Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

Characterization of Exact Lumpability of Smooth Dynamics on Manifolds

by

Leonhard Horstmeyer and Fatihcan M. Atay

Preprint no.: 70 2015

Characterization of Exact Lumpability of Smooth Dynamics on Manifolds

Leonhard Horstmeyer and Fatihcan M. Atay

October 27, 2015

Abstract

We characterize the exact lumpability of nonlinear differential equations on smooth manifolds. We derive necessary and sufficient conditions for lumpability and express them from four different perspectives, thus simplifying and generalizing various results from the literature that exist for Euclidean spaces. The conditions are formulated in terms of the differential of the lumping map, its Lie derivative, and their respective kernels. Two examples are discussed to illustrate the theory.

PACS numbers: 02.30.Hq, 02.30.Ik, 02.30.Yy
AMS classification scheme numbers: 34C40, 34C20, 34A05, 37J15

Keywords: lumping, aggregation, dimensional reduction

1 Introduction

Dimensional reduction is an important aspect in the study of smooth dynamical systems and in particular in modeling by ordinary differential equations (ODEs). Often a reduction can elucidate key mechanisms, reveal conserved quantities, make the problem computationally tractable, or rid it from redundancies. A dimensional reduction by which micro state variables are aggregated into macro state variables also goes by the name of *lumping*. Starting from a micro state dynamics, this aggregation induces a *lumped dynamics* on the macro state space. Whenever a non-trivial lumping, one that is neither the identity nor maps to a single point, confers the defining property to the induced dynamics, one calls the dynamics *exactly lumpable* and the map an *exact lumping*.

Our aim in this paper is to provide necessary and sufficient conditions for exact lumpability of smooth dynamics generated by a system of ODEs on smooth manifolds. To be more precise, let X and Y be two smooth manifolds of dimension n and m, respectively, with $0 < m < n$. Let $\pi_X : TX \to X$

and $\pi_Y : TY \to Y$ be their tangent bundles, whose fibres we take as spaces of derivations, and let v be an element of the smooth sections $\Gamma^\infty(X, TX)$ of TX over X, i.e. smooth maps from X to TX satisfying $\pi_X \circ v = \mathrm{id}_X$, the identity map on X. The integral curves $x(t)$ of v satisfy the equation

$$\frac{d}{dt} x(t) = v(x(t)) \ . \tag{1}$$

On a local coordinate patch $U \subseteq X$ we can write (1) as $\dot{x}^i = v^i(x)$ so that we recover an ODE on that patch. Consider a smooth surjective map $\pi : X \to Y$ and let $y = \pi(x)$. Since $\dim(Y) < \dim(X)$, the mapping π is many-to-one, and hence is called a *lumping*. The question is whether there exists a smooth dynamics on Y that is generated by another system of ODEs, say,

$$\frac{d}{dt} y(t) = \tilde{v}(y(t)) \tag{2}$$

for some smooth vector field $\tilde{v}$ on Y. If that is the case, we say that (1) is *exactly lumpable* for the map π.

In this paper we take the lumping to be a smooth and surjective map $\pi : X \to Y$ and for simplicity restrict ourselves to maps where the preimages are connected. In general one may want to allow for a broader class of lumpings where Y need not have a manifold structure. A natural extension would be to take Y and eventually X to be stratified spaces and π a morphism of stratified spaces. These spaces naturally occur for instance when the lumping is the quotient map of a proper Lie group action. In the following presentation, however, we will constrain ourselves to smooth manifolds.

The reduction of the state space dimension has been studied for Markov chains by Burke and Rosenblatt [1–3] in the 1960s. Kemeny and Snell [4] have studied its variants and called them weak and strong lumpability. Many conditions have been found, mostly in terms of linear algebra, for various forms of Markov lumpability [4–12]. Since Markov chains are characterized by linear transition kernels, most of these conditions carry over directly to the case of linear difference and differential equations. In 1969 Kuo and Wei studied exact [13] and approximate lumpability [14] in the context of monomolecular reaction systems, which are systems of linear first order ODEs of the form $\dot{x} = Ax$. They gave two equivalent conditions for exact lumpability in terms of the commutativity of the lumping map with the flow or with the matrix A respectively. Luckyanov [15] and Iwasa [16] studied exact lumpability in the context of ecological modelling and derived further conditions in terms of the Jacobian of the induced vector field and the pseudoinverse of the lumping map. Iwasa only considered those maps that have a non-degenerate differential, i.e., only submersions. The program was then continued by Li and Rabitz et al. who wrote a series of papers successively generalizing the setting, but remaining in the Euclidean realm. They first constrain the analysis to linear lumping maps [17], where they offer for the first time two construction methods in terms of matrix decompositions of the vector field Jacobian. These methods together with the observability concept [18] from control theory were employed to arrive at a scheme for approximate lumpings with linear maps [19]. They extended their analysis further to exact nonlinear lumpings of general nonlinear but differentiable dynamics [20], providing a set of necessary and sufficient conditions, extending and refining those obtained by Kuo, Wei, Luckyanov and Iwasa. By considering the spaces that are left invariant by the Jacobian of the vector field, they open up a new fruitful perspective, namely the tangent space distribution viewpoint in disguise. Finding lumpings reduces to finding those subspaces. They offer three methods to construct lumpings: Either one finds by an ingenious guess a constraint that is satisfied by some lumping, and this very lumping is then found by an iterative procedure starting from the constraint. Or alternatively, one has to find a set of generalized eigenfunctions to the differential operator given by the vector field of the dynamics when viewed as a derivation. This is as hard a task as finding the set of first integrals. Eventually they also discuss a Lie algebra method, which works in the case where symmetries are present. In each case it is not possible to construct all possible lumpings for a given dynamics. It is also still an open question to determine whether there even exist non-trivial lumpings.

The connection to control theory has been made explicit in [21]. Coxson notices that exact lumpability is an extreme case of non-observability, where the lumping map is viewed as the observable. She specifies another neccesary and sufficient condition by stating that the rank of the observability matrix ought to be that of the lumping map itself. The geometric theory of nonlinear control is outlined in [22]. There, Isidori employs the concept of an exterior differential system [23], though not stated explicitly, in combination with Frobenius Theorem to arrive at a condition of observability for a control system. He makes use of the Lie derivative of the exterior derivative of the lumping, which is analogous to our concept for the Lie derivative of the differential.

