m}}\check{\mathcal{M}}). \tag{5.2}$$

*Proof.* We prove the statement in index notation. Using (4.2), (4.6), (2.4), and (4.5), we obtain for all $\check{m} \in \check{\mathcal{M}}$, all $\check{v} \in T_{\check{m}}\check{\mathcal{M}}$ and all $1 \leq i \leq n$

$$\begin{aligned}
\left(\flat_{\check{\tau}}\left(\check{v}\right)\right)_{\underline{i}} &= \left(\check{\tau}|_{\check{m}}\right)_{\underline{ij}}\, \check{v}^{\underline{j}} = \left(\tau|_{\varphi(\check{m})}\right)_{\underline{\ell_1\ell_2}} \left.\frac{\partial \varphi^{\underline{\ell_1}}}{\partial x^{\underline{i}}}\right|_{\check{m}} \left.\frac{\partial \varphi^{\underline{\ell_2}}}{\partial x^{\underline{j}}}\right|_{\check{m}} \check{v}^{\underline{j}} \\
&= \left.\frac{\partial \varphi^{\underline{\ell_1}}}{\partial x^{\underline{i}}}\right|_{\check{m}} \left(\tau|_{\varphi(\check{m})}\right)_{\underline{\ell_1\ell_2}} \left.\frac{\partial \varphi^{\underline{\ell_2}}}{\partial x^{\underline{j}}}\right|_{\check{m}} \check{v}^{\underline{j}} = \left(\left(\mathrm{d}\varphi^*|_{\check{m}} \circ \flat_\tau \circ \mathrm{d}\varphi|_{\check{m}}\right)(\check{v})\right)_{\underline{i}}. \qquad \square
\end{aligned}$$

The additional structure allows us to construct an alternative reduction mapping to the MPG reduction map (3.11), which we refer to as the *generalized manifold Galerkin* (GMG)

$$R_{\mathsf{GMG}}\colon T\mathcal{M} \supseteq E_{\varphi(\check{\mathcal{M}})} \to T\check{\mathcal{M}}, \qquad (m, v) \mapsto \left(\varrho(m), \left(\sharp_{\check{\tau}} \circ \mathrm{d}\varphi^*|_{\varrho(m)} \circ \flat_\tau\right)(v)\right), \tag{5.3}$$

which is defined on the vector bundle

$$E_{\varphi(\check{\mathcal{M}})} := \dot{\bigcup}_{m \in \varphi(\check{\mathcal{M}})} T_m\mathcal{M}.$$

The domain $E_{\varphi(\check{\mathcal{M}})} \subseteq T\mathcal{M}$ of the GMG reduction map is in general smaller than in the original definition of a reduction map (▷ Definition 3.3). Nevertheless, all previous results are valid for reduction maps $R : E_{\varphi(\check{\mathcal{M}})} \to T\check{\mathcal{M}}$ as the reduction map is only used in the ROM to project $X|_{\varphi(\check{m})} \in T_{\varphi(\check{m})}\mathcal{M}$ which is part of $E_{\varphi(\check{\mathcal{M}})}$. We avoided introducing $E_{\varphi(\check{\mathcal{M}})}$ earlier for a better readability. The restriction of the domain for the GMG is necessary as $\mathrm{d}\varphi^*|_{\check{m}} : T^*_{\varphi(\check{m})}\mathcal{M} \to T^*_{\check{m}}\check{\mathcal{M}}$ is defined on $T^*_{\varphi(\check{m})}\mathcal{M}$ only.

By construction, (3.4a), Lemma 5.2, and (4.4), we obtain

$$R_{\mathsf{GMG}}|_{\varphi(\check{m})} \circ \mathrm{d}\varphi|_{\check{m}} = \sharp_{\check{\tau}} \circ \mathrm{d}\varphi^*|_{(\varrho\circ\varphi)(\check{m})} \circ \flat_{\tau} \circ \mathrm{d}\varphi|_{\check{m}} = \sharp_{\check{\tau}} \circ \flat_{\check{\tau}} = \mathrm{id}_{T_{\check{m}}\check{\mathcal{M}}},$$

which proves the following result.

**Theorem 5.3.** *The* GMG *reduction* (5.3) *is a reduction map for* $\varphi$*.*

The corresponding ROM (3.5) obtained with the GMG reduction map is called *GMG-ROM*. In bold notation, the associated reduced vector field for the FOM vector field $X \in \mathfrak{X}_{\mathcal{M}}$ reads with (3.4a)

$$\begin{aligned} \boldsymbol{R}_{\mathsf{GMG}}|_{\varphi(\check{m})}\left(\boldsymbol{X}|_{\varphi(\check{\boldsymbol{m}})}\right) &= \left(\boldsymbol{D\varphi}|_{(\boldsymbol{\varrho}\circ\boldsymbol{\varphi})(\check{\boldsymbol{m}})}^{\top}\boldsymbol{\tau}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})}\boldsymbol{D\varphi}|_{(\boldsymbol{\varrho}\circ\boldsymbol{\varphi})(\check{\boldsymbol{m}})}\right)^{-1}\boldsymbol{D\varphi}|_{(\boldsymbol{\varrho}\circ\boldsymbol{\varphi})(\check{\boldsymbol{m}})}^{\top}\boldsymbol{\tau}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})}\boldsymbol{X}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})} \\ &= \left(\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}^{\top}\boldsymbol{\tau}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})}\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}\right)^{-1}\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}^{\top}\boldsymbol{\tau}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})}\boldsymbol{X}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})} \in \mathbb{R}^n. \end{aligned} \tag{5.4}$$

To motivate the name GMG, we consider the special case that $\mathcal{M} = \mathbb{R}^N$, $\check{\mathcal{M}} = \mathbb{R}^n$ are vector spaces over $\mathbb{R}$ (with identity charts $x \equiv \mathrm{id}_{\mathbb{R}^N}$, $\check{x} \equiv \mathrm{id}_{\mathbb{R}^n}$) and the nondegenerate tensor field $\tau$ is a Riemannian metric that is constant in coordinates, i.e., $\boldsymbol{\tau}|_{\boldsymbol{m}} = \boldsymbol{\tau} = \mathrm{const}$. We then obtain with (5.4)

$$\begin{aligned} \boldsymbol{R}_{\mathsf{GMG}}|_{\varphi(\check{m})}\left(\boldsymbol{X}|_{\varphi(\check{\boldsymbol{m}})}\right) &= \left(\left(\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}^{\top}\boldsymbol{\tau}^{1/2}\right)\left(\boldsymbol{\tau}^{1/2}\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}\right)\right)^{-1}\left(\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}^{\top}\boldsymbol{\tau}^{1/2}\right)\boldsymbol{\tau}^{1/2}\boldsymbol{X}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})} \\ &= \left(\boldsymbol{\tau}^{1/2}\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}\right)^{\dagger}\boldsymbol{\tau}^{1/2}\boldsymbol{X}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})}, \end{aligned}$$

where $(\cdot)^{\dagger}$ denotes the Moore–Penrose pseudoinverse. In particular, we recover the *manifold Galerkin projection* introduced in [46, Rem 3.4], which allows interpreting the reduced vector field as the optimal projection w.r.t. the Riemannian metric $\tau$; see [46, Sec. 3.2].

**Example 5.4** (Special case: linear-subspace MOR)**.** *In the case of* $\varrho$, $\varphi$ *being linear as in Example 3.4 with* $\boldsymbol{\tau} \equiv \mathrm{const}$ *and* $\boldsymbol{V}^{\top}\boldsymbol{\tau}\boldsymbol{V} = \boldsymbol{I}_n$*, the* GMG *reduction is exactly the* ROM *obtained via standard Galerkin projection*

$$\boldsymbol{R}_{GMG}|_{\varphi(\check{\boldsymbol{m}})}(\boldsymbol{v}) = \left(\boldsymbol{V}^{\top}\boldsymbol{\tau}\boldsymbol{V}\right)^{-1}\boldsymbol{V}^{\top}\boldsymbol{\tau}\boldsymbol{v} = \boldsymbol{V}^{\top}\boldsymbol{\tau}\boldsymbol{v}, \qquad \left.\tfrac{\mathrm{d}}{\mathrm{d}t}\check{\boldsymbol{\gamma}}\right|_t = \boldsymbol{V}^{\top}\boldsymbol{\tau}\boldsymbol{X}|_{\check{\boldsymbol{\gamma}}(t)}.$$

As discussed in Section 4, the FOM vector field may possess additional structure in specific applications, such as Lagrangian or Hamiltonian dynamics. In the following, we show that the GMG reduction can be used to formulate structure-preserving MOR (on manifolds) for Lagrangian and Hamiltonian systems by choosing a specific nondegenerate tensor field.

5.2. **MOR on manifolds for Lagrangian systems.** As in Section 4.2, consider a $Q$-dimensional smooth manifold $\mathcal{Q}$ with a chart $(V, y)$ and the corresponding chart $(TV, \xi)$ of the tangent bundle $T\mathcal{Q}$ (▷ Section 4.2). The manifold to be reduced in the context of Lagrangian systems is the tangent bundle $T\mathcal{Q}$. To be consistent with the notation introduced before, we thus set $\mathcal{M} := T\mathcal{Q}$ with even dimension $N := \dim(\mathcal{M}) := 2Q$. Instead of working directly on $\mathcal{M}$, we still aim for a construction on $\mathcal{Q}$ by employing that the differential of a smooth map (▷ Section 2.5) is a mapping between the associated tangent spaces.

**Definition 5.5** (Lifted embedding and lifted point reduction)**.** *Consider an embedded submanifold $\varphi_{\mathcal{Q}}(\check{\mathcal{Q}}) \subseteq \mathcal{Q}$ defined by a $\check{Q}$-dimensional manifold $\check{\mathcal{Q}}$ and a smooth embedding $\varphi_{\mathcal{Q}}\colon \check{\mathcal{Q}} \to \varphi_{\mathcal{Q}}(\check{\mathcal{Q}})$. Then, we call*

$$\varphi := \mathrm{d}\varphi_{\mathcal{Q}}\colon T\check{\mathcal{Q}} \to T\Big(\varphi_{\mathcal{Q}}(\check{\mathcal{Q}})\Big), \qquad (\check{q}, \check{v}) \mapsto \Big(\varphi_{\mathcal{Q}}(\check{q}), \mathrm{d}\varphi_{\mathcal{Q}}\big|_{\check{q}}(\check{v})\Big)$$

*the* lifted embedding *for $\varphi_{\mathcal{Q}}$. Analogously, for a point reduction $\varrho_{\mathcal{Q}}\colon \mathcal{Q} \to \check{\mathcal{Q}}$, we define the* lifted point reduction

$$\varrho := \mathrm{d}\varrho_{\mathcal{Q}}\colon T\mathcal{Q} \to T\check{\mathcal{Q}}, \qquad (q, v) \mapsto \Big(\varrho_{\mathcal{Q}}(q), \mathrm{d}\varrho_{\mathcal{Q}}\big|_{q}(v)\Big).$$

Let us emphasize that $\varrho$ is indeed a point reduction on $\mathcal{M} = T\mathcal{Q}$ for the lifted embedding $\varphi$, which is a straightforward consequence of Theorem 3.9. For $\Big(\check{q}, \check{v}^{\underline{k}} \frac{\partial}{\partial \check{y}^{\underline{k}}}\Big|_{\check{q}}\Big) \in T\check{\mathcal{Q}}$ with a chart $(\check{V}, \check{y})$ for $\check{\mathcal{Q}}$ and $(T\check{V}, \check{\xi})$ for $T\check{\mathcal{Q}}$, we immediately obtain

$$\frac{\partial \varphi^{\underline{i}}}{\partial \check{\xi}^{\underline{j}}}\Bigg|_{\Big(\check{q}, \check{v}^{\underline{k}} \frac{\partial}{\partial \check{y}^{\underline{k}}}\big|_{\check{q}}\Big)} = \begin{cases} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{i}}}{\partial \check{y}^{\underline{j}}}\Big|_{\check{q}}, & 1 \leq i \leq Q,\ 1 \leq j \leq \check{Q}, \\ 0, & 1 \leq i \leq Q,\ 1 \leq j - \check{Q} \leq \check{Q}, \\ \frac{\partial^2 {\varphi_{\mathcal{Q}}}^{\underline{i-Q}}}{\partial \check{y}^{\underline{j}} \partial \check{y}^{\underline{k}}}\Big|_{\check{q}} \check{v}^{\underline{k}}, & 1 \leq i - Q \leq Q,\ 1 \leq j \leq \check{Q}, \\ \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{i-Q}}}{\partial \check{y}^{\underline{j-\check{Q}}}}\Big|_{\check{q}}, & 1 \leq i - Q \leq Q,\ 1 \leq j - \check{Q} \leq \check{Q}, \end{cases}$$

which reads in bold notation

$$\boldsymbol{\varphi}(\check{\boldsymbol{q}}, \check{\boldsymbol{v}}) = \begin{pmatrix} \boldsymbol{\varphi}_{\mathcal{Q}}(\check{\boldsymbol{q}}) \\ \boldsymbol{D}\boldsymbol{\varphi}_{\mathcal{Q}}\big|_{\check{\boldsymbol{q}}} \check{\boldsymbol{v}} \end{pmatrix} \in \mathbb{R}^{2Q}, \quad \boldsymbol{D}\boldsymbol{\varphi}\big|_{(\check{\boldsymbol{q}}, \check{\boldsymbol{v}})} = \begin{pmatrix} \boldsymbol{D}\boldsymbol{\varphi}_{\mathcal{Q}}\big|_{\check{\boldsymbol{q}}} & \boldsymbol{0} \\ \boldsymbol{D}\Big(\boldsymbol{D}\boldsymbol{\varphi}_{\mathcal{Q}}\big|_{(\cdot)} \check{\boldsymbol{v}}\Big)\Big|_{\check{\boldsymbol{q}}} & \boldsymbol{D}\boldsymbol{\varphi}_{\mathcal{Q}}\big|_{\check{\boldsymbol{q}}} \end{pmatrix} \in \mathbb{R}^{2Q \times 2\check{Q}}.$$

**Example 5.6.** *For a linear embedding $\boldsymbol{\varphi}_{\mathcal{Q}}(\check{\boldsymbol{q}}) = \boldsymbol{V}\check{\boldsymbol{q}}$ as in Example 3.4, the lifted embedding from Definition 5.5 is described by a block-diagonal basis matrix*

$$\boldsymbol{\varphi}(\check{\boldsymbol{q}}, \check{\boldsymbol{v}}) = \begin{pmatrix} \boldsymbol{V} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V} \end{pmatrix} \begin{pmatrix} \check{\boldsymbol{q}} \\ \check{\boldsymbol{v}} \end{pmatrix},$$

