On systems theoretic aspects of Koopman operator theoretic frameworks

Shen Zeng

Abstract—In this paper, we point out and study a connection between the recently flourishing consideration of Koopman operators and classical systems theoretic concepts such as aggregation and observability decompositions of nonlinear systems. The exploration of this newly unveiled cross-connection promotes a cross-fertilization of different methodologies and ideas intrinsic to the two different frameworks, resulting in a deeper understanding of both domains. In particular, the insights established in the paper connect intuitive systems theoretic viewpoints with the framework of Koopman operators.

I. INTRODUCTION

Inspired by a methodology being applied in econometrics during that time [1], [2], M. Aoki introduced in 1968 in a more encompassing systems theoretic framework the general concept of aggregation of complex and large-scale systems [3], see also [4] and references therein. From a more global point of view, this concept would substantially extend the fundamental notion of conserved quantities (or first integrals) of dynamical systems in physics to a framework that is of more interest when introducing control inputs into the picture.

For a simple introductory example that illustrates the general idea of aggregation (as it is well-known to physicists in one form or another), consider a large-scale particle system

\[ \dot{x}_i = f(x_i). \]

Let us assume that it admits a quantity \( h(x) \) which is a first integral, say the Hamiltonian of the particle system. The whole idea of aggregation is that instead of having to simulate forward the dynamics of the large-scale particle system and then evaluating the Hamiltonian of the particle system at a given time, we would already know before-hand that the Hamiltonian, the quantity of interest, had stayed constant over the given period of time, i.e.

\[ \frac{d}{dt} h(x) = 0. \]

Viewing this example through the lens of systems theory, this fundamental physical concept can in fact be significantly generalized to a much broader circle of ideas that far exceeds the original idea of conservation laws. To illustrate the general idea, consider another particle system, this time consisting of individual particles

\[ \dot{x}_i = Ax_i + Bu_i. \]

One can readily guess that in the present setup the mean

\[ \bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i, \]

might admit a special role. Indeed, we can compute that

\[ \frac{d}{dt} \bar{x} = A\bar{x} + B\bar{u} \]

i.e. the dynamics of \( \bar{x} \) can be described by a closed linear system in the state \( \bar{x} \). It is again that we do not need to simulate forward the dynamics of the large-scale particle system and then compute the mean, as the evolution of the quantity \( \bar{x} \) can be emulated by a much simpler low-dimensional system of the size of a single system. In this methodology, the analysis, and also the control (of the mean), is significantly simplified. Essentially, this approach is a manifestation of the pragmatism to “exclusively focus on what is crucial”.

In this paper, we revisit the somewhat forgotten / overlooked concept of aggregation and highlight its continuing relevance to a number of different systems theoretic problems. In particular, we illuminate a natural connection to recently considered Koopman operator theoretic methods, by which the more abstract functional analytic formalism associated with the Koopman operators are linked to more concrete and intuitive aspects of the underlying differential equations of the considered systems. Conversely, we discuss how the spectral analysis of the Koopman operator provides a natural computational framework to address the problem of finding approximate aggregate models of dynamical systems. This can serve as an important methodology for both analysis and controller design for complex and large-scale dynamical systems that otherwise would be too complex to analyze and to control. For the sake of clarity, we choose to present the ideas in the most concrete and elementary setup, which, in particular, avoids the presentation and discussion of unnecessarily general differential geometric concepts.

The paper is organized as follows. We briefly review the problem of aggregation in the linear framework and provide a novel unified viewpoint that will serve as a concise but comprehensive summary. We then start to connect the concept of aggregation of linear systems to a basic property fundamental to nonlinear systems analysis. This will then be taken as the basis for shedding some light into the systems theoretic aspects of the Koopman operator methodology. The deep connection between aggregation and Koopman operators is explored and highlighted by illustrative examples.
II. THE CONCEPT OF AGGREGATION

Consider a linear (large-scale) system
\[ \begin{align*}
\dot{x} &= Ax + Bu, \\
y &=Cx,
\end{align*} \tag{1} \]

with \( x(t) \in \mathbb{R}^n, \ y(t) \in \mathbb{R}^m \) and \( m \ll n \). The concept of aggregation, which appears to have slowly faded into obscurity over the years, deals with the very basic and interesting question of when
\[ \dot{y} = CAx + CBu \]
can be written as a differential equation in \( y \), i.e. the evolution of \( y \) would only depend on itself. This property that the dynamics of the output \( y \) of a system (1) would be governed by an autonomous system in the variable \( y \) is also referred to as a dynamic exactness of \( y \). In such a case, we may say that the quantity \( y \) can be emulated by a reduced system.