In this paper we tie together all these strands into one geometric theory of exact lumpability. The conditions obtained by Iwasa, Luckyanov, Coxson, Li, Rabitz and Toth are contained in this framework. Instead of considering the distribution spanned by the differential of the lumping map, as is done in [20] although not explicitly, we consider the vertical distribution which is defined by the kernel of the differential. We state the mathematical setting in Section 2, where we introduce the relevant objects. In Section 3 we define the notion of exact smooth lumpability and provide four propositions that characterize it. Based on these characterizations, we provide in Section 4 a method for the construction of lumpings and subsequently study two examples, checking for their lumpability in terms of the necessary and sufficient conditions derived in Section 3.

2 Preliminaries

The differential of π at point x is a $\mathbb{R}$-linear map $D\pi_x : T_xX \to T_{\pi(x)}Y$. For $w_x \in T_xX$ the vector $D\pi_x w_x$ can be defined via its action as a derivation $D\pi_x w_x[f] = w_x[f \circ \pi]$, where the argument of the derivation, a smooth and compact test function $f \in C_0^\infty(X, \mathbb{R})$, stands in square brackets. The lumping is a *submersion* whenever $D\pi_x$ is surjective with constant rank for all $x \in X$. The vector bundle over X whose fibers at x are $T_{\pi(x)}Y$ is called the pullback bundle π^*TY and a section is called a vector field along π. As a tangent bundle homomorphism $D\pi : TX \to TY$ is a map whose action on $w \in TX$ is given by $(x, w_x) \mapsto (\pi(x), D\pi_x w_x)$. As a map on smooth sections, $D\pi : \Gamma^\infty(X, TX) \to \Gamma^\infty(X, \pi^*TY)$ is a $\mathcal{C}^\infty$-linear map, taking the vector field w to a vector field along π that lives on X and not on Y. One can only define a vector field $\tilde{w}$ on Y when there is a unique vector $D\pi_x w(x)$ for all $x \in \pi^{-1}(y)$ and all $y \in Y$. Whenever this holds true, $\tilde{w}$ and w are called *π-related*. It is easy to check [24] that w and $\tilde{w}$ are π-related if and only if $w[f \circ \pi] = \tilde{w}[f] \circ \pi$ for test functions f. If $\theta : X \to X'$ is a diffeomorphism, where X' is some manifold of equal dimension, then all sections on X are θ-related to a unique section on X', since there is a one-to-one correspondence. We use the notation $\theta_* : \Gamma^\infty(X, TX) \to \Gamma^\infty(X', TX')$ to denote the *pushforward of vector fields* from X to θ-related vector fields on X'.

Let w be a vector field on X with local flow θ_t. The Lie derivative $\mathcal{L}_w v$ of v along w is defined by

$$\mathcal{L}_w v := \frac{d}{dt}\Big|_{t=0} (\theta_{-t})_*(v \circ \theta_t) \ , \tag{3}$$

which is again a smooth vector field on X and can be shown to be equivalent to the commutator $\big[w, v\big]$. The Lie derivative of a real-valued function g along w is $\mathcal{L}_w g = w[g]$.

Given a linear map $L : T_xX \to V$ into a vector space V and a diffeomorphism $\theta : X \to X'$, there exists an induced linear map $\theta_\sharp L : T_{x'}X' \to V$ of L:

$$\big(\theta_\sharp L\big)_{x'} := L \circ \big(D\theta^{-1}\big)_{x'} \ . \tag{4}$$

It is clearly again linear since the differential $(D\theta^{-1})_{x'}$ is linear and linearity is preserved under composition. Analogously to the Lie derivative of sections on the tangent bundle, we define the Lie derivative of the differential $D\pi$, a section on the vector bundle $\mathrm{Hom}(TX, \pi^*TY)$ over X, given here pointwise as

$$(\mathcal{L}_w D\pi)_x := \frac{d}{dt}\Big|_{t=0} \Big(\big(\theta_{-t}\big)_\sharp D\pi_{\theta_t(x)}\Big)_x \ . \tag{5}$$

However, (5) is a $\mathcal{C}^\infty$-linear map from TX to π^*TY. In order to write down the component form of (5), we choose a coordinate chart $\psi : U \subseteq X \to \mathbb{R}^n$

with indices labelled by i, j and a chart $\tilde{\psi} : V \subseteq Y \to \mathbb{R}^m$ whose indices are labelled by a, b. We invoke the definition of $\mathcal{L}_w D\pi$ to obtain

$$\begin{aligned}
(\mathcal{L}_w D\pi\,)^a_i &= \lim_{\epsilon\to 0} \frac{1}{\epsilon}\Big[\sum_j \Big(\frac{\partial \pi^a}{\partial x^j}\circ \theta_\epsilon\Big)\frac{\partial \theta^j_\epsilon}{\partial x^i} - \frac{\partial \pi^a}{\partial x^j}\delta^j_i\Big] \\
&= \lim_{\epsilon\to 0} \frac{1}{\epsilon}\Big[\Big(\frac{\partial \pi^a}{\partial x^j}\circ(\mathrm{id} + \epsilon w + \mathcal{O}(\epsilon^2))\Big)\frac{\partial}{\partial x^i}\Big(\mathrm{id}^j + \epsilon w^j + \mathcal{O}(\epsilon^2)\Big) - \frac{\partial \pi^a}{\partial x^i}\Big] \\
&= \lim_{\epsilon\to 0} \frac{1}{\epsilon}\Big[\epsilon \sum_j \Big(w^j \frac{\partial^2 \pi^a}{\partial x^i \partial x^j} + \frac{\partial \pi^a}{\partial x^j}\frac{\partial v^j}{\partial x^i}\Big) + \mathcal{O}(\epsilon^2)\Big] \\
&= \sum_j \frac{\partial}{\partial x^i}\Big(w^j \frac{\partial \pi^a}{\partial x^j}\Big).
\end{aligned}$$

We can use the action of the Lie derivative on differentials and on vector fields to also define its action on sections of the pullback bundle $\pi^* TY$, such that Leibniz's rule holds:

$$\mathcal{L}_w(D\pi\, v) := (\mathcal{L}_w D\pi\,)v + D\pi\, \mathcal{L}_w v\ . \tag{6}$$

We then get the following identity, spelled out in local coordinates,

$$\big(\mathcal{L}_w(D\pi\, v)\big)^a = \sum_{jk} w^j \frac{\partial}{\partial x^j}\Big(\frac{\partial \pi^a}{\partial x^k} v^k\Big) = \big((\mathcal{L}_v D\pi)\, w\big)^a.$$

Next we introduce the central object of this paper. Following [25], a (singular tangent space) distribution S is a choice of a subspace $S_x \subseteq T_x X$ of the tangent space at each point $x \in X$, so $S = \bigsqcup_{x\in X} S_x$, where $\bigsqcup$ denotes the disjoint union. A priori this choice need not be continuous in x or be of constant dimension. S is said to be a smooth distribution if each subspace S_x is given by the span of locally defined smooth vector fields at x. Smoothness of S does not entail that the distribution is regular, i.e. that it has constant dimension.