*which is frequently used in MOR for second-order systems (see, e.g., [68]).*

With these preparations, let us now assume that we have a Lagrangian system $(\mathcal{Q}, \mathscr{L})$ with initial value $(q_0, v_0) \in T\mathcal{Q}$ together with embedded submanifold $\varphi_{\mathcal{Q}}(\check{\mathcal{Q}}) \subseteq \mathcal{Q}$ with the embedding $\varphi_{\mathcal{Q}}$ and a point reduction $\varrho_{\mathcal{Q}}$ available. Let $\varphi$ and $\varrho$ denote the corresponding lifted embedding and lifted point reduction as in Definition 5.5. To preserve the Lagrangian system structure in the ROM, we do not aim for a projection of the Euler–Lagrange equations (4.9) but rather start by constructing a reduced Lagrangian via

$$\check{\mathscr{L}} := \varphi^* \mathscr{L} = \mathscr{L} \circ \varphi \in \mathcal{C}^\infty(T\check{\mathcal{Q}}) \tag{5.5}$$

and immediately obtain the reduced Lagrangian system $(\check{\mathcal{Q}}, \check{\mathscr{L}})$ with reduced initial value $(\check{q}_0, \check{v}_0) := \varrho(q_0, v_0) \in T\check{\mathcal{Q}} =: \check{\mathcal{M}}$. Note that with this strategy, we immediately obtain the ROM that itself is a Lagrangian system, which is not automatically guaranteed if we reduce the vector field (4.11). Straightforward calculations (see Appendix B.1) show that the Euler-Lagrange equations of the reduced Lagrangian system read

$$
\begin{aligned}
0 = \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{j}}}{\partial \check{y}^{\underline{i}}}\bigg|_{\check{\gamma}(t)} \Bigg( & \frac{\partial \mathscr{L}}{\partial \xi^{\underline{j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}}(t)\right)} - \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{k}} \partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}}(t)\right)} \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\mathrm{d}}{\mathrm{d}t} \check{\gamma}_{\mathcal{Q}}{}^{\underline{\ell}}\bigg|_t \\
& - \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{Q+k}} \partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}}(t)\right)} \frac{\partial^2 \varphi_{\mathcal{Q}}{}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}} \partial \check{y}^{\underline{p}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}|_t} \frac{\mathrm{d}}{\mathrm{d}t} \check{\gamma}_{\mathcal{Q}}{}^{\underline{p}}\bigg|_t \frac{\mathrm{d}}{\mathrm{d}t} \check{\gamma}_{\mathcal{Q}}{}^{\underline{\ell}}\bigg|_t \\
& - \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{Q+k}} \partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}}(t)\right)} \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\mathrm{d}^2}{\mathrm{d}t^2} \check{\gamma}_{\mathcal{Q}}{}^{\underline{\ell}}\bigg|_t \Bigg)
\end{aligned}
\tag{5.6}
$$

for $1 \leq i \leq n$ where the right-hand side sums over $1 \leq j,k \leq Q$ and $1 \leq \ell, p \leq \check{Q}$ by the Einstein summation convention (1.3). In bold notation, the reduced Euler–Lagrange equations read

$$
\begin{aligned}
\mathbf{0} = \boldsymbol{D}\boldsymbol{\varphi}_{\mathcal{Q}}|_{\check{\gamma}_{\mathcal{Q}}(t)}^{\top} \Bigg( & \boldsymbol{D}_{\boldsymbol{q}}\boldsymbol{\mathscr{L}}|_{\boldsymbol{\varphi}\left(\check{\gamma}_{\mathcal{Q}}(t), \frac{\mathrm{d}}{\mathrm{d}t}\check{\gamma}_{\mathcal{Q}}\big|_t\right)} - \boldsymbol{D}^2_{\boldsymbol{vq}}\boldsymbol{\mathscr{L}}\Big|_{\boldsymbol{\varphi}\left(\check{\gamma}_{\mathcal{Q}}(t), \frac{\mathrm{d}}{\mathrm{d}t}\check{\gamma}_{\mathcal{Q}}\big|_t\right)} \boldsymbol{D}\boldsymbol{\varphi}_{\mathcal{Q}}|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\mathrm{d}}{\mathrm{d}t}\check{\boldsymbol{\gamma}}_{\mathcal{Q}}\Big|_t \\
& - \boldsymbol{D}^2_{\boldsymbol{vv}}\boldsymbol{\mathscr{L}}\Big|_{\boldsymbol{\varphi}\left(\check{\gamma}_{\mathcal{Q}}(t), \frac{\mathrm{d}}{\mathrm{d}t}\check{\gamma}_{\mathcal{Q}}\big|_t\right)} \boldsymbol{D}^2\boldsymbol{\varphi}_{\mathcal{Q}}\Big|_{\check{\gamma}_{\mathcal{Q}}(t)} \left(\frac{\mathrm{d}}{\mathrm{d}t}\check{\boldsymbol{\gamma}}_{\mathcal{Q}}\Big|_t \otimes \frac{\mathrm{d}}{\mathrm{d}t}\check{\boldsymbol{\gamma}}_{\mathcal{Q}}\Big|_t\right) \\
& - \boldsymbol{D}^2_{\boldsymbol{vv}}\boldsymbol{\mathscr{L}}\Big|_{\boldsymbol{\varphi}\left(\check{\gamma}_{\mathcal{Q}}(t), \frac{\mathrm{d}}{\mathrm{d}t}\check{\gamma}_{\mathcal{Q}}\big|_t\right)} \boldsymbol{D}\boldsymbol{\varphi}_{\mathcal{Q}}|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\mathrm{d}^2}{\mathrm{d}t^2}\check{\boldsymbol{\gamma}}_{\mathcal{Q}}\Big|_t \Bigg) \in \mathbb{R}^{\check{Q}}.
\end{aligned}
\tag{5.7}
$$

By construction, the reduced Lagrangian system fulfills the Euler–Lagrange equations for the reduced Lagrangian $\check{\mathscr{L}}$. Thus, Theorem 4.1 guarantees that along the lift of the solution curve $\check{\gamma}_{\mathcal{Q}}$, the *reduced energy*

$$
\check{\mathcal{E}} \colon T\check{\mathcal{Q}} \to \mathbb{R}, \qquad \check{Y}_{\mathcal{Q}} = \left(\check{q}, \check{v}^{\underline{i}} \frac{\partial}{\partial \check{y}^{\underline{i}}}\bigg|_{\check{q}}\right) \mapsto \check{v}^{\underline{j}} \frac{\partial \check{\mathscr{L}}}{\partial \check{\xi}^{\underline{\check{Q}+j}}}\bigg|_{\check{Y}_{\mathcal{Q}}} - \check{\mathscr{L}}\left(\check{Y}_{\mathcal{Q}}\right)
$$

is preserved. Note however, that in general $\check{\mathcal{E}} \not\equiv \mathcal{E} \circ \varphi$ since for $\check{Y}_{\mathcal{Q}} = \left(\check{q}, \check{v}^{\underline{i}} \frac{\partial}{\partial \check{y}^{\underline{i}}}\Big|_{\check{q}}\right)$ straightforward calculations yield

$$
\check{\mathcal{E}}\left(\check{Y}_{\mathcal{Q}}\right) - (\mathcal{E} \circ \varphi)\left(\check{Y}_{\mathcal{Q}}\right) = \check{v}^{\underline{j}} \frac{\partial^2 \varphi_{\mathcal{Q}}{}^{\underline{k}}}{\partial \check{y}^{\underline{j}} \partial \check{y}^{\underline{\ell}}}\bigg|_{\check{q}} \check{v}^{\underline{\ell}} \frac{\partial \mathscr{L}}{\partial \xi^{\underline{Q+k}}}\bigg|_{\varphi\left(\check{Y}_{\mathcal{Q}}\right)}.
$$

Following the construction in Section 4.2 and assuming that the reduced Lagrangian $\check{\mathscr{L}}$ is regular, we can formulate a reduced vector field $\check{X}_{\check{\mathscr{L}}} \in \Gamma\left(T^{(1,0)}(T\check{\mathcal{M}})\right)$ for the reduced Euler–Lagrange equations (5.7). Indeed, we obtain for a point $\check{Y}_{\mathcal{Q}} = \left(\check{q}, \check{v}^{\underline{i}} \frac{\partial}{\partial \check{y}^{\underline{i}}}\Big|_{\check{q}}\right) \in T\check{\mathcal{Q}}$ as

$$
\check{X}_{\check{\mathscr{L}}}\Big|_{\check{Y}_{\mathcal{Q}}} =: \check{v}^{\underline{i}} \frac{\partial}{\partial \check{\xi}^{\underline{i}}}\bigg|_{\check{Y}_{\mathcal{Q}}} + \left(\sigma|_{\check{Y}_{\mathcal{Q}}}\right)^{\underline{i}}_{\underline{\ell}} \left(\left(X_{\mathscr{L}}|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)^{\underline{Q+\ell}} - \frac{\partial^2 \varphi_{\mathcal{Q}}{}^{\underline{\ell}}}{\partial \check{y}^{\underline{p}} \partial \check{y}^{\underline{r}}}\bigg|_{\check{q}} \check{v}^{\underline{p}} \check{v}^{\underline{r}}\right) \frac{\partial}{\partial \check{\xi}^{\underline{\check{Q}+i}}}\bigg|_{\check{Y}_{\mathcal{Q}}}
\tag{5.8}
$$

with indices $1 \leq k, \ell \leq Q$ and $1 \leq i, j, p, r \leq \check{Q}$ and

$$
\left(\sigma|_{\check{Y}_{\mathcal{Q}}}\right)^{\underline{i}}_{\underline{\ell}} := \left(\check{\tau}_v|_{\check{Y}_{\mathcal{Q}}}\right)^{\underline{ij}} \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{k}}}{\partial \check{y}^{\underline{j}}}\bigg|_{\check{q}} \left(\tau_v|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{k\ell}}.
$$

In order to relate this reduction to our framework, we show in the following that the reduced Euler–Lagrangian vector field (5.8) can be interpreted as a GMG reduction (5.3) of the Euler–Lagrangian vector field (4.11) if an appropriate tensor field is selected. We refer to this as the *Lagrangian manifold Galerkin* (LMG). With the nondegenerate tensor field $\tau_v$ from (4.10), we define a tensor field $\tau_{\mathsf{LMG}} \in \Gamma\left(T^{(0,2)}(T\mathcal{M})\right)$ on $\mathcal{M} = T\mathcal{Q}$ with

$$\tau_{\mathsf{LMG}}|_{Y_{\mathcal{Q}}} := \left(\tau_q|_{Y_{\mathcal{Q}}}\right)_{\underline{ij}} \mathrm{d}\xi^{\underline{Q+i}}\Big|_{Y_{\mathcal{Q}}} \otimes \mathrm{d}\xi^{\underline{j}}\Big|_{Y_{\mathcal{Q}}} + \left(\tau_v|_{Y_{\mathcal{Q}}}\right)_{\underline{ij}} \mathrm{d}\xi^{\underline{i}}\Big|_{Y_{\mathcal{Q}}} \otimes \mathrm{d}\xi^{\underline{Q+j}}\Big|_{Y_{\mathcal{Q}}}, \tag{5.9}$$

where $\left(\tau_q|_{Y_{\mathcal{Q}}}\right)_{\underline{ij}}$ are additional components. A typical choice could be $\left(\tau_q|_{Y_{\mathcal{Q}}}\right)_{\underline{ij}} = \left(\tau_v|_{Y_{\mathcal{Q}}}\right)_{\underline{ij}}$. In bold notation, the tensor field reads

$$\boldsymbol{\tau}_{\mathsf{LMG}}|_{\boldsymbol{Y}_{\mathcal{Q}}} = \begin{pmatrix} \mathbf{0} & \boldsymbol{\tau}_v|_{\boldsymbol{Y}_{\mathcal{Q}}} \\ \boldsymbol{\tau}_q|_{\boldsymbol{Y}_{\mathcal{Q}}} & \mathbf{0} \end{pmatrix}. \tag{5.10}$$

The associated reduced tensor field is denoted with $\check{\tau}_{\mathsf{LMG}}$ (as in Section 5.1). Assuming that $\check{\tau}_{\mathsf{LMG}}$ is nondegenerate, we define the LMG reduction map

$$R_{\mathsf{LMG}} \colon T\mathcal{M} \supseteq E_{\varphi(\check{\mathcal{M}})} \to T\check{\mathcal{M}}, \qquad (m, v) \mapsto \left(\varrho(m), \left(\sharp_{\check{\tau}_{\mathsf{LMG}}} \circ \mathrm{d}\varphi^*|_{\varrho(m)} \circ \flat_{\tau_{\mathsf{LMG}}}\right)(v)\right). \tag{5.11}$$

The LMG reduction map (5.11) is a particular case of a GMG reduction map, and thus, we immediately obtain from Theorem 5.3 that $R_{\mathsf{LMG}}$ is a reduction map for the lifted embedding $\varphi$.

**Theorem 5.7.** *Consider the ROM obtained by reducing the Euler–Lagrange vector field with $R_{\mathsf{LMG}}$. Then solving this ROM for $\check{\gamma}$ is equivalent to solving the reduced Euler–Lagrange equations (5.6) for $\check{\gamma}_{\mathcal{Q}}$ with $\check{\gamma}(t) = \Gamma_{\check{\gamma}_{\mathcal{Q}}}(t)$.*

*Proof.* (▷ Appendix B.2). □

We conclude this section with three remarks.

**Remark 5.8.** *In the special case of a classical MOR $\mathcal{Q} = \mathbb{R}^Q$, $\check{\mathcal{Q}} = \mathbb{R}^{\check{Q}}$ with a linear embedding $\boldsymbol{\varphi}_{\mathcal{Q}}(\check{\boldsymbol{q}}) = \boldsymbol{V}\check{\boldsymbol{q}}$ as in Example 5.6, a linear point reduction $\boldsymbol{\varrho}_{\mathcal{Q}}(\boldsymbol{q}) = \boldsymbol{V}^\top\boldsymbol{q}$, and a quadratic Lagrangian $\mathscr{L}$, the reduced Euler–Lagrange equations (5.7) recover the ROM from [44]. In our framework that relates to the choice $\mathcal{M} = \mathbb{R}^{2Q}$, $\check{\mathcal{M}} = \mathbb{R}^{2\check{Q}}$, $\varphi$ as in Example 5.6, and $\boldsymbol{R}_{\mathsf{LMG}}(\boldsymbol{v}_{\boldsymbol{q}}, \boldsymbol{v}_{\boldsymbol{v}}) = \left(\boldsymbol{v}_{\boldsymbol{q}}^\top\boldsymbol{V}^\top,\ \boldsymbol{v}_{\boldsymbol{v}}^\top\boldsymbol{V}^\top\right)^\top$.*

**Remark 5.9.** *In [44], the authors argue that the reduced Euler–Lagrange equations cannot be obtained from a projection with the embedding $\varphi_{\mathcal{Q}}$ of the first-order system (which is formulated with the Euler–Lagrange vector field (4.11) in the scope of this work). This is no contradiction to our work since we suggest a projection based on the lifted embedding $\varphi$ from Definition 5.5 to obtain the reduced Euler–Lagrange equations via a reduction of the Euler–Lagrange vector field.*

**Remark 5.10** (Second-order derivatives of $\varphi_{\mathcal{Q}}$)**.** *The reduced Euler–Lagrange equations require the computation of second-order derivatives of $\varphi_{\mathcal{Q}}$, which might be computationally intensive. Notably, the formulation of the ROM in structure-preserving MOR for Hamiltonian systems presented in the following subsection is independent of second-order derivatives of the embedding $\varphi_{\mathcal{Q}}$.*

5.3. **MOR on manifolds for Hamiltonian systems.** Lastly, we assume to be given a Hamiltonian system $(\mathcal{M}, \omega, \mathscr{H})$ as FOM and demonstrate how structure-preserving MOR on manifolds can be formulated. The procedure works analogously to the GMG from Section 5.1, while choosing the symplectic form $\omega$ as the nondegenerate tensor field $\tau = \omega$.