In fact, many important results in both linear and nonlinear systems theory are established by making use of this idea. For example, in nonlinear systems analysis, the concept of Lyapunov functions bears a close resemblance of this idea. For example, in nonlinear systems analysis, the concept of systems theory are established by making use of this idea. Here one is essentially aggregating state information into one scalar quantity, based on which stability of a higher-dimensional system may be inferred by showing e.g. \( V < 0 \). The situation is of course the easiest when one can achieve a closure of the form \( \dot{V} = \lambda V \) where \( \lambda < 0 \), in which case one can immediately conclude asymptotic stability. Indeed, especially in linear systems theory, many important arguments aimed at showing the convergence of some relevant quantity \( e \) to zero hinge on the idea of considering the derivative \( \dot{e} = Ae \) where \( A \) is Hurwitz. Of course, the main results of Lyapunov theory apply to more general cases such as when \( \dot{V} \leq 0 \) (cf. LaSalle’s invariance principle), i.e. no exact closure is needed in general, but the general flavor of the idea of aggregation can still be clearly recognized.

The question of dynamic exactness admits, at least conceptually, a quite straightforward answer. If there exists a matrix \( \bar{A} \in \mathbb{R}^{m \times m} \) such that the matrix equation
\[ CA = \bar{A}C \]
holds, then of course we have
\[ \dot{y} = CAx + CBu = \bar{A}Cx + CBu = \bar{A}y + CBu. \]

Based on this aggregation, one can then, for example, more easily design a controller
\[ u = Ky = (KC)x, \]
which is in fact a static output feedback controller. From a different perspective, when \( y \) does not play the role of an actual physical output, we may view this as a structured state feedback as the feedback matrix naturally factors into the product of the two matrices \( K \) and \( C \). Such low-rank structure may often simplify the control architecture significantly.

In [3], the concept of aggregation was considered in connection with the design of LQR controllers based on a cost functional of the form
\[ J = \int_0^\infty y(t)^\top Q y(t) + u(t)^\top Ru(t) \, dt, \tag{2} \]
subject to the reduced system \( \dot{\bar{y}} = \bar{A}y + CBu \). The obtained LQR controller for the reduced system is then taken as a candidate for controlling the original large-scale system. In [5], the question was somewhat reversed in that an LQR problem for a large-scale system \( \dot{x} = Ax + Bu \), was considered, in which the cost functional was of the form (2) and \( y = Cx \) was viewed as a performance output (not a physical output) summarizing specific quantities of interest. In our study of this problem, we already pointed towards a deep connection between the concept of aggregation and the concept of (un)observability. As we will see, this question has also been extensively considered in the nonlinear control literature [6], however with no explicit reference to the connection with the concept of aggregation. In the remainder of the paper, we work towards bringing together these different fundamental concepts.

III. QUOTIENT FLOWS AND (UN)OBSERVABILITY

An output \( y = Cx \) for which there exists a matrix \( \bar{A} \in \mathbb{R}^{m \times m} \) such that
\[ CA = \bar{A}C \]
defines a system with output that is in some sense “maximally unobservable”. This is because for any \( x \in \ker C \), we have
\[ C(Ax) = CAx = \bar{A}Cx = 0, \]
showing that \( A(\ker C) \subset \ker C \), i.e. \( \ker C \) is \( A \)-invariant. Therefore, \( \ker C \) is equal to the unobservable subspace, which is geometrically described as the largest \( A \)-invariant subspace contained in \( \ker C \).