The distribution $\ker D\pi\; = \bigsqcup_{x\in X} \ker D\pi_x$ can be shown to be smooth. This follows from the existence of a smooth local coframe (c.f. [24]) spanned by m smooth 1-forms $(d\pi^1, \ldots, d\pi^m)$ that annihilate $\ker D\pi$. The distribution $\ker D\pi$ has constant dimension m if and only if π is a submersion. The distribution $\ker \mathcal{L}_v D\pi\; = \bigsqcup_{x\in X} \ker \mathcal{L}_v D\pi_x$ can be shown to be smooth in much the same way. This time the annihilating m smooth 1-form fields are $(d\sigma^1, \ldots, d\sigma^m)$ where $\sigma^a = \langle d\pi^a, v\rangle$ is a smooth function on X and $\langle\cdot,\cdot\rangle : T^*X \times TX \to \mathbb{R}$ is the natural pairing of tangent and co-tangent vectors.

3 Characterization of Lumpability

In this section, through a sequence of propositions, we shall derive conditions for exact lumpability from four perspectives. We start by giving a precise definition of lumpability.

Definition 1 (Exact Smooth Lumpability)**.** *The system*

$$\dot{x} = v(x) \tag{7}$$

is called exactly smoothly lumpable (henceforth exactly lumpable) for π iff there exists a smooth vector field $\tilde{v} \in \Gamma^\infty(Y, TY)$ such that the dynamics of $y = \pi(x)$ is governed by

$$\dot{y} = \tilde{v}(y), \tag{8}$$

where x satisfies (7).

The Picard-Lindelöf Theorem guarantees a unique solution of (7) for sufficiently small times, but it may cease to exist at some point. Let $\mathcal{T}_x \subseteq \mathbb{R}$ be the times for which a solution with initial point x exists. We introduce $\mathcal{T}_X := \{(\mathcal{T}_x, x) : x \in X\}$ and define the flow $\Phi : \mathcal{T}_X \subseteq \mathbb{R} \times X \to X$ by the map $(t, x(0)) \mapsto x(t)$. Given the flow Φ, we denote by $\Phi_x : \mathcal{T}_x \to X$ the integral curves with starting point x, and by $\Phi_t : \mathcal{X}_t \to X$ the flow map parametrized by time, with $\mathcal{X}_t := \{x \in X : t \in \mathcal{T}_x\}$ being the domain of definition.

Formally, equation (7) should be understood as the pushforward of the unit section $\frac{\partial}{\partial t}$ on $\mathcal{T}_x$ by the integral curve $\Phi_x : \mathcal{T}_x \to X$:

$$\frac{d}{dt}\Big|_t \Phi_x := \left(D\Phi_x\right)_t \frac{\partial}{\partial t} = v(\Phi_x(t)), \tag{9}$$

and likewise for (8). Given v and $\tilde{v}$, both curves Φ_x and $\tilde{\Phi}_{\pi(x)}$ are guaranteed to exist at least for small times $\mathcal{T}_x$ and $\widetilde{\mathcal{T}}_y$ respectively. There is no apriori connection between those times; however, we will see later that Proposition 2 relates them.

Proposition 1. *The system $\dot{x} = v(x)$ is exactly lumpable for π iff there exists a smooth vector field $\tilde{v} \in \Gamma^\infty(Y, TY)$ such that*

$$D\pi_x v(x) = \tilde{v}(\pi(x)) \tag{10}$$

for all $x \in X$.

Proof. Consider the time derivatives of $\pi \circ \Phi_x$ and $\tilde{\Phi}_{\pi(x)}$:

$$\frac{d}{dt}\Big|_t \pi \circ \Phi_x = D(\pi \circ \Phi_x)_t \frac{\partial}{\partial t} = D\pi_{\Phi_x(t)}(D\Phi_x)_t \frac{\partial}{\partial t} = D\pi_{\Phi_x(t)} v(\Phi_x(t)), \tag{11}$$

$$\frac{d}{dt}\Big|_t \tilde{\Phi}_{\pi(x)} = \left(D\tilde{\Phi}_{\pi(x)}\right)_t \frac{\partial}{\partial t} = \tilde{v}(\tilde{\Phi}_{\pi(x)}(t)), \tag{12}$$

and take $t = 0$, where the equality $\pi(x) = \pi \circ \Phi_x(0) = \tilde{\Phi}_{\pi(x)}(0)$ holds. By the definition of exact lumpability we know that $\pi(x)$ is governed by $\tilde{v}$; in other words, $\tilde{v}(\pi(x)) \stackrel{!}{=} \frac{d}{dt}|_0\ \pi \circ \Phi_x = D\pi_x v(x)$, where the latter equality comes from (11). Conversely if we assume condition (10) for all x, then (11) equals (12) for $t = 0$. Thus the infinitesimal dynamics of $\pi \circ \Phi_x$ is governed by $\tilde{v}(\pi(x))$, which is the definition of exact lumpability. □

Remark 1. *Alternatively, we can say that $\dot{x} = v(x)$ is exactly lumpable for π iff there exists a smooth vector field $\tilde{v} \in \Gamma^\infty(Y, TY)$ such that $\tilde{v}$ and v are π-related, i.e. $v[f \circ \pi] = \tilde{v}[f] \circ \pi$ for any smooth and compact test function f.*

Proposition 2. *The system $\dot{x} = v(x)$ is exactly lumpable for π iff for all $y \in Y$ the time domain $\widetilde{\mathcal{T}}_y = \mathcal{T}_x$ is independent of the choice $x \in \pi^{-1}(y)$, and there is a smooth map $\tilde{\Phi} : \widetilde{\mathcal{T}}_Y \to Y$ such that*

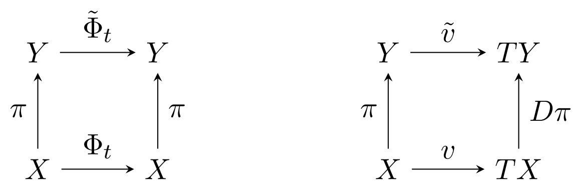
$$\tilde{\Phi}_t \circ \pi(x) = \pi \circ \Phi_t(x) \tag{13}$$

for all $x \in X$ and all times $t \in \widetilde{\mathcal{T}}_{\pi(x)}$.