First, we assume that the Approximation step is completed and we are given a reduced manifold $\check{\mathcal{M}}$ and a smooth embedding $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$ fulfilling Assumption 5.1, i.e., $\varphi^*\omega$ is nondegenerate. We show at the end of this section (▷ Lemma 5.13) that this assumption is enough that $\check{\omega} := \varphi^*\omega$ is a symplectic form and $(\check{\mathcal{M}}, \check{\omega})$ is a symplectic manifold. In this case, the embedding $\varphi\colon (\check{\mathcal{M}}, \check{\omega}) \to (\varphi(\check{\mathcal{M}}), \omega|_{\varphi(\check{\mathcal{M}})})$ is a symplectomorphism. Second, we use the reduction map

$$R_{\mathsf{SMG}}\colon T\mathcal{M} \supseteq E_{\varphi(\check{\mathcal{M}})} \to T\check{\mathcal{M}}, \qquad (m, v) \mapsto \left(\varrho(m), \left(\sharp_{\check{\omega}} \circ \mathrm{d}\varphi^*|_{\varrho(m)} \circ \flat_\omega\right)(v)\right), \tag{5.12}$$

which we refer to as the *symplectic manifold Galerkin* (SMG) reduction map. The SMG reduction map is a special case of the GMG reduction map (5.3) with $\tau = \omega$ and $\check{\tau} = \check{\omega}$, and, thus, we obtain from Theorem 5.3 that $R_{\mathsf{SMG}}$ is a reduction map for $\varphi$. Hence, the SMG reduction fits our MOR framework from Section 3.1 and it defines a ROM by (3.5), which we refer to as the SMG-ROM. It remains to show that the SMG-ROM indeed is a Hamiltonian system, which was the motivation for preserving the underlying structure.

**Theorem 5.11.** *The SMG-ROM is a Hamiltonian system $(\check{\mathcal{M}}, \check{\omega}, \check{\mathscr{H}})$ with the* reduced Hamiltonian $\check{\mathscr{H}} := \varphi^*\mathscr{H} = \mathscr{H} \circ \varphi$.

*Proof.* The ROM vector field with the SMG reduction (5.12) reads with (a) equations (3.4a) and (4.4), and (b) equation (4.8)

$$\begin{aligned} R_{\mathsf{SMG}}|_{\varphi(\check{m})}(X_{\mathscr{H}}|_{\varphi(\check{m})}) = {} & \left(\sharp_{\check{\omega}} \circ \mathrm{d}\varphi^*|_{(\varrho\circ\varphi)(\check{m})} \circ \flat_\omega\right)\left(\sharp_\omega\left(\mathrm{d}\mathscr{H}|_{\varphi(\check{m})}\right)\right) \\ & \overset{\text{(a)}}{=} \sharp_{\check{\omega}}\left(\mathrm{d}\varphi^*|_{\check{m}}\left(\mathrm{d}\mathscr{H}|_{\varphi(\check{m})}\right)\right) \overset{\text{(b)}}{=} \sharp_{\check{\omega}}\left(\mathrm{d}\check{\mathscr{H}}\Big|_{\check{m}}\right), \end{aligned} \tag{5.13}$$

which is exactly the Hamiltonian vector field of the Hamiltonian system $(\check{\mathcal{M}}, \check{\omega}, \check{\mathscr{H}})$. ☐

Using (5.4), the reduced vector field in the SMG-ROM in bold notation reads

$$\boldsymbol{R}_{\mathsf{SMG}}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})}\left(\boldsymbol{X}_{\mathscr{H}}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})}\right) = \left(\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}^\top \boldsymbol{\omega}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})} \boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}\right)^{-1} \underbrace{\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}^\top \boldsymbol{\omega}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})} \boldsymbol{X}_{\mathscr{H}}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})}}_{=\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}^\top \boldsymbol{D\mathscr{H}}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})}^\top = \boldsymbol{D}\check{\boldsymbol{\mathscr{H}}}\big|_{\check{\boldsymbol{m}}}^\top} \in \mathbb{R}^n. \tag{5.14}$$

For a canonical Hamiltonian system, our generalization of the SMG-ROM is consistent with the definitions existing in the literature, which is shown by the following lemma.

**Lemma 5.12.** *For a canonical Hamiltonian system $(\mathbb{R}^{2N}, \mathbb{J}_{2N}^\top, \boldsymbol{\mathscr{H}})$ and reduced symplectic manifold $(\check{\mathcal{M}}, \check{\omega}) = (\mathbb{R}^{2n}, \mathbb{J}_{2n}^\top)$, it holds that*

*(i) the SMG reduction evaluated at the base point $\boldsymbol{\varphi}(\check{\boldsymbol{m}})$ equals the symplectic inverse*

$$\boldsymbol{R}_{\mathsf{SMG}}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})}(\boldsymbol{v}) = \boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}^+ \boldsymbol{v} := \mathbb{J}_{2n}\, \boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}^\top \mathbb{J}_{2N}^\top \boldsymbol{v} \qquad \text{for all } \boldsymbol{v} \in \mathbb{R}^{2N},$$

*(ii) the SMG-ROM is consistent with [16], and*

*(iii) if, moreover, the embedding $\boldsymbol{\varphi}$ is linear, the SMG-ROM equals the symplectic Galerkin ROM introduced in [47, 60].*

*Proof.* By assumption, it holds $\mathcal{M} = \mathbb{R}^{2N}$, $\boldsymbol{\omega} = \mathbb{J}_{2N}^\top$, $\check{\mathcal{M}} = \mathbb{R}^{2n}$, $\check{\boldsymbol{\omega}} = \mathbb{J}_{2n}^\top$. (i) Inserting the quantities in (5.14) yields the statement. (ii) For $\boldsymbol{\varphi}$ to be a symplectomorphism, i.e., $(\varphi^*\omega)|_{\check{m}} = \check{\omega}|_{\check{m}}$ for all $\check{m} \in \check{\mathcal{M}}$, is with (5.1) equivalent to

$$\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}^\top \mathbb{J}_{2N}^\top \boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}} = \mathbb{J}_{2n}^\top \qquad \text{for all } \check{\boldsymbol{m}} \in \mathbb{R}^{2n}. \tag{5.15}$$

Considering $\mathbb{J}_{2N}^\top = -\mathbb{J}_{2N}$ and $\mathbb{J}_{2n}^\top = -\mathbb{J}_{2n}$ and multiplying the previous equation on both sides with $(-1)$, gives exactly the definition of a symplectic embedding from [16, Def. 2]. Thus, the assumptions on the embedding are equivalent (up to smoothness requirements). Moreover, the SMG-ROM in [16] is projected with the symplectic inverse which (by point (i)) is equivalent to the SMG reduction map for the case assumed in the present lemma ($\mathcal{M} = \mathbb{R}^{2N}$, $\boldsymbol{\omega} = \mathbb{J}_{2N}^\top$, $\check{\mathcal{M}} = \mathbb{R}^{2n}$, $\check{\boldsymbol{\omega}} = \mathbb{J}_{2n}^\top$).

(iii) If the embedding is linear, then there exists $\boldsymbol{V} \in \mathbb{R}^{2N\times 2n}$ such that $\boldsymbol{\varphi}(\check{\boldsymbol{m}}) = \boldsymbol{V}\check{\boldsymbol{m}}$. Then, the requirement of $\varphi$ to be a symplectomorphism is equivalent to $\boldsymbol{V}^\top \mathbb{J}_{2N} \boldsymbol{V} = \mathbb{J}_{2n}$, which is in [60, Equation 3.2] formulated as the condition that $\boldsymbol{V}$ is a *symplectic matrix*. Moreover, the symplectic inverse of $\boldsymbol{V}$ is used to obtain the ROM, which is, again, by point (i), equivalent to our approach in this particular case. $\square$

However, our approach extends the existing methods, as it also works (i) on general smooth manifolds (not just $\mathcal{M} = \mathbb{R}^{2N}$) and (ii) even in the case $\mathcal{M} = \mathbb{R}^{2N}$ for noncanonical symplectic forms $\boldsymbol{\omega} \neq \mathbb{J}_{2N}^\top$. Structure-preserving MOR for noncanonical Hamiltonian systems (for the particular case of a linear embedding) is discussed in [48]. Compared to that approach, however, we use the noncanonical symplectic form prescribed by the FOM, which generalizes the symplectic inverse straightforwardly.

It remains to show that assuming nondegeneracy of $\varphi^*\omega$ is enough that $\varphi^*\omega$ is a symplectic form, which we show in the following.

**Lemma 5.13.** *Consider a symplectic manifold $(\mathcal{M}, \omega)$, a smooth manifold $\check{\mathcal{M}}$, and a smooth embedding $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$ such that $\check{\omega} := \varphi^*\omega$ is nondegenerate. Then $\check{\omega}$ is a symplectic form, $(\check{\mathcal{M}}, \check{\omega})$ is a symplectic manifold, and $\varphi$ is a symplectomorphism.*

*Proof.* It is sufficient to show that $\check{\omega} = \varphi^*\omega$ is a symplectic form, which in this case results in showing that $\check{\omega}$ is skew-symmetric and closed. The skew-symmetry is inherited for all points $\check{m} \in \check{\mathcal{M}}$ since with (4.6)

$$(\check{\omega}|_{\check{m}})_{\underline{j_1 j_2}} = (\omega|_{\varphi(\check{m})})_{\underline{\ell_1 \ell_2}} \left.\frac{\partial \varphi^{\underline{\ell_1}}}{\partial \check{x}^{\underline{j_1}}}\right|_{\check{m}} \left.\frac{\partial \varphi^{\underline{\ell_2}}}{\partial \check{x}^{\underline{j_2}}}\right|_{\check{m}} = -(\omega|_{\varphi(\check{m})})_{\underline{\ell_2 \ell_1}} \left.\frac{\partial \varphi^{\underline{\ell_2}}}{\partial \check{x}^{\underline{j_2}}}\right|_{\check{m}} \left.\frac{\partial \varphi^{\underline{\ell_1}}}{\partial \check{x}^{\underline{j_1}}}\right|_{\check{m}} = -(\check{\omega}|_{\check{m}})_{\underline{j_2 j_1}}.$$

Closedness is inherited since the pullback of a closed form is closed again [45, proof of Prop. 17.2]. $\square$

Note that this is a central difference to reduced Riemannian metrics, which are automatically nondegenerate due to positive definiteness. The following example shows that the reduced tensor field can degenerate if arbitrary embeddings $\varphi$ in combination with a symplectic form are considered.

**Example 5.14** (Example for degenerate $\varphi^*\omega$)**.** *For an arbitrary $n$ with $2n \leq N$, consider $\mathcal{M} = \mathbb{R}^{2N}$, $\boldsymbol{\omega} = \mathbb{J}_{2N}^\top$, $\check{\mathcal{M}} = \mathbb{R}^{2n}$, and the embedding*

$$\boldsymbol{\varphi}(\check{\boldsymbol{m}}) = \boldsymbol{E}\check{\boldsymbol{m}} \qquad \text{with} \qquad \boldsymbol{E} := \begin{pmatrix} \boldsymbol{I}_{2n} \\ \boldsymbol{0}_{2N-2n} \end{pmatrix} \in \mathbb{R}^{2N\times 2n}.$$

*In this case, it holds $\boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}} = \boldsymbol{E}$ and the reduced tensor field is*

$$\check{\boldsymbol{\omega}}|_{\check{\boldsymbol{m}}} = \boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}}^\top \boldsymbol{\omega}|_{\boldsymbol{\varphi}(\check{\boldsymbol{m}})} \boldsymbol{D\varphi}|_{\check{\boldsymbol{m}}} = \begin{pmatrix} \boldsymbol{I}_{2n} & \boldsymbol{0}_{2N-2n} \end{pmatrix} \begin{pmatrix} \boldsymbol{0}_N & \boldsymbol{I}_N \\ -\boldsymbol{I}_N & \boldsymbol{0}_N \end{pmatrix} \begin{pmatrix} \boldsymbol{I}_{2n} \\ \boldsymbol{0}_{2N-2n} \end{pmatrix} = \boldsymbol{0} \in \mathbb{R}^{2n\times 2n},$$

*which is clearly degenerate.*

**Table 3** – MOR techniques from different references that are covered by MPG (3.11), GMG (5.3), LMG (5.11), and SMG (5.12) introduced in our work.

| name | ref. | | details |
|---:|---|---|---|
| QPROM | [4, 23, 37] | MPG | $\boldsymbol{\varrho}$ linear, $\boldsymbol{\varphi}$ quadratic |
| EncROM | [57] | MPG | $\boldsymbol{\varrho}, \boldsymbol{\varphi}$ autoencoders |
| qmf | [7] | GMG | $\boldsymbol{\tau} \equiv \boldsymbol{I}_N$, $\boldsymbol{\varrho}$ linear, $\boldsymbol{\varphi}$ quadratic |
| manifold Galerkin | [46] | GMG | $\boldsymbol{\tau}$ independent of $\boldsymbol{m}$, symmetric, pos. def. |
| | [18, 44] | LMG | $\boldsymbol{\tau}$ as in (5.10), $\boldsymbol{\varrho}, \boldsymbol{\varphi}$ linear |
| symplectic Galerkin | [47, 60] | SMG | $\boldsymbol{\tau} \equiv \mathbb{J}_{2N}^{\top}$, $\boldsymbol{\varrho}$, $\boldsymbol{\varphi}$ linear |
| SMG | [16] | SMG | $\boldsymbol{\tau} \equiv \mathbb{J}_{2N}^{\top}$, $\boldsymbol{\varrho}, \boldsymbol{\varphi}$ autoencoders |
| SMG-QMCL | [72] | SMG | $\boldsymbol{\tau} \equiv \mathbb{J}_{2N}^{\top}$, $\boldsymbol{\varrho}$ linear, $\boldsymbol{\varphi}$ from manifold cotangent lift |

## 6. Snapshot-based generation of embedding and point reduction

Another key task in MOR is the choice of a particular embedding $\varphi$ (the Approximation step in Section 3.1.4). In this section, we thus consider the construction of the embedding in a data-driven setting, which is directly combined with the construction of a point reduction $\varrho$. We first introduce the data-driven setting (▷ Section 6.1) and then detail four techniques to generate an embedding and a corresponding point reduction. In Table 3, we present an overview of selected methods discussed in the literature and how they fit into our general framework. Throughout the section, we assume to be given the $N$-dimensional smooth manifold $\mathcal{M}$ and a metric $d_{\mathcal{M}} \colon \mathcal{M} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$.

6.1. **Snapshot-based generation.** In the scope of the present work, we focus on *snapshot-based generation of an embedding and a point reduction*. Consider a finite subset $S_{\mathrm{train}} \subseteq S$ of the set of all solutions $S \subseteq \mathcal{M}$ from (3.2), which is referred to as the *(training-)set of snapshots* and its elements $m_{\mathrm{train}} \in S_{\mathrm{train}}$ as *snapshots*. Typically, the embedding and the point reduction are determined by searching in a given family of functions

$$\mathcal{F}_{\varphi,\varrho} := \left\{ (\varphi, \varrho) \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M}) \times \mathcal{C}^\infty(\mathcal{M}, \check{\mathcal{M}}) \,\middle|\, \varrho \text{ is a point reduction for } \varphi \text{ (3.4a)} \right\}$$

by optimizing over a functional $L \colon \mathcal{F}_{\varphi,\varrho} \to \mathbb{R}_{\geq 0}$ that measures the quality of approximation based on the snapshots $m_{\mathrm{train}} \in S_{\mathrm{train}}$, i.e.,

$$(\varphi^\star, \varrho^\star) := \operatorname*{arg\,min}_{(\varphi,\varrho) \in \mathcal{F}_{\varphi,\varrho}} L(\varphi, \varrho). \tag{6.1}$$

We emphasize that Lemma 2.1 guarantees that searching within $\mathcal{F}_{\varphi,\varrho}$ automatically yields that $\varphi$ is a smooth embedding and $\varphi(\check{\mathcal{M}})$ is an embedded submanifold. Note that for practical purposes, which we do not further consider, one might want to relax the smoothness assumptions in $\mathcal{F}_{\varphi,\varrho}$.

One well-established functional is the *mean squared error* (MSE)

$$L_{\mathsf{MSE}}(\varphi, \varrho) := \frac{1}{|S_{\mathrm{train}}|} \sum_{m_{\mathrm{train}} \in S_{\mathrm{train}}} \left( d_{\mathcal{M}}\left(m_{\mathrm{train}}, (\varphi \circ \varrho)(m_{\mathrm{train}})\right) \right)^2 \in \mathbb{R}_{\geq 0}. \tag{6.2}$$

The motivation of minimizing the MSE is that if $L_{\mathsf{MSE}}(\varphi, \varrho) = 0$, it is guaranteed that all snapshots $m_{\mathrm{train}} \in S_{\mathrm{train}}$ are in the image of the embedding $\varphi$ and thus directly lay on the embedded submanifold, i.e., $S_{\mathrm{train}} \subseteq \varphi(\check{\mathcal{M}})$. In general, however, the MSE is not equal to zero. Nevertheless, then we know that for each addend of (6.2) it holds that

$$\left(d_{\mathcal{M}}\left(m_{\mathrm{train}},\, (\varphi \circ \varrho)\left(m_{\mathrm{train}}\right)\right)\right)^2 \leq |S_{\mathrm{train}}| \cdot L_{\mathsf{MSE}}\left(\varphi, \varrho\right) \tag{6.3}$$

for all snapshots $m_{\mathrm{train}} \in S_{\mathrm{train}}$ due to non-negativity of the respective addends.

In the following we present four examples for snapshot-based generation for the case where $\mathcal{M} = \mathbb{R}^N$, $\check{\mathcal{M}} = \mathbb{R}^n$, $T_m\mathcal{M} = \mathbb{R}^N$, $T_{\check{m}}\check{\mathcal{M}} = \mathbb{R}^n$ are Euclidean vector spaces with chart mappings $x \equiv \mathrm{id}_{\mathbb{R}^N}$, $\check{x} \equiv \mathrm{id}_{\mathbb{R}^n}$ and the metric $d_{\mathcal{M}}$ is defined by a symmetric, positive-definite matrix $\boldsymbol{g} \in \mathbb{R}^{N\times N}$ with

$$\|\boldsymbol{m}\|_{\boldsymbol{g}} := \sqrt{\boldsymbol{m}^\top \boldsymbol{g}\boldsymbol{m}}, \qquad \boldsymbol{d}_{\mathcal{M}}\left(\boldsymbol{m},\, \boldsymbol{w}\right) = \|\boldsymbol{m} - \boldsymbol{w}\|_{\boldsymbol{g}}, \qquad \text{for } \boldsymbol{m}, \boldsymbol{w} \in \mathbb{R}^N.$$

With this choice, the MSE (6.2) in coordinates reads

$$\boldsymbol{L}_{\mathsf{MSE}}\left(\boldsymbol{\varphi}, \boldsymbol{\varrho}\right) = \frac{1}{|S_{\mathrm{train}}|} \sum_{\boldsymbol{m}_{\mathrm{train}} \in S_{\mathrm{train}}} \|\boldsymbol{m}_{\mathrm{train}} - \left(\boldsymbol{\varphi} \circ \boldsymbol{\varrho}\right)\left(\boldsymbol{m}_{\mathrm{train}}\right)\|_{\boldsymbol{g}}^2. \tag{6.4}$$

For each of the four presented approaches, we

(i) formulate the respective family of functions as a subset of $\mathcal{F}_{\varphi,\varrho}$,
(ii) describe how the MSE functional (6.2) is optimized,
(iii) refer to existing work that uses the respective technique.

6.2. **Linear subspaces.** As discussed in Example 3.4, linear-subspace MOR is included in our framework if the embedding $\boldsymbol{\varphi}$ and the point reduction $\boldsymbol{\varrho}$ are linear maps

$$\boldsymbol{\varphi}_{\mathrm{lin}}\left(\check{\boldsymbol{m}}\right) := \boldsymbol{V}\check{\boldsymbol{m}}, \qquad\qquad \boldsymbol{\varrho}_{\mathrm{lin}}\left(\boldsymbol{m}\right) := \boldsymbol{W}^\top \boldsymbol{m}, \tag{6.5}$$

based on the matrices $\boldsymbol{V}, \boldsymbol{W} \in \mathbb{R}^{N\times n}$ with $n \ll N$. Due to the linearity of the mapping $\boldsymbol{\varphi}_{\mathrm{lin}}$, we get that $\boldsymbol{\varphi}_{\mathrm{lin}}\left(\mathbb{R}^n\right)$ is a linear subspace of $\mathcal{M} = \mathbb{R}^N$, which is why we refer to this technique as *linear-subspace MOR*. We formulate the respective family of functions by

$$\mathcal{F}_{\varphi,\varrho,\mathrm{lin}} := \left\{ \left(\boldsymbol{\varphi}_{\mathrm{lin}}, \boldsymbol{\varrho}_{\mathrm{lin}}\right) \text{ from (6.5)} \,\middle|\, \boldsymbol{V}, \boldsymbol{W} \in \mathbb{R}^{N\times n} \text{ such that } \boldsymbol{W}^\top \boldsymbol{V} = \boldsymbol{I}_n \right\}.$$

**Proposition 6.1.** *The family of functions $\mathcal{F}_{\varphi,\varrho,\mathrm{lin}}$ is a subset of $\mathcal{F}_{\varphi,\varrho}$.*

*Proof.* The assumption $\boldsymbol{W}^\top \boldsymbol{V} = \boldsymbol{I}_n$ implies the point projection property (3.4a) with

$$\left(\boldsymbol{\varrho}_{\mathrm{lin}} \circ \boldsymbol{\varphi}_{\mathrm{lin}}\right)\left(\check{\boldsymbol{m}}\right) = \boldsymbol{W}^\top \boldsymbol{V}\check{\boldsymbol{m}} = \check{\boldsymbol{m}}. \qquad \square$$

In the case of a *Galerkin projection*, i.e., $\boldsymbol{W}^\top = \boldsymbol{V}^\top \boldsymbol{g}$, solving (6.1) for $\mathcal{F}_{\varphi,\varrho,\mathrm{lin}}$ simplifies to determine the basis matrix

$$\boldsymbol{V}^* = \underset{\substack{\boldsymbol{V} \in \mathbb{R}^{N\times n} \\ \boldsymbol{V}^\top \boldsymbol{g}\boldsymbol{V} = \boldsymbol{I}_n}}{\arg\min} \sum_{\boldsymbol{m}_{\mathrm{train}} \in S_{\mathrm{train}}} \left\| \left(\boldsymbol{I}_N - \boldsymbol{V}\boldsymbol{V}^\top \boldsymbol{g}\right)\boldsymbol{m}_{\mathrm{train}} \right\|_{\boldsymbol{g}}^2, \tag{6.6}$$

which is known as the *proper orthogonal decomposition* (POD). An optimal solution can be computed with the truncated singular value decomposition (see, e.g., [77]).

For structure-preserving linear-subspace MOR techniques, $\mathcal{F}_{\varphi,\varrho,\mathrm{lin}}$ may have to be restricted to a class that preserves the respective structure. For example, for the structure-preserving MOR for Hamiltonian systems, the SMG is used, which assumes $\boldsymbol{\varphi}$ to be a symplectomorphism and $\boldsymbol{W} = \mathbb{J}_{2N}\boldsymbol{V}\mathbb{J}_{2n}^\top$ is the so-called symplectic inverse. Such MOR techniques are used, e.g., in [14, 47, 60].

6.3. **Quadratic manifolds.** Recently, so-called *MOR on quadratic manifolds* has become an active field of research [4,7,23,37,72]. In our terms, the embedding and point reduction are set to

$$\boldsymbol{\varphi}_{\text{quad}}(\check{\boldsymbol{m}}) := \boldsymbol{A}_2\check{\boldsymbol{m}}^{\otimes^{\text{s}}2} + \boldsymbol{A}_1\check{\boldsymbol{m}} + \boldsymbol{A}_0, \qquad \boldsymbol{\varrho}_{\text{quad}}(\boldsymbol{m}) := \boldsymbol{A}_1^\top(\boldsymbol{m} - \boldsymbol{A}_0), \tag{6.7}$$

with $\boldsymbol{A}_2 \in \mathbb{R}^{N\times n(n+1)/2}$, $\boldsymbol{A}_1 \in \mathbb{R}^{N\times n}$, $\boldsymbol{A}_0 \in \mathbb{R}^N$. By $(\cdot)^{\otimes^{\text{s}}2} \colon \mathbb{R}^n \to \mathbb{R}^{n(n+1)/2}$, $\check{\boldsymbol{m}} \mapsto \check{\boldsymbol{m}}^{\otimes^{\text{s}}2}$, we denote the *symmetric Kronecker product*, which produces all pairwise products of components $[\check{\boldsymbol{m}}]^i \in \mathbb{R}$ of $\check{\boldsymbol{m}}$ for $1 \le i \le n$ while neglecting redundant entries, i.e.,

$$\check{\boldsymbol{m}}^{\otimes^{\text{s}}2} = \Big[[\check{\boldsymbol{m}}]^1 \cdot [\check{\boldsymbol{m}}]^1,\ [\check{\boldsymbol{m}}]^1 \cdot [\check{\boldsymbol{m}}]^2,\ [\check{\boldsymbol{m}}]^2 \cdot [\check{\boldsymbol{m}}]^2,\ \dots,\ [\check{\boldsymbol{m}}]^n \cdot [\check{\boldsymbol{m}}]^n\Big]^\top \in \mathbb{R}^{n(n+1)/2}.$$

The respective family of functions is

$$\mathcal{F}_{\varphi,\varrho,\text{quad}} := \left\{(\boldsymbol{\varphi}_{\text{quad}}, \boldsymbol{\varrho}_{\text{quad}}) \text{ from (6.7)} \,\middle|\, \boldsymbol{A}_1^\top\boldsymbol{A}_1 = \boldsymbol{I}_n \text{ and } \boldsymbol{A}_1^\top\boldsymbol{A}_2 = \boldsymbol{0}_{n\times n(n+1)/2}\right\}. \tag{6.8}$$

**Proposition 6.2.** *The family $\mathcal{F}_{\varphi,\varrho,\text{quad}}$ is a subset of $\mathcal{F}_{\varphi,\varrho}$.*

*Proof.* The assumptions (6.8) on $\boldsymbol{A}_2$ and $\boldsymbol{A}_1$ imply the point projection property (3.4a),

$$(\boldsymbol{\varrho}_{\text{quad}} \circ \boldsymbol{\varphi}_{\text{quad}})(\check{\boldsymbol{m}}) = \boldsymbol{A}_1^\top\left(\boldsymbol{A}_2\check{\boldsymbol{m}}^{\otimes^{\text{s}}2} + \boldsymbol{A}_1\check{\boldsymbol{m}} + \boldsymbol{A}_0 - \boldsymbol{A}_0\right) \overset{(6.8)}{=} \check{\boldsymbol{m}}. \qquad \square$$

The matrices $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ are obtained in [4,7,23,37,72] from the MSE functional (6.4). In this setting, the assumptions in (6.8) allow to determine the matrices $\boldsymbol{A}_0$, $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ sequentially: First, $\boldsymbol{A}_0$ is chosen, e.g., as the mean of $S_{\text{train}}$. Then, (6.4) is optimized for $\boldsymbol{A}_1$ (similarly to (6.6)). Finally, (6.4) is optimized for $\boldsymbol{A}_2$, which results in a (regularized) linear least squares problem. The preceding papers use different tangent reductions to derive the ROM, which can be classified with the framework introduced in the present paper: [4,23,37] use the MPG reduction map (3.11), while [7] relies on the GMG reduction map (5.3) (but neglects a few higher order terms). The major difference between using MPG or GMG in that context is that the MPG projects the FOM vector field with the tangent reduction

$$\boldsymbol{R}_{\mathsf{MPG}}|_{\boldsymbol{\varphi}_{\text{quad}}(\check{\boldsymbol{m}})} = \boldsymbol{D}\boldsymbol{\varrho}_{\text{quad}}|_{\boldsymbol{\varphi}_{\text{quad}}(\check{\boldsymbol{m}})} = \boldsymbol{A}_1^\top$$

which is constant, while the GMG uses the tangent reduction from (5.4) with $\boldsymbol{\tau}|_{\boldsymbol{m}} = \boldsymbol{I}_N$ for all $\boldsymbol{m} \in \mathbb{R}^N$ and $\boldsymbol{D}\boldsymbol{\varphi}|_{\check{\boldsymbol{m}}} = \boldsymbol{A}_2\boldsymbol{B}_2(\check{\boldsymbol{m}}) + \boldsymbol{A}_1$ with a linear function $\boldsymbol{B}_2 \colon \mathbb{R}^n \to \mathbb{R}^{n(n+1)/2\times n}$ describing the derivative of $(\cdot)^{\otimes^{\text{s}}2}$, resulting in

$$\begin{aligned}\boldsymbol{R}_{\mathsf{GMG}}|_{\boldsymbol{\varphi}_{\text{quad}}(\check{\boldsymbol{m}})} &= \left(\boldsymbol{D}\boldsymbol{\varphi}|_{\check{\boldsymbol{m}}}^\top\, \boldsymbol{\tau}|_{\boldsymbol{\varphi}_{\text{quad}}(\check{\boldsymbol{m}})}\boldsymbol{D}\boldsymbol{\varphi}|_{\check{\boldsymbol{m}}}\right)^{-1}\boldsymbol{D}\boldsymbol{\varphi}|_{\check{\boldsymbol{m}}}^\top\, \boldsymbol{\tau}|_{\boldsymbol{\varphi}_{\text{quad}}(\check{\boldsymbol{m}})}\\ &= \left(\boldsymbol{I}_n + (\boldsymbol{B}_2(\check{\boldsymbol{m}}))^\top\boldsymbol{A}_2^\top\boldsymbol{A}_2\boldsymbol{B}_2(\check{\boldsymbol{m}})\right)^{-1}(\boldsymbol{A}_2\boldsymbol{B}_2(\check{\boldsymbol{m}}) + \boldsymbol{A}_1)^\top,\end{aligned}$$

which is typically nonlinear in $\check{\boldsymbol{m}}$, and, thus, so is the reduced vector field in general.

In [72], structure-preserving MOR of Hamiltonian systems on quadratic manifolds is investigated. Two approaches are presented and compared: (i) The *blockwise quadratic* approach uses an embedding of a comparable structure as (6.7) in combination with the MPG tangent reduction. In contrast, (ii) the *quadratic manifold cotangent lift* uses the SMG-ROM. In order to construct a symplectomorphism from a quadratic embedding, the so-called proper symplectic decomposition cotangent lift from [60] (which generates a linear embedding $\boldsymbol{\varphi}$) is generalized to the case of nonlinear embeddings $\boldsymbol{\varphi}$ by introducing the so-called *manifold cotangent lift*. Based on this idea, the authors construct an embedding $\boldsymbol{\varphi} : \mathbb{R}^{2n} \to \mathbb{R}^{2N}$, where the first $N$ component functions are of the structure (6.7) and the last $N$ component functions are rational functions. The SMG (as a special case of the GMG) is then used for a structure-preserving tangent reduction of the Hamiltonian vector field.