Proof. One implication is obtained by taking time derivatives on both sides of (13) at $t = 0$, which gives rise to (10) and by Proposition 1 implies exact lumpability.

For the other direction we consider the curve $\tilde{\Theta}_x = \pi \circ \Phi_x : \mathcal{T}_x \to Y$ with $\tilde{\Theta}_x(0) = \pi(x) = y$. From the condition (10) for exact lumpability we see that $\tilde{v}(\pi \circ \Phi_x(t)) = D\pi_{\,\Phi_x(t)} v(\Phi_x(t)) = \frac{d}{dt}\big(\pi \circ \Phi_x\big)(t)$, so $\tilde{v}$ is tangent to $\tilde{\Theta}_x(t)$ for all times $t \in \mathcal{T}_x$. Thus $\tilde{\Theta}_x$ is an integral curve of the vector field $\tilde{v}$ going through y. For $\tilde{v}$ there exists already an integral curve $\tilde{\Phi}_y$, which by uniqueness must coincide with $\tilde{\Theta}_x$, and so we have $\tilde{\Phi}_{\pi(x)}(t) = \pi \circ \Phi_x(t)$ for all $t \in \mathcal{T}_x$. In terms of flow maps this is equivalent to (13). This argument is independent of the choice of $x \in \pi^{-1}(y)$; so the domain of definition has to be $\widetilde{\mathcal{T}}_y = \mathcal{T}_x$ for all $x \in \pi^{-1}(y)$. □

Remark 2. *The* $\operatorname{rank} D\pi : X \to \mathbb{N}$ *is to be understood as an allocation of the* $\operatorname{rank} D\pi_x \in \mathbb{N}$ *for every point $x \in X$. So, if π is a submersion then* $\operatorname{rank} D\pi \equiv m$*, but otherwise not. Propositions 1 and 2 can be cast into commuting diagrams:*



The left one says that $\tilde{\Phi}_t \circ \pi = \pi \circ \Phi_t$ for all times of definition t. The right one reads $\tilde{v}(\pi(x)) = D\pi_x v(x)$ for all $x \in X$.

Proposition 3. *The system $\dot{x} = v(x)$ is exactly lumpable for π iff* $\ker D\pi$ *is invariant under $\mathcal{L}_v$, or equivalently,* $\ker D\pi \subseteq \ker \mathcal{L}_v D\pi$*.*

Proof. First we show that $\ker D\pi$ is invariant under $\mathcal{L}_v$ iff $\ker D\pi \subseteq \ker \mathcal{L}_v D\pi$. Recall from the Leibniz's rule (6) that

$$\big(\mathcal{L}_v D\pi\big)w = \mathcal{L}_v\big(D\pi\, w\big) - D\pi\, \mathcal{L}_v w\ . \tag{14}$$

Take $w \in \ker D\pi$. If $\ker D\pi$ is invariant under $\mathcal{L}_v$, then the right hand side of (14) vanishes and $w \in \ker \mathcal{L}_v D\pi$. Conversely, if $w \in \ker \mathcal{L}_v D\pi$, then by (14) the Lie derivative $\mathcal{L}_v w$ is again in $\ker D\pi$, since $w \in \ker D\pi$.

Second, we show that exact lumpability implies the invariance of $\ker D\pi$ under $\mathcal{L}_v$. By exact lumpability we know from (10) that there is a vector field $\tilde{v}$ such that $v[f\circ\pi] = \tilde{v}[f]\circ\pi$ for any test function $f \in \mathcal{C}^\infty(Y,\mathbb{R})$. Therefore,

$$\begin{aligned} \big(D\pi\,\mathcal{L}_v w\big)[f] = D\pi\,\big[v,w\big][f] &= v[w[f\circ\pi]] - w[v[f\circ\pi]] \qquad (15)\\ &= v[w[f\circ\pi]] - w[\tilde{v}[f]\circ\pi] \\ \Leftrightarrow \qquad \big(D\pi\,\mathcal{L}_v w\big)[f] &= v[D\pi\, w[f]] - D\pi\, w[\tilde{v}[f]] \qquad (16) \end{aligned}$$

and thus $w \in \ker D\pi$ implies $\mathcal{L}_v w \in \ker D\pi$.

Third, we show that the condition $\ker D\pi \subseteq \ker \mathcal{L}_v D\pi$ implies exact lumpability. We want to define the vector field $\tilde{v}$ as a smooth function of y such that $\tilde{v}_{\pi(x)} = D\pi_x v(x)$ for all $x \in X$. This would imply exact lumpability due to (10). We first note that $\tilde{v}$ is well defined because π is surjective and smooth and $D\pi_x v(x)$ is constant along the connected fibers $\pi^{-1}(y)$. The latter can be seen by considering a vector field $w \in \ker D\pi$ and its local flow Θ. We fix local coordinate patches $\tilde{\psi} : V \subseteq Y \to \mathbb{R}^m$ and $\psi : U \subset X \to \mathbb{R}^n$, with $\tilde{\psi}^a(y) = y^a$, $\tilde{\psi}^a \circ \pi = \pi^a$, and $\psi^i(x) = x^i$. Now $D\pi\, v$ does not change along the flow:

$$\frac{\partial}{\partial t}\left(D\pi^{\,a} v\circ\Theta_x\right)_{t=0} = \sum_{ij}\left(\frac{\partial}{\partial x^i}\left(\frac{\partial \pi^a}{\partial x^j}v^j\right)\right)\left(\frac{\partial \Theta_x^i}{\partial t}\right)_{t=0} = \sum_i (\mathcal{L}_v D\pi\,)^a_i w^i = 0\ ,$$

since $w \in \ker D\pi$ implies $w \in \ker \mathcal{L}_v D\pi$ by assumption.