6.4. **Nonlinear compressive approximation.** Following the idea of the previous subsection, the embedding and the point reduction can be defined more generally with

$$\varphi_{\mathsf{NCA}}(\check{\boldsymbol{m}}) := \boldsymbol{A}_2 \boldsymbol{f}(\check{\boldsymbol{m}}) + \boldsymbol{A}_1 \check{\boldsymbol{m}} + \boldsymbol{A}_0, \qquad \varrho_{\mathsf{NCA}}(\boldsymbol{m}) := \boldsymbol{B}^\top (\boldsymbol{m} - \boldsymbol{A}_0), \tag{6.9}$$

where $\boldsymbol{A}_2 \in \mathbb{R}^{N \times \widetilde{n}}$, $\boldsymbol{A}_1 \in \mathbb{R}^{N \times n}$, $\boldsymbol{A}_0 \in \mathbb{R}^N$, $\boldsymbol{B} \in \mathbb{R}^{N \times n}$ are matrices, and $\boldsymbol{f} \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\widetilde{n}})$ is a smooth nonlinear mapping for a given $\widetilde{n} \in \mathbb{N}$. Following [19], we refer to this approach as *nonlinear compressive approximation* (NCA). The respective family of functions is

$$\mathcal{F}_{\varphi,\varrho,\mathsf{NCA}} := \left\{ (\varphi_{\mathsf{NCA}}, \varrho_{\mathsf{NCA}}) \text{ from (6.9)} \,\middle|\, \boldsymbol{B}^\top \boldsymbol{A}_1 = \boldsymbol{I}_n,\ \boldsymbol{B}^\top \boldsymbol{A}_2 = \boldsymbol{0}_{n \times \widetilde{n}} \right\}. \tag{6.10}$$

**Proposition 6.3.** *The family $\mathcal{F}_{\varphi,\varrho,\mathsf{NCA}}$ is a subset of $\mathcal{F}_{\varphi,\varrho}$.*

*Proof.* The assumptions on $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, and $\boldsymbol{B}$ in (6.10) imply the point projection property (3.4a) with

$$(\varrho_{\mathsf{NCA}} \circ \varphi_{\mathsf{NCA}})(\check{\boldsymbol{m}}) = \boldsymbol{B}^\top (\boldsymbol{A}_2 \boldsymbol{f}(\check{\boldsymbol{m}}) + \boldsymbol{A}_1 \check{\boldsymbol{m}} + \boldsymbol{A}_0 - \boldsymbol{A}_0) = \check{\boldsymbol{m}}. \qquad \square$$

The MSE for this approach may be optimized sequentially as in the MOR on quadratic manifolds discussed in the previous section using $\boldsymbol{B} = \boldsymbol{A}_1$. This method is, e.g., used in [5], where $\boldsymbol{f}$ is a neural network.

Multiple works investigate NCA: First, the quadratic embedding (6.7) discussed in the previous subsection is a special case of the NCA when choosing $\boldsymbol{f}(\check{\boldsymbol{m}}) = \check{\boldsymbol{m}}^{\otimes^s 2}$. Similarly, $\boldsymbol{f}$ can be chosen as a higher-order polynomial in $\check{\boldsymbol{m}}$ to obtain a more general approximation. Second, in [5], $\boldsymbol{f}$ is learned with an artificial neural network, while a time-discrete setting is considered for the reduction, which is not covered by our methods. Third, [19] analyzes the approximation of a set of traveling wave solutions with (and without) varying support on the PDE level using decision trees and random forests in their numerical experiments. Interestingly, the authors show that a linear point reduction is enough to reproduce the set of traveling wave solutions. Fourth, in [15], it is shown that the NCA has its limitations in terms of the Kolmogorov $(\widetilde{n} + n)$-width since the solution is contained in an $(\widetilde{n} + n)$-dimensional linear subspace of $\mathbb{R}^N$.

6.5. **Autoencoders.** Autoencoders are a well-known technique in nonlinear dimension reduction (see, e.g., [25, Cha. 14]). In the understanding of the present work, MOR with autoencoders chooses

$$\varphi_{\mathsf{AE}} \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^N), \qquad \varrho_{\mathsf{AE}} \in \mathcal{C}^\infty(\mathbb{R}^N, \mathbb{R}^n), \tag{6.11}$$

where both functions are artificial neural networks (ANNs) with network parameters $\boldsymbol{\theta} \in \mathbb{R}^{n_{\boldsymbol{\theta}}}$ (like weights and biases). Since $\varrho_{\mathsf{AE}} \colon \mathbb{R}^N \to \mathbb{R}^n$ and $\varphi_{\mathsf{AE}} \colon \mathbb{R}^n \to \mathbb{R}^N$, the concatenation $\varphi_{\mathsf{AE}} \circ \varrho_{\mathsf{AE}}$ maps from $\mathbb{R}^N$ over $\mathbb{R}^n$ back to $\mathbb{R}^N$. The in-between compression to $\mathbb{R}^n$ is typically referred to as the *bottleneck*, $\varrho_{\mathsf{AE}}$ as the *encoder*, $\varphi_{\mathsf{AE}}$ as the *decoder*, and the concatenation $\varphi_{\mathsf{AE}} \circ \varrho_{\mathsf{AE}}$ as an *autoencoder*. The respective family is

$$\mathcal{F}_{\varphi,\varrho,\mathsf{AE}} := \{ (\varphi_{\mathsf{AE}}, \varrho_{\mathsf{AE}}) \text{ from (6.11)} \mid \boldsymbol{\theta} \in \mathbb{R}^{n_{\boldsymbol{\theta}}} \text{ network parameters} \}.$$

Without special assumptions about the architecture of the ANNs, it is generally impossible to show the point projection property (3.4a). However, whenever the minimum of the cost functional (6.1) is small, then we show in the following that the point projection property (3.4a) holds approximately. We assume to be given a norm $\|\cdot\|_{\check{\boldsymbol{g}}} : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ such that

$\boldsymbol{\varphi}_{\mathsf{AE}}$ and $\boldsymbol{\varrho}_{\mathsf{AE}}$ are *Lipschitz continuous*, i.e., there exists a constant $C_\varphi \geq 0$ such that for all points $\check{\boldsymbol{m}}, \check{\boldsymbol{w}} \in \mathbb{R}^n$

$$\left\|\boldsymbol{\varphi}_{\mathsf{AE}}\left(\check{\boldsymbol{m}}\right) - \boldsymbol{\varphi}_{\mathsf{AE}}\left(\check{\boldsymbol{w}}\right)\right\|_{\boldsymbol{g}} \leq C_\varphi \left\|\check{\boldsymbol{m}} - \check{\boldsymbol{w}}\right\|_{\check{\boldsymbol{g}}} \tag{6.12}$$

and a constant $C_\varrho \geq 0$ such that for all points $\boldsymbol{m}, \boldsymbol{w} \in \mathbb{R}^N$

$$\left\|\boldsymbol{\varrho}_{\mathsf{AE}}\left(\boldsymbol{m}\right) - \boldsymbol{\varrho}_{\mathsf{AE}}\left(\boldsymbol{w}\right)\right\|_{\check{\boldsymbol{g}}} \leq C_\varrho \left\|\boldsymbol{m} - \boldsymbol{w}\right\|_{\boldsymbol{g}}. \tag{6.13}$$

**Theorem 6.4.** *For a given tuple $(\boldsymbol{\varphi}_{\mathsf{AE}}, \boldsymbol{\varrho}_{\mathsf{AE}}) \in \mathcal{F}_{\varphi,\varrho,\mathsf{AE}}$ from the family of functions for $\mathsf{MOR}$ with autoencoders with an $\mathsf{MSE}$ value of $L_{\mathsf{MSE}}(\boldsymbol{\varphi}_{\mathsf{AE}}, \boldsymbol{\varrho}_{\mathsf{AE}}) \geq 0$, the point projection property* (3.4a) *is fulfilled approximately in the sense that for each $\check{\boldsymbol{m}} \in \mathbb{R}^n$*

$$\begin{aligned}\left\|\left(\boldsymbol{\varrho}_{\mathsf{AE}} \circ \boldsymbol{\varphi}_{\mathsf{AE}}\right)\left(\check{\boldsymbol{m}}\right) - \check{\boldsymbol{m}}\right\|_{\check{\boldsymbol{g}}} \leq{}& C_\varrho \sqrt{\left|S_{\mathrm{train}}\right| L_{\mathsf{MSE}}\left(\boldsymbol{\varphi}_{\mathsf{AE}}, \boldsymbol{\varrho}_{\mathsf{AE}}\right)} \\ &+ \left(C_\varrho C_\varphi + 1\right) \min_{\boldsymbol{w}_{\mathrm{train}} \in S_{\mathrm{train}}} \left\|\check{\boldsymbol{m}} - \boldsymbol{\varrho}_{\mathsf{AE}}\left(\boldsymbol{w}_{\mathrm{train}}\right)\right\|_{\check{\boldsymbol{g}}}\end{aligned}$$

Thus, for a bounded set $\check{M} \subsetneq \mathbb{R}^n$, a fine sampling in $S_{\mathrm{train}}$ of $\check{M}$ and small values of the $\mathsf{MSE}$ functional such that the term $\left|S_{\mathrm{train}}\right| L_{\mathsf{MSE}}\left(\boldsymbol{\varphi}_{\mathsf{AE}}, \boldsymbol{\varrho}_{\mathsf{AE}}\right)$ is small, the point projection property (3.4a) holds approximately on $\check{M}$, i.e., $\boldsymbol{\varrho}_{\mathsf{AE}} \circ \boldsymbol{\varphi}_{\mathsf{AE}}|_{\check{M}} \approx \mathrm{id}_{\check{M}}$.

*Proof.* The proof is split in two parts. In the first part, we show that the inequality holds in the encoded training points $\check{\boldsymbol{w}}_{\mathrm{train}} := \boldsymbol{\varrho}_{\mathsf{AE}}\left(\boldsymbol{w}_{\mathrm{train}}\right)$ with $\boldsymbol{w}_{\mathrm{train}} \in S_{\mathrm{train}}$. Then, we show that the inequality holds for general $\check{\boldsymbol{m}} \in \mathbb{R}^n$ by applying Lipschitz continuity.

Consider a fixed but arbitrary training point $\boldsymbol{w}_{\mathrm{train}} \in S_{\mathrm{train}}$. Using (6.13) and (6.3), it holds

$$\begin{aligned}\left\|\left(\boldsymbol{\varrho}_{\mathsf{AE}} \circ \boldsymbol{\varphi}_{\mathsf{AE}}\right)\left(\check{\boldsymbol{w}}_{\mathrm{train}}\right) - \check{\boldsymbol{w}}_{\mathrm{train}}\right\|_{\check{\boldsymbol{g}}} &= \left\|\left(\boldsymbol{\varrho}_{\mathsf{AE}} \circ \boldsymbol{\varphi}_{\mathsf{AE}} \circ \boldsymbol{\varrho}_{\mathsf{AE}}\right)\left(\boldsymbol{w}_{\mathrm{train}}\right) - \boldsymbol{\varrho}_{\mathsf{AE}}\left(\boldsymbol{w}_{\mathrm{train}}\right)\right\|_{\check{\boldsymbol{g}}} \\ &\leq C_\varrho \left\|\left(\boldsymbol{\varphi}_{\mathsf{AE}} \circ \boldsymbol{\varrho}_{\mathsf{AE}}\right)\left(\boldsymbol{w}_{\mathrm{train}}\right) - \boldsymbol{w}_{\mathrm{train}}\right\|_{\boldsymbol{g}} \\ &\leq C_\varrho \sqrt{\left|S_{\mathrm{train}}\right| L_{\mathsf{MSE}}\left(\boldsymbol{\varphi}_{\mathsf{AE}}, \boldsymbol{\varrho}_{\mathsf{AE}}\right)}.\end{aligned}$$

This property can be generalized to points $\check{\boldsymbol{m}} \in \mathbb{R}^n \setminus \boldsymbol{\varrho}_{\mathsf{AE}}\left(S_{\mathrm{train}}\right)$. The idea is that for a general Lipschitz continuous function $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n$ with Lipschitz constant $C \geq 0$ and $\left\|\boldsymbol{f}\left(\check{\boldsymbol{w}}_{\mathrm{train}}\right)\right\|_{\check{\boldsymbol{g}}} \leq C_B$ for some $C_B \geq 0$, it holds for all $\check{\boldsymbol{m}} \in \mathbb{R}^n$ by adding a zero, triangle inequality, and Lipschitz continuity

$$\left\|\boldsymbol{f}\left(\check{\boldsymbol{m}}\right)\right\|_{\check{\boldsymbol{g}}} \leq \left\|\boldsymbol{f}\left(\check{\boldsymbol{w}}_{\mathrm{train}}\right)\right\|_{\check{\boldsymbol{g}}} + \left\|\boldsymbol{f}\left(\check{\boldsymbol{m}}\right) - \boldsymbol{f}\left(\check{\boldsymbol{w}}_{\mathrm{train}}\right)\right\|_{\check{\boldsymbol{g}}} \leq C_B + C \left\|\check{\boldsymbol{m}} - \check{\boldsymbol{w}}_{\mathrm{train}}\right\|_{\check{\boldsymbol{g}}}.$$

For our case, we use $\boldsymbol{f}\left(\check{\boldsymbol{m}}\right) = \left(\boldsymbol{\varrho}_{\mathsf{AE}} \circ \boldsymbol{\varphi}_{\mathsf{AE}}\right)\left(\check{\boldsymbol{m}}\right) - \check{\boldsymbol{m}}$ with Lipschitz constant $C = C_\varrho C_\varphi + 1$, bound $C_B = C_\varrho \sqrt{\left|S_{\mathrm{train}}\right| L_{\mathsf{MSE}}\left(\boldsymbol{\varphi}_{\mathsf{AE}}, \boldsymbol{\varrho}_{\mathsf{AE}}\right)}$, and points $\check{\boldsymbol{w}}_{\mathrm{train}} = \boldsymbol{\varrho}_{\mathsf{AE}}\left(\boldsymbol{w}_{\mathrm{train}}\right)$. Thus, it holds

$$\begin{aligned}\left\|\left(\boldsymbol{\varrho}_{\mathsf{AE}} \circ \boldsymbol{\varphi}_{\mathsf{AE}}\right)\left(\check{\boldsymbol{m}}\right) - \check{\boldsymbol{m}}\right\|_{\check{\boldsymbol{g}}} \leq{}& C_\varrho \sqrt{\left|S_{\mathrm{train}}\right| L_{\mathsf{MSE}}\left(\boldsymbol{\varphi}_{\mathsf{AE}}, \boldsymbol{\varrho}_{\mathsf{AE}}\right)} \\ &+ \left(C_\varrho C_\varphi + 1\right) \left\|\check{\boldsymbol{m}} - \boldsymbol{\varrho}_{\mathsf{AE}}\left(\boldsymbol{w}_{\mathrm{train}}\right)\right\|_{\check{\boldsymbol{g}}}\end{aligned}$$

Taking the minimum over all $\boldsymbol{w}_{\mathrm{train}} \in S_{\mathrm{train}}$ on the right-hand side yields the result. □

**Remark 6.5** (Constrained autoencoders)**.** *In [57], the authors introduce a novel autoencoder architecture, which aims at fulfilling the point projection property* (3.4a) *exactly. The architecture of the encoder is chosen to invert the decoder layer-wise based on the assumption that the linear layers of the decoder and encoder are pairwise biorthogonal. This biorthogonality is penalized in the loss functional with an additional term.*

One of the first works to combine autoencoders with projection-based MOR is [46]. As discussed in Section 5.1, the time-continuous formulation in that work considers the GMG reduction for a state-independent Riemannian metric. A structure-preserving formulation for Hamiltonian systems in combination with autoencoders is discussed in [16]. As shown in Lemma 5.12, this work is based on the SMG reduction.

## 7. Conclusions

This work proposed a differential geometric framework for MOR on manifolds in order to analyze two important efforts in MOR jointly: (i) The use of nonlinear projections and (ii) structure preservation. The key ingredient for our framework is an embedding for a low-dimensional submanifold and a compatible reduction map. The joint abstraction allowed us to derive shared theoretical properties, such as the exact reproduction result. As two possible reduction mappings, we discussed (a) the *manifold Petrov–Galerkin* (MPG) using the differential of the point reduction and (b) the *generalized manifold Galerkin* (GMG), which is based on a nondegenerate tensor field. Moreover, we showed that structure-preserving MOR on manifolds for Lagrangian and Hamiltonian systems can be accomplished by choosing specific nondegenerate tensor fields in the GMG reduction map, which we refer to as the *Lagrangian manifold Galerkin* (LMG) and *symplectic manifold Galerkin* (SMG). In order to connect our framework to existing work in the field, we demonstrated how different techniques for data-driven construction of the embedding and point reduction map are reflected in our approach. We discussed four approximation types (linear, quadratic, nonlinear compressive, and autoencoders) and linked each of these types to multiple existing works in the field.

We believe that our framework can be extended in several regards: First, other structure-preserving MOR techniques might be formulated in the framework, such as Poisson systems [36], port-Hamiltonian (descriptor) systems [51, 76], or (port-)metriplectic systems [28, 34, 54]. Thereby, all the described nonlinear projections (quadratic, nonlinear compressive, autoencoders) become available for such structured systems. Second, as the high-dimensional differential equations we consider as FOM are often obtained from the semi-discretization of systems of partial differential equations (PDEs), the framework should be extended to the PDE level. Such a formulation would forward the structures formulated on the PDE level to the ODE level.

**Acknowledgements** PB, BH, and BU are funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2075 - 390740016 and acknowledge the support by the Stuttgart Center for Simulation Science (SimTech). PB and BH acknowledge the funding of DFG Project No. 314733389. BU acknowledges funding by the BMBF (grant no. 05M22VSA).

## References

[1] R. Abraham and J. E. Marsden. *Foundations of Mechanics*. Addison-Wesley Publishing Company, Redwood City, CA, second edition, 1987.

[2] D. Amsallem, J. Cortial, K. Carlberg, and C. Farhat. A method for interpolating on manifolds structural dynamics reduced-order models. *Internat. J. Numer. Methods Engrg.*, 80(9), 2009.

[3] A. C. Antoulas, C. Beattie, and S. Gugercin. *Interpolatory Methods for Model Reduction*. SIAM, Philadelphia, PA, USA, 2020.

[4] J. Barnett and C. Farhat. Quadratic approximation manifold for mitigating the Kolmogorov barrier in nonlinear projection-based model order reduction. *J. Comput. Phys.*, 464:111348, 2022.

[5] J. L. Barnett, C. Farhat, and Y. Maday. Mitigating the Kolmogorov barrier for the reduction of aerodynamic models using neural-network-augmented reduced-order models. In *AIAA SCITECH 2023 Forum*, 2023.

[6] P. Benner, A. Cohen, M. Ohlberger, and K. Willcox. *Model Reduction and Approximation*. Advances in Design and Control. SIAM, Philadelphia, PA, USA, 2017.

[7] P. Benner, P. Goyal, J. Heiland, and I. Pontes Duff. A quadratic decoder approach to nonintrusive reduced-order modeling of nonlinear dynamical systems. *PAMM*, 23(1):e202200049, 2023.

[8] W. J. Beyn and V. Thümmler. Freezing solutions of equivariant evolution equations. *SIAM J. Appl. Dyn. Syst.*, 3(2):85–116, 2004.

[9] R. Bishop and S. Goldberg. *Tensor analysis on manifolds*. Macmillan, New York, 1968.

[10] F. Black, P. Schulze, and B. Unger. Projection-based model reduction with dynamically transformed modes. *ESAIM: Math. Model. Numer. Anal.*, 54(6):2011–2043, 2020.

[11] F. Black, P. Schulze, and B. Unger. Efficient wildland fire simulation via nonlinear model order reduction. *Fluids*, 6(8):280, 2021.

[12] F. Black, P. Schulze, and B. Unger. Modal decomposition of flow data via gradient-based transport optimization. In R. King and D. Peitsch, editors, *Active Flow and Combustion Control 2021*, pages 203–224, Cham, 2022. Springer International Publishing.

[13] G. Boncoraglio and C. Farhat. Active manifold and model-order reduction to accelerate multidisciplinary analysis and optimization. *AIAA Journal*, 59(11):4739–4753, 2021.

[14] P. Buchfink, A. Bhatt, and B. Haasdonk. Symplectic model order reduction with non-orthonormal bases. *Math. Comput. Appl.*, 24(2), 2019.

[15] P. Buchfink, S. Glas, and B. Haasdonk. Approximation bounds for model reduction on polynomially mapped manifolds. *ArXiv e-print 2312.00724*, 2023.

[16] P. Buchfink, S. Glas, and B. Haasdonk. Symplectic model reduction of Hamiltonian systems on nonlinear manifolds and approximation with weakly symplectic autoencoder. *SIAM J. Sci. Comput.*, 45(2):A289–A311, 2023.

[17] N. Cagniart, Y. Maday, and B. Stamm. Model order reduction for problems with large convection effects. In B. N. Chetverushkin, W. Fitzgibbon, Y. A. Kuznetsov, P. Neittaanmäki, J. Periaux, and O. Pironneau, editors, *Contributions to Partial Differential Equations and Applications*, Computational Methods in Applied Sciences, pages 131–150. Springer-Verlag, Cham, Switzerland, 2019.

[18] K. Carlberg, R. Tuminaro, and P. Boggs. Preserving Lagrangian structure in nonlinear model reduction with application to structural dynamics. *SIAM J. Sci. Comput.*, 37(2):B153–B184, 2015.

[19] A. Cohen, C. Farhat, Y. Maday, and A. Somacal. Nonlinear compressive reduced basis approximation for PDE's. *Comptes Rendus. Mécanique*, pages 1–8, 2023. Online first.

[20] P. G. Constantine. *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*. SIAM Spotlights. SIAM, 2015.

[21] D. M. Copeland, S. W. Cheung, K. Huynh, and Y. Choi. Reduced order models for Lagrangian hydrodynamics. *Comput. Meth. Appl. Mech. Eng.*, 388:114259, 2022.

[22] A. Ferrero, T. Taddei, and L. Zhang. Registration-based model reduction of parameterized two-dimensional conservation laws. *J. Comput. Phys.*, 457:111068, 2022.

[23] R. Geelen, S. Wright, and K. Willcox. Operator inference for non-intrusive model reduction with quadratic manifolds. *Comput. Meth. Appl. Mech. Eng.*, 403:115717, 2023.

[24] Y. Gong, Q. Wang, and Z. Wang. Structure-preserving Galerkin POD reduced-order modeling of Hamiltonian systems. *Comput. Meth. Appl. Mech. Eng.*, 315:780–798, 2017.

[25] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.

[26] C. Greif and K. Urban. Decay of the Kolmogorov $n$-width for wave problems. *Appl. Math. Lett.*, 96:216–222, 2019.

[27] S. Greydanus, M. Dzamba, and J. Yosinski. Hamiltonian neural networks. In *NIPS'19: Proceedings of the 33rd International Conference on Neural Information Processing Systems*, pages 15379–15389, 2019.

[28] A. Gruber, M. Gunzburger, L. Ju, and Z. Wang. Energetically consistent model reduction for metriplectic systems. *Computer Methods in Applied Mechanics and Engineering*, 404:115709, 2023.

[29] A. Gruber and I. Tezaur. Canonical and noncanonical Hamiltonian operator inference. *Comput. Meth. Appl. Mech. Eng.*, 416:116334, 2023.

[30] C. Gu. QLMOR: A projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems. *IEEE Trans. Comput.-Aided Des. Integr.*, 30(9):1307–1320, 2011.
[31] B. Haasdonk. Reduced basis methods for parametrized PDEs—a tutorial introduction for stationary and instationary problems. In P. Benner, M. Ohlberger, A. Cohen, and K. Willcox, editors, *Model Reduction and Approximation*, pages 65–136. SIAM, 2017.
[32] H. Harma, Z. Wang, and B. Kramer. Hamiltonian operator inference: Physics-preserving learning of reduced-order models for canonical Hamiltonian systems. *Phys. D*, 431, 2022.
[33] D. Hartmann and L. K. Mestha. A deep learning framework for model reduction of dynamical systems. In *Proc. IEEE Conference on Control Technology and Applications (CCTA), Kohala Coast, USA*, pages 1917–1922, 2017.
[34] Q. Hernández, A. Badías, F. Chinesta, and E. Cueto. Port-metriplectic neural networks: thermodynamics-informed machine learning of complex physical systems. *Comput. Mech.*, 72:553–561, 2023.
[35] J. S. Hesthaven, G. Rozza, and B. Stamm. *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*. SpringerBriefs in Mathematics. Springer-Verlag, Cham, Switzerland, 2016.
[36] J. S. Hesthaven and C. Pagliantini. Structure-preserving reduced basis methods for Poisson systems. *Math. Comp.*, 90(330):1701–1740, 2021.
[37] O. Issan and B. Kramer. Predicting solar wind streams from the inner-heliosphere to earth via shifted operator inference. *J. Comput. Phys.*, 473:111689, 2023.
[38] K. Kashima. Nonlinear model reduction by deep autoencoder of noise response data. In *Proc. 55th IEEE Conference on Decision and Control (CDC), Las Vegas, USA*, pages 5750–5755, 2016.
[39] Y. Kim, Y. Choi, D. Widemann, and T. Zohdi. A fast and accurate physics-informed neural network reduced order model with shallow masked autoencoder. *J. Comput. Phys.*, 451:110841, 2022.
[40] M. Kirby and D. Armbruster. Reconstructing phase space from PDE simulations. *Z. Angew. Math. Phys.*, 42:999–1022, 1992.
[41] A. Kolmogoroff. Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse. *Ann. of Math. (2)*, 37(1):107–110, 1936.
[42] P. Krah, S. Büchholz, M. Häringer, and J. Reiss. Front transport reduction for complex moving fronts. *J. Sci. Comput.*, 96:28, 2023.
[43] B. Kramer and K. E. Willcox. Nonlinear model order reduction via lifting transformations and proper orthogonal decomposition. *AIAA Journal*, 57(6):2297–2307, 2019.
[44] S. Lall, P. Krysl, and J. E. Marsden. Structure-preserving model reduction for mechanical systems. *Phys. D*, 184(1):304–318, 2003.
[45] J. M. Lee. *Introduction to Smooth Manifolds*. Graduate Texts in Mathematics. Springer-Verlag, New York, NY, 2nd edition, 2012.
[46] K. Lee and K. T. Carlberg. Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders. *J. Comput. Phys.*, 404:108973, 2020.
[47] B. Maboudi Afkham and J. S. Hesthaven. Structure preserving model reduction of parametric Hamiltonian systems. *SIAM J. Sci. Comput.*, 39(6):A2616–A2644, 2017.
[48] B. Maboudi Afkham, A. Bhatt, B. Haasdonk, and J. S. Hesthaven. Symplectic model-reduction with a weighted inner product. *ArXiv e-print 1803.07799*, 2018.
[49] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*. Springer New York, 1999.
[50] E. Massart, P.-Y. Gousenbourger, N. T. Son, T. Stykel, and P.-A. Absil. Interpolation on the manifold of fixed-rank positive-semidefinite matrices for parametric model order reduction: preliminary results. In *Proceedings of the 27th European Symposium on Artifical Neural Networks, Computational Intelligence and Machine Learning (ESANN2019)*, pages 281–286, 2019.
[51] V. Mehrmann and B. Unger. Control of port-Hamiltonian differential-algebraic systems and applications. *Acta Numer.*, 32:395–515, 2023.
[52] A. Mendible, S. L. Brunton, A. Y. Aravkin, W. Lowrie, and J. N. Kutz. Dimensionality reduction and reduced-order modeling for traveling wave physics. *Theor. Comput. Fluid Dyn.*, 34:385–400, 2020.
[53] R. Mojgani and M. Balajewicz. Lagrangian basis method for dimensionality reduction of convection dominated nonlinear flows. *ArXiv e-print 1701.04343v1*, 2017.

[54] P. J. Morrison. A paradigm for joined Hamiltonian and dissipative systems. *Phys. D*, 18(1):410–419, 1986.
[55] M. Ohlberger and S. Rave. Nonlinear reduced basis approximation of parameterized evolution equations via the method of freezing. *C. R. Math. Acad. Sci. Paris*, 351(23–24):901–906, 2013.
[56] M. Ohlberger and S. Rave. Reduced basis methods: Success, limitations and future challenges. *Proceedings of the Conference Algoritmy*, pages 1–12, 2016.
[57] S. E. Otto, G. R. Macchio, and C. W. Rowley. Learning nonlinear projections for reduced-order modeling of dynamical systems using constrained autoencoders. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 33(11):113130, 2023.
[58] S. E. Otto, A. Padovan, and C. W. Rowley. Optimizing oblique projections for nonlinear systems using trajectories. *SIAM J. Sci. Comput.*, 44(3):A1681–A1702, 2022.
[59] M. Oulghelou and C. Allery. Non intrusive method for parametric model order reduction using a bi-calibrated interpolation on the Grassmann manifold. *J. Comput. Phys.*, 426:109924, 2021.
[60] L. Peng and K. Mohseni. Symplectic model reduction of Hamiltonian systems. *SIAM J. Sci. Comput.*, 38(1):A1–A27, 2016.
[61] A. Pinkus. *N-widths in approximation theory*. Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer-Verlag, Heidelberg, Germany, 1985.
[62] E. Qian, B. Kramer, B. Peherstorfer, and K. Willcox. Lift & Learn: Physics-informed machine learning for large-scale nonlinear dynamical systems. *Phys. D*, 406:132401, 2020.
[63] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced Basis Methods for Partial Differential Equations: An Introduction*. UNITEXT. Springer-Verlag, Cham, Switzerland, 2016.
[64] A. Quarteroni and G. Rozza, editors. *Reduced Order Methods for Modeling and Computational Reduction*. Number 9 in MS&A - Modeling, simulation and applications. Springer, Cham, Switzerland, 2014.
[65] J. Reiss, P. Schulze, J. Sesterhenn, and V. Mehrmann. The shifted proper orthogonal decomposition: A mode decomposition for multiple transport phenomena. *SIAM J. Sci. Comput.*, 40(3):A1322–A1344, 2018.
[66] D. Rim, S. Moe, and R. J. LeVeque. Transport reversal for model reduction of hyperbolic partial differential equations. *SIAM/ASA J. Uncertain. Quantif.*, 6(1):118–150, 2018.
[67] C. W. Rowley and J. E. Marsden. Reconstruction equations and the Karhunen–Loève expansion for systems with symmetry. *Phys. D*, 142(1–2):1–19, 2000.
[68] T. Ruiner, J. Fehr, B. Haasdonk, and P. Eberhard. A-posteriori error estimation for second order mechanical systems. *Acta Mech. Sin.*, 28(3):854–862, 2012.
[69] P. Schulze. *Energy-based model reduction of transport-dominated phenomena*. Dissertation, Technische Universität Berlin, Institut für Mathematik, 2023.
[70] P. Schulze, J. Reiss, and V. Mehrmann. Model reduction for a pulsed detonation combuster via shifted proper orthogonal decomposition. In R. King, editor, *Active Flow and Combustion Control 2018*, pages 271–286. Springer, Cham, Switzerland, 2019.
[71] P. Schwerdtner, P. Schulze, J. Berman, and B. Peherstorfer. Nonlinear embeddings for conserving Hamiltonians and other quantities with Neural Galerkin schemes. *ArXiv e-print 2310.07485*, 2023.
[72] H. Sharma, H. Mu, P. Buchfink, R. Geelen, S. Glas, and B. Kramer. Symplectic model reduction of Hamiltonian systems using data-driven quadratic manifolds. *Comput. Meth. Appl. Mech. Eng.*, 417:116402, 2023.
[73] N. T. Son. A real time procedure for affinely dependent parametric model order reduction using interpolation on Grassmann manifolds. *Internat. J. Numer. Methods Engrg.*, 93(8):818–833, 2012.
[74] T. Taddei. A registration method for model order reduction: Data compression and geometry reduction. *SIAM J. Sci. Comput.*, 42(2):A997–A1027, 2020.
[75] B. Unger and S. Gugercin. Kolmogorov $n$-widths for linear dynamical systems. *Adv. Comput. Math.*, 45(5–6):2273–2286, 2019.
[76] A. van der Schaft and D. Jeltsema. Port-Hamiltonian systems theory: An introductory overview. *Foundations and Trends in Systems and Control*, 1(2-3):173–378, 2014.
[77] S. Volkwein. Proper orthogonal decomposition: Theory and reduced-order modelling, August 2013. Lecture Notes, University of Konstanz.