It remains to show that $\tilde{v}$ is a smooth function of y. This is the case if for any smooth curve $\tilde{\gamma}_y : (-\epsilon,\epsilon) \to Y$ the composition $\tilde{v}\circ\tilde{\gamma}_y$ is a smooth function in time. But any such curve can be viewed as the composition of π with a curve $\gamma_x : (-\epsilon,\epsilon) \to X$, where $\pi(x) = y$. Since for any γ_x the equality $\tilde{v}\circ\pi\circ\gamma_x = D\pi\, v\circ\gamma_x$ holds, and since the right hand side is a composition of smooth functions and is thus also smooth, it follows that $\tilde{v}$ must be smooth. □

Proposition 4. *A distribution Ω is invariant under $\mathcal{L}_v$ (i.e. $w \in \Omega \Rightarrow \mathcal{L}_v w \in \Omega$) iff it is invariant under the corresponding flow ϕ_t (i.e. $D\phi_t\Omega_x = \Omega_{\phi_t(x)}$ for all $x \in X$).*

Proof. This can be seen by considering $w \in \Omega$, $\sigma \in \Omega^\perp$ and the pairing $\langle \sigma(x), D\phi_{-t} w(\phi_t(x)) \rangle$. For $t = 0$ we have $\langle \sigma(x), w(x) \rangle = 0$ by definition. Upon taking time derivatives we get

$$\frac{d}{dt}\Big|_0 \Big\langle \sigma(x), D\phi_{-t} w(\phi_t(x)) \Big\rangle = \Big\langle \sigma(x), \frac{d}{dt}\Big|_0 D\phi_{-t} w(\phi_t(x)) \Big\rangle = \Big\langle \sigma(x), \mathcal{L}_v w(x) \Big\rangle .$$

Hence $w \in \Omega \Rightarrow \mathcal{L}_v w \in \Omega$ implies and is implied by $D\phi_{-t} w(\phi_t(x)) \in \Omega_x$ for all t and x. Upon multiplying by $D\phi_t$ the latter becomes $w(\phi_t(x)) \in D\phi_t \Omega_x$. The above argument can be repeated for $-v$, with $-\mathcal{L}_v w(x) = \mathcal{L}_{-v} w(x) = \frac{d}{dt}\Big|_0 D\phi_t w(\phi_{-t}(x))$, implying this time $D\phi_t w(\phi_{-t}(x)) \in \Omega_x$. By taking $x \to \phi_t(x)$ this becomes $D\phi_t w(x) \in \Omega_{\phi_t(x)}$. In summary, $w \in \Omega \Rightarrow \mathcal{L}_v w \in \Omega$ is equivalent to $D\phi_t \Omega_x = \Omega_{\phi_t(x)}$ for all x and t in the domain of definition. □

We make the connection to control theory by introducing the 2-*observability map*:

$$\mathcal{O}_2 := \begin{pmatrix} D\pi \\ \mathcal{L}_v D\pi \end{pmatrix} : TX \to \pi^* TY \oplus \pi^* TY \ , \tag{17}$$

given locally by

$$(x; w) \mapsto \big(\pi(x); (D\pi_x w, \mathcal{L}_v D\pi_x w)\big) \ .$$

The n-observability map $\mathcal{O}_n : TX \to \bigoplus^n \pi^* TY$ is defined analogously with higher-order Lie derivatives. In the linear case, where $\dot{x} = v(x) = Ax$ and $y = \pi(x) = Cx$, we have $\mathcal{O}_2 = \begin{pmatrix} C \\ CA \end{pmatrix}$; furthermore, $\mathcal{O}_n$ is just the standard observability matrix

$$\mathcal{O}_n = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} \tag{18}$$

familiar from linear control theory [26], where the system is called *observable* if $\operatorname{rank} \mathcal{O}_n = n$. We will show that a system is exactly lumpable iff $\operatorname{rank} \mathcal{O}_2 = \operatorname{rank} D\pi$.

Proposition 5. *The system $\dot{x} = v(x)$ is exactly lumpable for π iff* $\operatorname{rank} \mathcal{O}_2 = \operatorname{rank} D\pi$ *, or equivalently, iff locally on each coordinate patch,* $\big(\mathcal{L}_v D\pi\big)^a \in \operatorname{span}(D\pi^1, \ldots, D\pi^m)$ *for all a with $1 \leq a \leq m$.*

Proof. First, we fix coordinate patches $\tilde{\psi} : V \subseteq Y \to \mathbb{R}^m$ and $\psi : U \subset X \to \mathbb{R}^n$ with $\tilde{\psi}^a(y) = y^a$ and $\psi^i(x) = x^i$. Note that in local coordinates the

2-observability map is given by the $2m \times n$ matrix

$$\mathcal{O}_2 = \begin{pmatrix} D\pi^1 \\ \vdots \\ D\pi^m \\ \mathcal{L}_v D\pi^1 \\ \vdots \\ \mathcal{L}_v D\pi^m \end{pmatrix} = \begin{pmatrix} \frac{\partial \pi^1}{\partial x^1} & \cdots & \frac{\partial \pi^1}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \pi^m}{\partial x^1} & \cdots & \frac{\partial \pi^1}{\partial x^n} \\ \frac{\partial}{\partial x^1}\left(\sum_j v^j \frac{\partial \pi^1}{\partial x^j}\right) & \cdots & \frac{\partial}{\partial x^n}\left(\sum_j v^j \frac{\partial \pi^1}{\partial x^j}\right) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x^1}\left(\sum_j v^j \frac{\partial \pi^m}{\partial x^j}\right) & \cdots & \frac{\partial}{\partial x^n}\left(\sum_j v^j \frac{\partial \pi^m}{\partial x^j}\right) \end{pmatrix}. \quad (19)$$

The first m rows of $\mathcal{O}_2$ have rank $D\pi$ by construction. So, every further row $\mathcal{L}_v D\pi^a$ needs to be a $\mathcal{C}^\infty$-linear combination of rows $D\pi^b$, and hence lies in $\mathrm{span}_{\mathcal{C}^\infty}(D\pi^1, \ldots, D\pi^m)$. Likewise, if $\mathcal{L}_v D\pi^a$ lie in $\mathrm{span}_{\mathcal{C}^\infty}(D\pi^1, \ldots, D\pi^m)$, then $\mathrm{rank}\, \mathcal{O}_2 = \mathrm{rank}\, D\pi$.