[78] S. Yildiz, P. Goyal, T. Bendokat, and P. Benner. Data-driven identification of quadratic symplectic representations of nonlinear Hamiltonian systems. *ArXiv e-print 2308.01084*, 2023.
[79] S. Yildiz, P. Goyal, and P. Benner. Linearly implicit global energy preserving reduced-order models for cubic Hamiltonian systems. *ArXiv e-print 2308.02625*, 2023.
[80] R. Zimmermann. Manifold interpolation. In P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. Schilders, and L. M. Silveira, editors, *Model Order Reduction - Volume 1: System- and Data-Driven Methods and Algorithms*, pages 229–274. De Gruyter, Berlin, Boston, 2021.

## Appendix A. Topological spaces and topological manifolds

A.1. **Fundamentals.** Consider a set $\mathcal{M}$. A *topology on* $\mathcal{M}$ is a collection $\mathcal{T}$ of subsets of $\mathcal{M}$ (which are called *open subsets of* $\mathcal{M}$) that satisfy that (i) both the empty set $\emptyset$ and the set itself $\mathcal{M}$ are open, (ii) each union of open subsets is open, (iii) each intersection of finitely many open subsets is open. The pair $(\mathcal{M}, \mathcal{T})$ is called a *topological space*. If the specific topology is clear from the context or not particularly relevant for the discussion, then we simply write $\mathcal{M}$ instead of $(\mathcal{M}, \mathcal{T})$.

For two topological spaces $\mathcal{M}$ and $\mathcal{Q}$, a map $F\colon \mathcal{M} \to \mathcal{Q}$ is called *continuous*, if for every open subset $V \subseteq \mathcal{Q}$, the *preimage* $\{m \in \mathcal{M} \mid F(m) \in V\}$ is open in $\mathcal{M}$. We call $F$ a *homeomorphism*, if (i) it is bijective (and thus the *inverse* $F^{-1}\colon \mathcal{Q} \to \mathcal{M}$ exists) and (ii) both $F$ and $F^{-1}$ are continuous. Correspondingly, two topological spaces $\mathcal{M}$ and $\mathcal{Q}$ are called *homeomorphic* if there exists a homeomorphism from $\mathcal{M}$ to $\mathcal{Q}$. Moreover, $\mathcal{M}$ is called *locally homeomorphic to* $\mathbb{R}^N$ for $N \in \mathbb{N}$ if for every point $m \in \mathcal{M}$ there exists an open set $U \subseteq \mathcal{M}$ with $m \in U$, which is homeomorphic to an open subset of $\mathbb{R}^N$.

A topological space $\mathcal{M}$ is called a *topological manifold* of *dimension* $N$ if it is locally homeomorphic to $\mathbb{R}^N$ (and additionally Hausdorff and second-countable, see e.g. [45, Cha. 1 and App. A]).

A.2. **Proof of Lemma 2.1.** By assumption, $\varphi \in \mathcal{C}^\infty(\check{\mathcal{M}}, \mathcal{M})$ and $\varrho \in \mathcal{C}^\infty(\mathcal{M}, \check{\mathcal{M}})$ are smooth maps. Then, the restrictions to $\varphi(\check{\mathcal{M}}) \subseteq \mathcal{M}$ are smooth maps in the subspace topology, i.e., $\varphi \in \mathcal{C}^\infty(\mathcal{M}, \varphi(\check{\mathcal{M}}))$ and $\varrho|_{\varphi(\check{\mathcal{M}})} \in \mathcal{C}^\infty(\varphi(\check{\mathcal{M}}), \check{\mathcal{M}})$. Thus, $\varphi$ is a smooth diffeomorphism onto its image in the subspace topology. By [45, Prop. 4.8. (a)], $\varphi$ is a smooth immersion and thus a smooth embedding.

## Appendix B. Proofs for Lagrangian systems

B.1. **Derivation of the reduced Euler-Lagrange equations.** The derivatives of $\check{\mathscr{L}}$ can be computed for $\check{Y}_{\mathcal{Q}} = \left(\check{q}, \check{v}^{\underline{i}} \frac{\partial}{\partial \check{y}^{\underline{i}}}\Big|_{\check{q}}\right) \in T\check{\mathcal{Q}}$ with $1 \leq j, k \leq Q$ and $1 \leq i \leq \check{Q}$ as

$$\begin{aligned}
\left.\frac{\partial \check{\mathscr{L}}}{\partial \check{\xi}^{\underline{i}}}\right|_{\check{Y}_{\mathcal{Q}}} &= \left.\frac{\partial \mathscr{L}}{\partial \xi^{\underline{j}}}\right|_{\varphi(\check{Y}_{\mathcal{Q}})} \left.\frac{\partial \varphi^{\underline{j}}}{\partial \check{\xi}^{\underline{i}}}\right|_{\check{Y}_{\mathcal{Q}}} + \left.\frac{\partial \mathscr{L}}{\partial \xi^{\underline{Q+j}}}\right|_{\varphi(\check{Y}_{\mathcal{Q}})} \left.\frac{\partial \varphi^{\underline{Q+j}}}{\partial \check{\xi}^{\underline{i}}}\right|_{\check{Y}_{\mathcal{Q}}} \\
&= \left.\frac{\partial \mathscr{L}}{\partial \xi^{\underline{j}}}\right|_{\varphi(\check{Y}_{\mathcal{Q}})} \left.\frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{i}}}\right|_{\check{q}} + \left.\frac{\partial \mathscr{L}}{\partial \xi^{\underline{Q+j}}}\right|_{\varphi(\check{Y}_{\mathcal{Q}})} \left.\frac{\partial^2 {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{i}} \partial \check{y}^{\underline{k}}}\right|_{\check{q}} \check{v}^{\underline{k}} \\
\left.\frac{\partial \check{\mathscr{L}}}{\partial \check{\xi}^{\underline{\check{Q}+i}}}\right|_{\check{Y}_{\mathcal{Q}}} &= \left.\frac{\partial \mathscr{L}}{\partial \xi^{\underline{j}}}\right|_{\varphi(\check{Y}_{\mathcal{Q}})} \left.\frac{\partial \varphi^{\underline{j}}}{\partial \check{\xi}^{\underline{\check{Q}+i}}}\right|_{\check{Y}_{\mathcal{Q}}} + \left.\frac{\partial \mathscr{L}}{\partial \xi^{\underline{Q+j}}}\right|_{\varphi(\check{Y}_{\mathcal{Q}})} \left.\frac{\partial \varphi^{\underline{Q+j}}}{\partial \check{\xi}^{\underline{\check{Q}+i}}}\right|_{\check{Y}_{\mathcal{Q}}} \\
&= \left.\frac{\partial \mathscr{L}}{\partial \xi^{\underline{Q+j}}}\right|_{\varphi(\check{Y}_{\mathcal{Q}})} \left.\frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{i}}}\right|_{\check{q}}.
\end{aligned}$$

Evaluation for the lifted curve $\Gamma_{\check{\gamma}_{\mathcal{Q}}} \in \mathcal{C}^\infty(\mathcal{I}, T\check{\mathcal{Q}})$ and derivation with respect to the time, yields for $1 \leq j,k \leq Q$ and $1 \leq i,\ell,p \leq \check{Q}$

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left( \frac{\partial \check{\mathscr{L}}}{\partial \check{\xi}^{\underline{\check{Q}+i}}}\bigg|_{\Gamma_{\check{\gamma}_{\mathcal{Q}}(\cdot)}} \right)\bigg|_t
&= \frac{\mathrm{d}}{\mathrm{d}t}\left( \frac{\partial \mathscr{L}}{\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(\cdot)}\right)} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{i}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(\cdot)} \right)\bigg|_t \\
&= \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{k}}\,\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{i}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\partial \varphi^{\underline{k}}}{\partial \check{\xi}^{\underline{\ell}}}\bigg|_{\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}} \frac{\mathrm{d}}{\mathrm{d}t}\left( {\Gamma_{\check{\gamma}_{\mathcal{Q}}(\cdot)}}^{\underline{\ell}} \right)\bigg|_t \\
&\quad + \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{k}}\,\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{i}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\partial \varphi^{\underline{k}}}{\partial \check{\xi}^{\underline{\check{Q}+\ell}}}\bigg|_{\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}} \frac{\mathrm{d}}{\mathrm{d}t}\left( {\Gamma_{\check{\gamma}_{\mathcal{Q}}(\cdot)}}^{\underline{\check{Q}+\ell}} \right)\bigg|_t \\
&\quad + \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{Q+k}}\,\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{i}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\partial \varphi^{\underline{Q+k}}}{\partial \check{\xi}^{\underline{\ell}}}\bigg|_{\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}} \frac{\mathrm{d}}{\mathrm{d}t}\left( {\Gamma_{\check{\gamma}_{\mathcal{Q}}(\cdot)}}^{\underline{\ell}} \right)\bigg|_t \\
&\quad + \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{Q+k}}\,\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{i}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\partial \varphi^{\underline{Q+k}}}{\partial \check{\xi}^{\underline{\check{Q}+\ell}}}\bigg|_{\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}} \frac{\mathrm{d}}{\mathrm{d}t}\left( {\Gamma_{\check{\gamma}_{\mathcal{Q}}(\cdot)}}^{\underline{\check{Q}+\ell}} \right)\bigg|_t \\
&\quad + \frac{\partial \mathscr{L}}{\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} \frac{\partial^2 {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{k}}\,\partial \check{y}^{\underline{i}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\mathrm{d}}{\mathrm{d}t}\left( {\Gamma_{\check{\gamma}_{\mathcal{Q}}(\cdot)}}^{\underline{k}} \right)\bigg|_t \\
&= \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{k}}\,\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{i}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\mathrm{d}}{\mathrm{d}t} {\check{\gamma}_{\mathcal{Q}}}^{\underline{\ell}}\bigg|_t \\
&\quad + \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{Q+k}}\,\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{i}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\partial^2 {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}\,\partial \check{y}^{\underline{p}}}\bigg|_{\check{q}} \frac{\mathrm{d}}{\mathrm{d}t} {\check{\gamma}_{\mathcal{Q}}}^{\underline{p}}\bigg|_t \frac{\mathrm{d}}{\mathrm{d}t} {\check{\gamma}_{\mathcal{Q}}}^{\underline{\ell}}\bigg|_t \\
&\quad + \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{Q+k}}\,\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{i}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\mathrm{d}^2}{\mathrm{d}t^2} {\check{\gamma}_{\mathcal{Q}}}^{\underline{\ell}}\bigg|_t \\
&\quad + \frac{\partial \mathscr{L}}{\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} \frac{\partial^2 {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{\ell}}\,\partial \check{y}^{\underline{i}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\mathrm{d}}{\mathrm{d}t} {\check{\gamma}_{\mathcal{Q}}}^{\underline{\ell}}\bigg|_t
\end{aligned}
$$

In total, it holds for $1 \leq j,k \leq Q$ and $1 \leq i,\ell,p \leq \check{Q}$ in

$$
\begin{aligned}
0 &= \frac{\partial \check{\mathscr{L}}}{\partial \check{\xi}^{\underline{i}}}\bigg|_{\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}} - \frac{\mathrm{d}}{\mathrm{d}t}\left( \frac{\partial \check{\mathscr{L}}}{\partial \check{\xi}^{\underline{\check{Q}+i}}}\bigg|_{\Gamma_{\check{\gamma}_{\mathcal{Q}}(\cdot)}} \right)\bigg|_t \\
&= \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{i}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \Bigg( \frac{\partial \mathscr{L}}{\partial \xi^{\underline{j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} - \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{k}}\,\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\mathrm{d}}{\mathrm{d}t} {\check{\gamma}_{\mathcal{Q}}}^{\underline{\ell}}\bigg|_t \\
&\qquad - \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{Q+k}}\,\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} \frac{\partial^2 {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}\,\partial \check{y}^{\underline{p}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}|_t} \frac{\mathrm{d}}{\mathrm{d}t} {\check{\gamma}_{\mathcal{Q}}}^{\underline{p}}\bigg|_t \frac{\mathrm{d}}{\mathrm{d}t} {\check{\gamma}_{\mathcal{Q}}}^{\underline{\ell}}\bigg|_t \\
&\qquad - \frac{\partial^2 \mathscr{L}}{\partial \xi^{\underline{Q+k}}\,\partial \xi^{\underline{Q+j}}}\bigg|_{\varphi_{\mathcal{Q}}\left(\Gamma_{\check{\gamma}_{\mathcal{Q}}(t)}\right)} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}}\bigg|_{\check{\gamma}_{\mathcal{Q}}(t)} \frac{\mathrm{d}^2}{\mathrm{d}t^2} {\check{\gamma}_{\mathcal{Q}}}^{\underline{\ell}}\bigg|_t \Bigg).
\end{aligned}
$$

B.2. **Proof of Theorem 5.7.** In order to show that the systems are equivalent, we show that the underlying vector fields are identical, i.e., we show that the LMG reduction (5.11) of the Euler–Lagrange vector field (4.11) results in the reduced Euler–Lagrangian vector field (5.8). To simplify the notation, we use $\tau = \tau_{\mathsf{LMG}}$ and $\check{\tau} = \check{\tau}_{\mathsf{LMG}}$ in the following, with

$\tau_{\mathsf{LMG}}$ as in (5.9). Let $\check{Y}_{\mathcal{Q}} = (\check{q}, \check{v}) = \big(\check{q}, \check{v}^{\underline{i}} \frac{\partial}{\partial \check{y}^{\underline{i}}}\big|_{\check{q}}\big) \in T\check{\mathcal{Q}}$. The reduced tensor field $\check{\tau} = \mathrm{d}\varphi^* \tau$ reads for $1 \leq \alpha, \beta \leq 2Q$

$$
\begin{aligned}
\check{\tau}|_{\check{Y}_{\mathcal{Q}}} &= \frac{\partial \varphi^{\underline{\gamma}}}{\partial \xi^{\underline{\alpha}}}\Big|_{\check{Y}_{\mathcal{Q}}} \left(\tau|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{\gamma\delta}} \frac{\partial \varphi^{\underline{\delta}}}{\partial \xi^{\underline{\beta}}}\Big|_{\check{Y}_{\mathcal{Q}}} \mathrm{d}\xi^{\underline{\alpha}}|_{\check{Y}_{\mathcal{Q}}} \otimes \mathrm{d}\xi^{\underline{\beta}}\big|_{\check{Y}_{\mathcal{Q}}} \\
&= \left(\tau_q|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{k\ell}} \left( \frac{\partial^2 \varphi_{\mathcal{Q}}{}^{\underline{k}}}{\partial \check{y}^{\underline{p}} \partial \check{y}^{\underline{i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \check{v}^{\underline{p}} \, \mathrm{d}\check{\xi}^{\underline{i}}\Big|_{\check{Y}_{\mathcal{Q}}} + \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{k}}}{\partial \check{y}^{\underline{i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \mathrm{d}\check{\xi}^{\underline{\check{Q}+i}}\Big|_{\check{Y}_{\mathcal{Q}}} \right) \otimes \left( \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{\ell}}}{\partial \check{y}^{\underline{j}}}\Big|_{\check{Y}_{\mathcal{Q}}} \mathrm{d}\check{\xi}^{\underline{j}}\Big|_{\check{Y}_{\mathcal{Q}}} \right) \\
&\quad + \left(\tau_v|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{k\ell}} \left( \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{k}}}{\partial \check{y}^{\underline{i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \mathrm{d}\check{\xi}^{\underline{i}}\Big|_{\check{Y}_{\mathcal{Q}}} \right) \otimes \left( \frac{\partial^2 \varphi_{\mathcal{Q}}{}^{\underline{\ell}}}{\partial \check{y}^{\underline{p}} \partial \check{y}^{\underline{j}}}\Big|_{\check{Y}_{\mathcal{Q}}} \check{v}^{\underline{p}} \, \mathrm{d}\check{\xi}^{\underline{j}}\Big|_{\check{Y}_{\mathcal{Q}}} + \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{\ell}}}{\partial \check{y}^{\underline{j}}}\Big|_{\check{Y}_{\mathcal{Q}}} \mathrm{d}\check{\xi}^{\underline{\check{Q}+j}}\Big|_{\check{Y}_{\mathcal{Q}}} \right) \\
&= \left(\check{\tau}_q|_{\check{Y}_{\mathcal{Q}}}\right)_{\underline{ij}} \mathrm{d}\check{\xi}^{\underline{\check{Q}+i}}\Big|_{\check{Y}_{\mathcal{Q}}} \otimes \mathrm{d}\check{\xi}^{\underline{j}}\Big|_{\check{Y}_{\mathcal{Q}}} + \left(\check{\tau}_{qv}|_{\check{Y}_{\mathcal{Q}}}\right)_{\underline{ij}} \mathrm{d}\check{\xi}^{\underline{i}}\Big|_{\check{Y}_{\mathcal{Q}}} \otimes \mathrm{d}\check{\xi}^{\underline{j}}\Big|_{\check{Y}_{\mathcal{Q}}} \\
&\quad + \left(\check{\tau}_v|_{\check{Y}_{\mathcal{Q}}}\right)_{\underline{ij}} \mathrm{d}\check{\xi}^{\underline{i}}\Big|_{\check{Y}_{\mathcal{Q}}} \otimes \mathrm{d}\check{\xi}^{\underline{\check{Q}+j}}\Big|_{\check{Y}_{\mathcal{Q}}}
\end{aligned}
$$

with the abbreviations

$$
\begin{aligned}
\left(\check{\tau}_q|_{\check{Y}_{\mathcal{Q}}}\right)_{\underline{ij}} &:= \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{k}}}{\partial \check{y}^{\underline{i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \left(\tau_q|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{k\ell}} \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{\ell}}}{\partial \check{y}^{\underline{j}}}\Big|_{\check{Y}_{\mathcal{Q}}}, \\
\left(\check{\tau}_v|_{\check{Y}_{\mathcal{Q}}}\right)_{\underline{ij}} &:= \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{k}}}{\partial \check{y}^{\underline{i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \left(\tau_v|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{k\ell}} \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{\ell}}}{\partial \check{y}^{\underline{j}}}\Big|_{\check{Y}_{\mathcal{Q}}}, \\
\left(\check{\tau}_{qv}|_{\check{Y}_{\mathcal{Q}}}\right)_{\underline{ij}} &:= \frac{\partial^2 \varphi_{\mathcal{Q}}{}^{\underline{k}}}{\partial \check{y}^{\underline{p}} \partial \check{y}^{\underline{i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \check{v}^{\underline{p}} \left(\tau_q|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{k\ell}} \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{\ell}}}{\partial \check{y}^{\underline{j}}}\Big|_{\check{Y}_{\mathcal{Q}}} + \frac{\partial \varphi_{\mathcal{Q}}{}^{\underline{k}}}{\partial \check{y}^{\underline{i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \left(\tau_v|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{k\ell}} \frac{\partial^2 \varphi_{\mathcal{Q}}{}^{\underline{\ell}}}{\partial \check{y}^{\underline{p}} \partial \check{y}^{\underline{j}}}\Big|_{\check{Y}_{\mathcal{Q}}} \check{v}^{\underline{p}}
\end{aligned}
$$

for $1 \leq k, \ell \leq Q$ and $1 \leq i, j, p \leq \check{Q}$. It is easy to verify that the inverse of $\check{\tau}$ is given by

$$
\begin{aligned}
\check{\tau}|_{\check{Y}_{\mathcal{Q}}}^{-1} &= (\check{\tau}_q|_{\check{Y}_{\mathcal{Q}}})^{\underline{ij}} \frac{\partial}{\partial \check{\xi}^{\underline{i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \otimes \frac{\partial}{\partial \check{\xi}^{\underline{\check{Q}+j}}}\Big|_{\check{Y}_{\mathcal{Q}}} \\
&\quad - (\check{\tau}_v|_{\check{Y}_{\mathcal{Q}}})^{\underline{ik}} (\check{\tau}_{qv}|_{\check{Y}_{\mathcal{Q}}})_{\underline{k\ell}} (\check{\tau}_q|_{\check{Y}_{\mathcal{Q}}})^{\underline{\ell j}} \frac{\partial}{\partial \check{\xi}^{\underline{\check{Q}+i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \otimes \frac{\partial}{\partial \check{\xi}^{\underline{\check{Q}+j}}}\Big|_{\check{Y}_{\mathcal{Q}}} \\
&\quad + (\check{\tau}_v|_{\check{Y}_{\mathcal{Q}}})^{\underline{ij}} \frac{\partial}{\partial \check{\xi}^{\underline{\check{Q}+i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \otimes \frac{\partial}{\partial \check{\xi}^{\underline{j}}}\Big|_{\check{Y}_{\mathcal{Q}}}
\end{aligned} \tag{B.1}
$$

with $1 \leq i, j, k, \ell \leq \check{Q}$, which in bold notation reads

$$
\check{\boldsymbol{\tau}}|_{\check{\boldsymbol{Y}}_{\mathcal{Q}}} = \begin{pmatrix} \check{\boldsymbol{\tau}}_{qv}|_{\check{\boldsymbol{Y}}_{\mathcal{Q}}} & \check{\boldsymbol{\tau}}_v|_{\check{\boldsymbol{Y}}_{\mathcal{Q}}} \\ \check{\boldsymbol{\tau}}_q|_{\check{\boldsymbol{Y}}_{\mathcal{Q}}} & \mathbf{0} \end{pmatrix} \qquad \text{and} \qquad \check{\boldsymbol{\tau}}|_{\check{\boldsymbol{Y}}_{\mathcal{Q}}}^{-1} = \begin{pmatrix} \mathbf{0} & \check{\boldsymbol{\tau}}_q|_{\check{\boldsymbol{Y}}_{\mathcal{Q}}}^{-1} \\ \check{\boldsymbol{\tau}}_v|_{\check{\boldsymbol{Y}}_{\mathcal{Q}}}^{-1} & -\check{\boldsymbol{\tau}}_v|_{\check{\boldsymbol{Y}}_{\mathcal{Q}}}^{-1} \check{\boldsymbol{\tau}}_{qv}|_{\check{\boldsymbol{Y}}_{\mathcal{Q}}} \check{\boldsymbol{\tau}}_q|_{\check{\boldsymbol{Y}}_{\mathcal{Q}}}^{-1} \end{pmatrix}.
$$

Moreover, we obtain for the indices $1 \leq \beta, \gamma \leq 2Q$ and $1 \leq \alpha \leq 2\check{Q}$ and $1 \leq j, k \leq Q$ and $1 \leq \ell, p, r \leq \check{Q}$

$$
\begin{aligned}
\flat_{\check{g}}\left(R_{\mathsf{LMG}}(X_{\mathscr{L}})\right)|_{\check{Y}_{\mathcal{Q}}} &= \left(\mathrm{d}\varphi^*|_{\check{Y}_{\mathcal{Q}}}\left(\flat_g\left(X_{\mathscr{L}}|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)\right)\right)_{\underline{\alpha}} \mathrm{d}\check{\xi}^{\underline{\alpha}}\Big|_{\check{Y}_{\mathcal{Q}}} \\
&= \frac{\partial \varphi^{\underline{\beta}}}{\partial \check{\xi}^{\underline{\alpha}}}\Big|_{\check{Y}_{\mathcal{Q}}} \left(\tau|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{\beta\gamma}} \left(X_{\mathscr{L}}|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)^{\underline{\gamma}} \mathrm{d}\check{\xi}^{\underline{\alpha}}\Big|_{\check{Y}_{\mathcal{Q}}} \\
&= \Bigg(\frac{\partial \varphi^{\underline{j}}}{\partial \check{\xi}^{\underline{\alpha}}}\Big|_{\check{Y}_{\mathcal{Q}}} \left(\tau_v|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{jk}} \left(X_{\mathscr{L}}|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)^{\underline{Q+k}} \\
&\qquad + \frac{\partial \varphi^{\underline{\check{Q}+j}}}{\partial \check{\xi}^{\underline{\alpha}}}\Big|_{\check{Y}_{\mathcal{Q}}} \left(\tau|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{jk}} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}}\Big|_{\check{q}} \check{v}^{\underline{\ell}}\Bigg) \mathrm{d}\check{\xi}^{\underline{\alpha}}\Big|_{\check{Y}_{\mathcal{Q}}} \qquad \text{(B.2)} \\
&= \Bigg(\frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{r}}}\Big|_{\check{q}} \left(\tau_v|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{jk}} \left(X_{\mathscr{L}}|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)^{\underline{Q+k}} \\
&\qquad + \frac{\partial^2 {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{r}} \partial \check{y}^{\underline{p}}}\Big|_{\check{q}} \check{v}^{\underline{p}} \left(\tau_q|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{jk}} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}}\Big|_{\check{q}} \check{v}^{\underline{\ell}}\Bigg) \mathrm{d}\check{\xi}^{\underline{r}}\Big|_{\check{Y}_{\mathcal{Q}}} \\
&\quad + \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{r}}}\Big|_{\check{q}} \left(\tau_q|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{jk}} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}}\Big|_{\check{q}} \check{v}^{\underline{\ell}} \mathrm{d}\check{\xi}^{\underline{\check{Q}+r}}\Big|_{\check{Y}_{\mathcal{Q}}}
\end{aligned}
$$

and observe that the last term equals $\left(\check{\tau}_q|_{\check{Y}_{\mathcal{Q}}}\right)_{\underline{r\ell}}$. Combining (B.2) with (B.1), the $\mathsf{LMG}$ reduction (5.11) of the Euler–Lagrange vector field (4.11) can be written (with the indices $1 \leq \beta, \gamma \leq 2Q$ and $1 \leq \alpha \leq 2\check{Q}$ and $1 \leq j, k \leq Q$ and $1 \leq i, \ell, p, r \leq \check{Q}$) as

$$
\begin{aligned}
R_{\mathsf{LMG}}(X_{\mathscr{L}})|_{\check{Y}_{\mathcal{Q}}} &= \left(\left(\sharp_{\check{g}} \circ \mathrm{d}\varphi^*|_{\check{Y}_{\mathcal{Q}}} \circ \flat_g\right)\left(X_{\mathscr{L}}|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)\right)^{\underline{\alpha}} \frac{\partial}{\partial \check{\xi}^{\underline{\alpha}}}\Big|_{\check{Y}_{\mathcal{Q}}} \\
&= \left(\check{\tau}_q|_{\check{Y}_{\mathcal{Q}}}\right)^{\underline{ir}} \left(\check{\tau}_q|_{\check{Y}_{\mathcal{Q}}}\right)_{\underline{r\ell}} \check{v}^{\underline{\ell}} \frac{\partial}{\partial \check{\xi}^{\underline{i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \\
&\quad + \left(\check{\tau}_v|_{\check{Y}_{\mathcal{Q}}}\right)^{\underline{ir}} \Bigg(\frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{r}}}\Big|_{\check{q}} \left(\tau_v|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{jk}} \left(X_{\mathscr{L}}|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)^{\underline{Q+k}} \\
&\qquad + \frac{\partial^2 {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{r}} \partial \check{y}^{\underline{p}}}\Big|_{\check{q}} \check{v}^{\underline{p}} \left(\tau_q|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{jk}} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}}\Big|_{\check{q}} \check{v}^{\underline{\ell}} \\
&\qquad - \left(\check{\tau}_{qv}|_{\check{Y}_{\mathcal{Q}}}\right)_{\underline{r\ell}} \left(\check{\tau}_q|_{\check{Y}_{\mathcal{Q}}}\right)^{\underline{\ell p}} \left(\check{\tau}_q|_{\check{Y}_{\mathcal{Q}}}\right)_{\underline{ps}} \check{v}^{\underline{s}}\Bigg) \frac{\partial}{\partial \check{\xi}^{\underline{\check{Q}+i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \\
&= \check{v}^{\underline{i}} \frac{\partial}{\partial \check{\xi}^{\underline{i}}}\Big|_{\check{Y}_{\mathcal{Q}}} + \left(\check{\tau}_v|_{\check{Y}_{\mathcal{Q}}}\right)^{\underline{ir}} \Bigg(\frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{r}}}\Big|_{\check{q}} \left(\tau_v|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{jk}} \left(X_{\mathscr{L}}|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)^{\underline{Q+k}} \\
&\qquad + \frac{\partial^2 {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{r}} \partial \check{y}^{\underline{p}}}\Big|_{\check{q}} \check{v}^{\underline{p}} \left(\tau_q|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{jk}} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}}\Big|_{\check{q}} \check{v}^{\underline{\ell}} \\
&\qquad - \frac{\partial^2 {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{p}} \partial \check{y}^{\underline{r}}}\Big|_{\check{q}} \check{v}^{\underline{p}} \left(\tau_q|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{jk}} \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{y}^{\underline{\ell}}}\Big|_{\check{q}} \check{v}^{\underline{\ell}} \\
&\qquad - \frac{\partial {\varphi_{\mathcal{Q}}}^{\underline{j}}}{\partial \check{y}^{\underline{r}}}\Big|_{\check{q}} \left(\tau_v|_{\varphi(\check{Y}_{\mathcal{Q}})}\right)_{\underline{jk}} \frac{\partial^2 {\varphi_{\mathcal{Q}}}^{\underline{k}}}{\partial \check{x}^{\underline{\ell}} \partial \check{y}^{\underline{p}}}\Big|_{\check{q}} \check{v}^{\underline{\ell}} \check{v}^{\underline{p}}\Bigg) \frac{\partial}{\partial \check{\xi}^{\underline{\check{Q}+i}}}\Big|_{\check{Y}_{\mathcal{Q}}} \\
&= \check{X}_{\check{\mathscr{L}}}\Big|_{\check{Y}_{\mathcal{Q}}}
\end{aligned}
$$

Thus, the vector field obtained with the $\mathsf{LMG}$ reduction (5.11) with the $\mathsf{LMG}$ tensor field $\tau_{\mathsf{LMG}}$ from (5.9) results in the reduced Euler–Lagrange vector field (5.8), which is equivalent to solving the reduced Euler–Lagrange equations (5.6) by construction.

[†] Institute of Applied Analysis and Numerical Simulation, University of Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany

*Email address*: {patrick.buchfink,haasdonk}@ians.uni-stuttgart.de

[‡] Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

*Email address*: s.m.glas@utwente.nl

[⋆] Stuttgart Center for Simulation Science (SC SimTech), University of Stuttgart, Universitätsstr. 32, 70569 Stuttgart, Germany

*Email address*: benjamin.unger@simtech.uni-stuttgart.de