It remains to show the equivalence of the span-condition and exact lumpability. The easy direction follows by writing the span-condition

$$\mathcal{L}_v D\pi^a = \sum_b \phi^a_b D\pi^b \quad (20)$$

with smooth coefficient functions ϕ^a_b. Now taking $w \in \ker D\pi$ implies $w \in \ker \mathcal{L}_v D\pi$ and, by Proposition 3, yields exact lumpability. On the other hand, if the system is exactly lumpable, then by Proposition 1 we have $D\pi^a v(x) = \tilde{v}^a(\pi(x))$ for all x. Taking partial derivatives in the x^i-direction on both sides gives

$$\sum_b \frac{\partial \tilde{v}^a}{\partial y^b} \frac{\partial \pi^b}{\partial x^i} = \frac{\partial}{\partial x^i}(\tilde{v}^a \circ \pi) = \frac{\partial}{\partial x^i}(D\pi\, v)^a = \sum_j \frac{\partial}{\partial x^i}\Big(v^j \frac{\partial \pi^a}{\partial x^j}\Big). \quad (21)$$

The rightmost expression is the local coordinate form of $(\mathcal{L}_v D\pi)^a_i$ and the leftmost side can be read as the linear combination of $D\pi^b$ with coefficients $\frac{\partial \tilde{v}^a}{\partial y^b}$, which are smooth because $\tilde{v}$ is smooth. In this way, the span-condition (20) follows. □

Remark 3. *If v and π are linear and X and Y are both Euclidean spaces, then the conditions reduce to the ones known for linear ODEs [13, 27].*

In closing this section, we discuss the relation of lumpability to the symmetries of the system. We shall show that the action of a Lie group that leaves the vector field invariant results in an exact lumping; however, the converse is not true.

Let G be a finite Lie Group with Lie Algebra $\mathfrak{g}$. We denote by $\mathcal{A} : G \times X \to X$ a smooth action (left or right action) on X and by $a : \mathfrak{g} \to \Gamma^\infty(X, TX)$ the corresponding action of the Lie Algebra into the smooth vector fields on X. The action on the whole algebra is a smooth distribution $a(\mathfrak{g})$, because it is by definition spanned by smooth vector fields.

Proposition 6. *Let $\mathcal{A}$ be a proper and free G-action on X and suppose v satisfies the condition*

$$\mathcal{L}_v a(\mathfrak{g}) \subseteq a(\mathfrak{g}). \tag{22}$$

Then the system $\dot{x} = v(x)$ is exactly lumpable for the quotient map $\pi : X \to X/G$.

Proof. Given a proper G-action, X can be decomposed into smooth submanifolds, called orbit types, that share the same stabiliser group up to conjugacy. The orbit space X/G inherits a decomposition into the quotients of orbit types, which are again smooth submanifolds and form the strata of a Whitney-stratification [28]. If the group also acts freely, then all orbits are of the same orbit type and the quotient has a natural smooth manifold structure. The quotient map is a submersion [24] onto the orbit space. the tangent space to any orbit is spanned by $a(\mathfrak{g})$ and mapped to zero by $D\pi$. Condition (22) then implies that $\ker D\pi$ is invariant under $\mathcal{L}_v$, which, by Proposition 3, implies exact lumpability. □

Remark 4. *Requiring that the vector field be invariant under the symmetry, namely that $\mathcal{L}_v a(\mathfrak{g}) = 0$, is a special case of, and thus stronger than, the condition (22).*

Remark 5. *The converse statement to Proposition 6 is not true. Given a vector field v and a lumping π, the level sets need not be orbits of a Lie group action. This can already be seen for very simple vector fields: For instance, if one considers the zero vector field, then π can be chosen arbitrarily, and thus in such a way that level sets cannot ever come from the action of the same group.*

4 Construction of Lumping and Examples

4.1 Construction of lumping maps

We now briefly consider the problem of finding a lumping map π under which the system is exactly lumpable.

By virtue of Proposition 5, $\mathcal{L}_v D\pi^a$ has to be a $\mathcal{C}^\infty$-combination of the $(D\pi^b)_{b=1}^m$ on each coordinate patch. To formalize the condition of linear dependence, consider the vector bundle whose fibers consist of the exterior algebra $\Lambda(T_x^* X)$. The algebra should be over the linear maps $T_x X \to T_{\pi^b(x)}\mathbb{R} \cong \mathbb{R}$, which is isomorphic to $T_x^* X$. So, the condition for exact lumpability becomes

$$\Omega^a := D\pi^1 \wedge \cdots \wedge D\pi^m \wedge \mathcal{L}_v D\pi^a = 0, \quad \forall a\ , \tag{23}$$

which is a system of second order PDEs for π. Solving (23) analytically is in general not easy but may be possible in specific cases; see Section 4.2

for an example. Alternatively, (23) may be approximated or numerically investigated, which will not be pursued in the present paper.

For maps $\pi : \mathbb{R}^n \to \mathbb{R}$, the condition (23) reads

$$0 = \Omega_{ij} = \sum_k \frac{\partial \pi}{\partial x^i} \frac{\partial}{\partial x^j} \left(\frac{\partial \pi}{\partial x^k} v^k \right) - \sum_k \frac{\partial \pi}{\partial x^j} \frac{\partial}{\partial x^i} \left(\frac{\partial \pi}{\partial x^k} v^k \right), \qquad \forall i, j \ . \quad (24)$$

One can also phrase (24) as a variational problem, which has the additional advantage that constraints can be added via the Lagrange multiplier formalism. We first introduce the variables $\gamma_i = \frac{\partial \pi}{\partial x^i}$ and their derivatives $\gamma_{ij} = \frac{\partial \gamma_i}{\partial x^j} = \frac{\partial^2 \pi}{\partial x^j \partial x^i}$. The Lagrangian $\mathcal{L} = \sum_{ij} \Omega_{ij}^2$ is a non-negative function of γ_i and γ_{ij}. Depending on v it will also have an explicit x-dependence. Hence, the variational problem consists of finding γ such that the integral

$$I = \int_{U \subseteq X} \mathcal{L} \, dx$$

is minimal over some region $U \subseteq X$.

In the remainder of the paper, we discuss two examples and study them in terms of the neccessary and sufficient conditions derived in Propositions 1-3 and 5.

4.2 Lotka-Volterra type dynamics

Consider the set of scalar ODEs defined on $\mathbb{R}^n$ by

$$\dot{x}_i = x_i \Big(1 - \sum_{j=1}^n a_j x_j \Big) = v_i(x), \qquad i = 1, \dots, n, \quad (25)$$

for some set of real coefficients a_i which are not all zero. The system (25) can exhibit non-trivial behaviour, such as limit cycles, yet we will see that the aggregated description in terms of a weighted average of the system variables satisfies a simple first order ODE.

We first demonstrate that our conditions apply. Plugging (25) into condition (24) gives

$$0 \stackrel{!}{=} \Omega_{ij} = (\gamma_i a_j - \gamma_j a_i) \sum_k \gamma_k x_k + \sum_k v_k (\gamma_i \gamma_{kj} - \gamma_j \gamma_{ki}) \, . \quad (26)$$

One sees immediately that $\gamma_i = \frac{\partial \pi}{\partial x^i} = a_i$ is a solution. Integrating, we find the lumpings $\pi_c : \mathbb{R}^n \to \mathbb{R}$ given by $\pi_c(x) = \sum_j a_j x_j + c$. Exact lumpability can then be checked for $\pi = \pi_0$ using the characterizations presented in this paper, as follows.

First, we can find a vector field $\tilde{v}$, namely $\tilde{v}(y) = y(1 - y)$, such that

$$D\pi_x v(x) = \sum_i \frac{\partial \pi}{\partial x_i} v_i(x) = \sum_i a_i x_i \Big(1 - \sum_j a_j x_j \Big) = \tilde{v}(\pi(x))$$

for all $x \in \mathbb{R}^n$, as required by Proposition 1.

Second, we consider $\ker D\pi = \left\{\xi : \sum_i a_i \xi_i = 0\right\}$. In explicit terms we can pick an $a_j \neq 0$ so that

$$\ker D\pi = \operatorname{span}_{\mathbb{R}} \left\{ \begin{pmatrix} a_j \\ 0 \\ \vdots \\ a_1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ a_j \\ \vdots \\ a_2 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ a_j \\ a_{j-1} \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a_{j+1} \\ a_j \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a_n \\ \vdots \\ 0 \\ a_j \end{pmatrix} \right\} .$$

This is a fixed $n-1$ dimensional linear subspace independent of x. We take $\xi \in \ker D\pi$ and compute its Lie derivative as

$$\left(\mathcal{L}_v \xi\right)_i = \sum_j \left(v_j \frac{\partial \xi_i}{\partial x^j} - \xi_j \frac{\partial v_i}{\partial x^j} \right) = -\Big(1 - \sum_j a_j x_j\Big) \xi_i \,,$$

where we have used the condition $\sum_j a_j \xi_j = 0$. This shows that the Lie derivative of a vector in the kernel is proportional to itself and thus again in the kernel, as required by Proposition 3.

Finally, we compute $\mathcal{L}_v D\pi$ as

$$(\mathcal{L}_v D\pi)_i = \Big(1 - 2\sum_j a_j x_j\Big) a_i = \Big(1 - 2\sum_j a_j x_j\Big) D\pi_i \,, \tag{27}$$

which is proportional to $D\pi$ and thus belongs to its span. Furthermore, it has the same kernel as $D\pi$ unless $\sum_j a_j x_j = \frac{1}{2}$, when its null space is all of $\mathbb{R}^n$. So, in both cases $\ker D\pi \subseteq \ker \mathcal{L}_v D\pi$, as required by Proposition 5. Hence, the system (25) is exactly lumpable for the map π given by $\pi(x) = \sum_j a_j x_j$. The geometry of the lumping is illustrated in Figure 1.

4.3 Two-dimensional system with cubic nonlinearity

The next example we consider is a two-dimensional nonlinear equation given by

$$\begin{array}{rclcl} \dot{x}_1 & = & 4x_1 x_2^2 + x_1^3 & = & v_1(x) \\ \dot{x}_2 & = & -2x_1^2 x_2 + x_2^3 & = & v_2(x) \end{array} . \tag{28}$$

Consider the nonlinear lumping $\pi : \mathbb{R}^2 \to \mathbb{R}^{\geq 0}$ given by

$$\pi(x) = x_1^2 + x_2^2 \,,$$

which maps a point x to its squared distance from the origin. That π is an exact lumping can also be checked by the condition (24). Note that $\mathbb{R}^{\geq 0}$ is a manifold with boundary and thus admits a stratification, but does

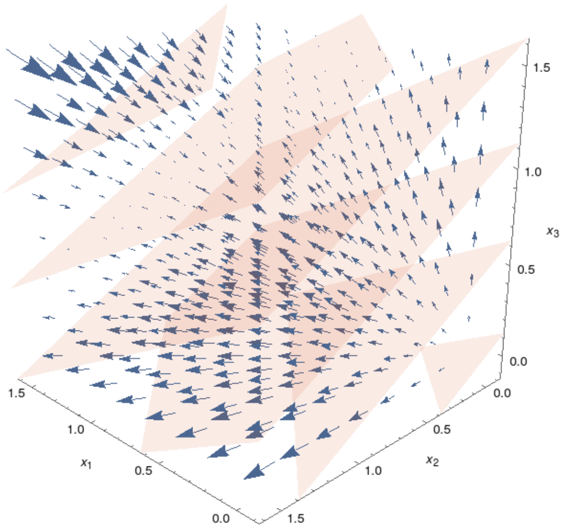


Figure 1: The vector field for the Lotka-Volterra system (25) for three variables. The colored planes are the level sets of $\pi(x) = \sum_j x_j$ for values $\pi = -1, -\frac{1}{2}, 0, \frac{1}{2}, 1$ and $\frac{3}{2}$.

not technically have a smooth manifold structure. However, by subtracting the boundary stratum $\{0\}$ we are back to the manifold scenario, and π is a submersion. Exact lumpability can be checked as before: We can find a vector field $\tilde{v}$, namely $\tilde{v}(y) = 2y^2$, such that $D\pi_x v(x) = 2(x_1^2 + x_2^2)^2 = \tilde{v}(\pi(x))$. The null space, which now depends on the point x, is given by

$$\ker D\pi \; = \text{span}_{\mathbb{R}} \left\{ \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix} \right\}$$

and is invariant under the Lie derivative

$$\mathcal{L}_v \xi = 3(x_1^2 - x_2^2)\xi \; ,$$

where ξ is in $\ker D\pi$. In particular, for $x_1 = \pm x_2$ we have $\mathcal{L}_v \xi = 0$. The Lie derivative of the differential

$$(\mathcal{L}_v D\pi)_i = 4(x_1^2 + x_2^2) D\pi_i$$

is proportional to itself and therefore is in the span of $D\pi$. Anything that is annihilated by $D\pi$ is also annihilated by $\mathcal{L}_v D\pi$. Hence, the system (28) is exactly lumpable for π. The geometry of the lumping is illustrated in Figure 2.

Acknowledgement. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 318723 (MATHEMACS).

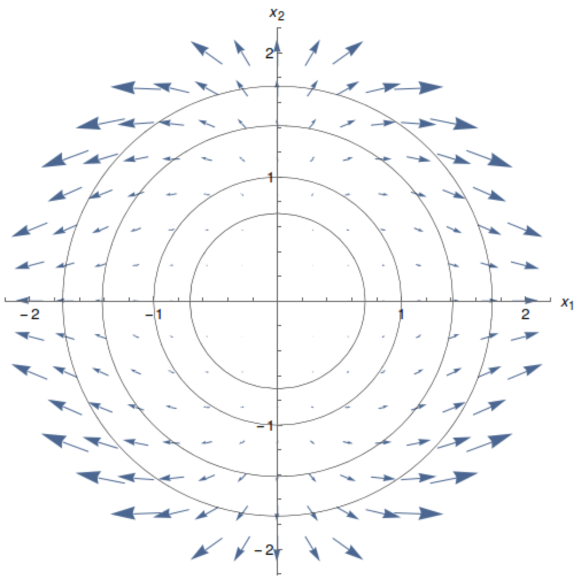


Figure 2: The vector field for the cubic system (28). The circles are the contours of $\pi(x) = x_1^2 + x_2^2$ for values $\pi = \frac{1}{2}, 1, 2$ and 3.

References

References

[1] C J Burke and M Rosenblatt. A markovian function of a markov chain. *The Annals of Mathematical Statistics*, pages 1112–1122, 1958.

[2] C J Burke and M Rosenblatt. Consolidation of probability matrices. *Theorie Statistique*, pages 7–8, 1959.

[3] M Rosenblatt. An aggregation problem for markov chains. *Information and Decision Processes (RE Machol, ed.)*, pages 87–92, 1960.

[4] J G Kemeny and J L Snell. *Finite Markov chains: with a new appendix" Generalization of a fundamental matrix"*. Springer, 1970.

[5] D R Barr and M U Thomas. Technical note - an eigenvector condition for markov chain lumpability. *Operations Research*, 25(6):1028–1031, 1977.

[6] G Rubino and B Sericola. On weak lumpability in markov chains. *Journal of Applied Probability*, pages 446–457, 1989.

[7] G Rubino and B Sericola. A finite characterization of weak lumpable markov processes. part i: The discrete time case. *Stochastic processes and their applications*, 38(2):195–204, 1991.

[8] G Rubino and B Sericola. A finite characterization of weak lumpable markov processes. part ii: The continuous time case. *Stochastic processes and their applications*, 45(1):115–125, 1993.

[9] F Ball and G F Yeo. Lumpability and marginalisability for continuous-time markov chains. *Journal of Applied Probability*, pages 518–528, 1993.

[10] P Buchholz. Exact and ordinary lumpability in finite markov chains. *Journal of applied probability*, pages 59–75, 1994.

[11] M N Jacobi and O Görnerup. A spectral method for aggregating variables in linear dynamical systems with application to cellular automata renormalization. *Advances in Complex Systems*, 12(02):131–155, 2009.

[12] M N Jacobi. A robust spectral method for finding lumpings and meta stable states of non-reversible markov chains. *Electronic Transactions on Numerical Analysis*, 37(1):296–306, 2010.

[13] J Wei and J C W Kuo. Lumping analysis in monomolecular reaction systems. analysis of the exactly lumpable system. *Industrial & Engineering chemistry fundamentals*, 8(1):114–123, 1969.

[14] J Wei and J C W Kuo. Lumping analysis in monomolecular reaction systems. analysis of approximately lumpable system. *Industrial & Engineering chemistry fundamentals*, 8(1):124–133, 1969.

[15] N K Luckyanov, Y M Svirezhev, and O V Voronkova. Aggregation of variables in simulation models of water ecosystems. *Ecological Modelling*, 18(3):235–240, 1983.

[16] Y Iwasa, V Andreasen, and S Levin. Aggregation in model ecosystems. i. perfect aggregation. *Ecological Modelling*, 37(3):287–302, 1987.

[17] G Li and H Rabitz. A general analysis of exact lumping in chemical kinetics. *Chemical engineering science*, 44(6):1413–1430, 1989.

[18] D G Luenberger. Observing the state of a linear system. *Military Electronics, IEEE Transactions on*, 8(2):74–80, 1964.

[19] G Li and H Rabitz. A general analysis of approximate lumping in chemical kinetics. *Chemical engineering science*, 45(4):977–1002, 1990.

[20] G Li, H Rabitz, and J Tóth. A general analysis of exact nonlinear lumping in chemical kinetics. *Chemical Engineering Science*, 49(3):343–361, 1994.

[21] P G Coxson. Lumpability and observability of linear systems. *Journal of mathematical analysis and applications*, 99(2):435–446, 1984.

[22] A Isidori. *Nonlinear control systems*, volume 1. Springer Science & Business Media, 1995.

[23] R L Bryant, S-S Chern, R B Gardner, H L Goldschmidt, and P A Griffiths. *Exterior differential systems*, volume 18. Springer Science & Business Media, 2013.

[24] J Lee. *Introduction to smooth manifolds*, volume 218. Springer Science & Business Media, 2012.

[25] I Kolář, P W Michor, and J Slovák. Natural operations in differential geometry. *Springer Verlag, New York*, 1993.

[26] H Trentelman, A A Stoorvogel, and M Hautus. *Control theory for linear systems*. Springer Science & Business Media, 2012.

[27] M A Fatihcan and L Roncoroni. On the lumpability of linear evolution equations. Preprint, Max Planck Institute for Mathematics in the Sciences, 2013.

[28] M Pflaum. *Ein Beitrag zur Geometrie und Analysis auf stratifizierten Räumen*. 2001.