The above insight can be further discussed in a geometric control theoretic setting, see e.g. [7]. First of all, as for any subspace, we can associate to \( \ker C \) an equivalence relation by virtue of
\[ x' \sim x'': \Leftrightarrow x' - x'' \in \ker C. \]

This equivalence relation induces equivalence classes \([x]\) associated to \( x \in \mathbb{R}^n \), and the family of these equivalence classes yields the quotient space \( \mathbb{R}^n / \ker C \). The elements \([x]\in \mathbb{R}^n / \ker C \) of this abstract vector space can be very naturally identified with the values \( y = Cx \in \mathbb{R}^m \).

For example, for the canonical choice
\[ C = \begin{pmatrix} 0 & I_{m \times m} \end{pmatrix} \]
an element \( \bar{x} \in \mathbb{R}^n / \ker C \) can naturally be identified with the vector \( y = Cx \in \mathbb{R}^m \), which corresponds to a reduction of a state \( x \in \mathbb{R}^n \) to its last \( m \) components, as illustrated in Figure 1.
We note that the above matrix \( C \) is in fact precisely a matrix representation of the canonical projection

\[
\Pi : \mathbb{R}^n \to \mathbb{R}^n / \ker C, \ x \mapsto [x]
\]

with respect to a basis of \( \mathbb{R}^n \) adapted to \( \ker C \), i.e. a basis \( \{v_1, \ldots, v_n-m, v_{n-m+1}, \ldots, v_n\} \) where the first \( n-m \) basis elements form a basis of \( \ker C \).

More generally, for an arbitrary \( C \in \mathbb{R}^{m \times n} \), the value of the output \( y \) can be taken as a characterization of the particular fiber

\[
C^{-1}(\{y\}) := \{x \in \mathbb{R}^n : Cx = y\}
\]
of \( \mathbb{R}^n \). This is an instance of the much broader principle that level sets of output mappings can be naturally used to describe submanifolds.

Now, given a matrix \( A \in \mathbb{R}^{n \times n} \) for which \( \ker C \) is an invariant subspace, one can introduce a well-defined quotient map, as well as a well-defined quotient flow. We observe that for two vectors \( x' \sim x'' \) in the same equivalence class, we have

\[
Ce^{At}(x' - x'') = e^{At}C(x' - x'') = 0,
\]
i.e. \( e^{At}x' \sim e^{At}x'' \) for all \( t \geq 0 \). More explicitly, for any pair of points \( x', x'' \in \mathbb{R}^n \) related by \( Cx' = Cx'' \), we have

\[
e^{At}Cx' = e^{At}Cx'' \tag{3}
\]
for all times \( t \geq 0 \), i.e. the equivalence relation continues to hold into the future.

In this way, we obtain a well-defined flow on the quotient space. By inspecting (3) it becomes evident that an infinitesimal representation of this quotient flow is given by the system \( \dot{y} = Ay \), which describes the evolution of output directly in terms of the current value of the output, and, in particular, does not require an evaluation of information about the full state \( x \) beyond its corresponding output value.

Geometrically, a vector field that induces a well-defined quotient flow on the quotient space \( \mathbb{R}^n / \ker C \) is one in which the component of the vector field that causes a change in the output value (the component normal to \( \ker C \)) is exactly the same for any points on the same slice.

These results can be summarized in the following theorem.

**Theorem 1**: Consider a linear system \( \dot{x} = Ax \) with an output \( y = Cx \), where \( A \in \mathbb{R}^{n \times n} \) and \( C \in \mathbb{R}^{m \times n} \) has full row rank. The following statements are equivalent:

(i) There exists a matrix \( \tilde{A} \in \mathbb{R}^{m \times m} \) such that \( CA = \tilde{A}C \).

(ii) The dynamics of the output can be written as \( \dot{y} = \tilde{A}y \).

(iii) The unobservable subspace of \( (A, C) \) is exactly \( \ker C \).

(iv) The system can be transformed into the form

\[
\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},
\]

where \( y = x_2 \).

For completeness, we include the proof.

**Proof**: (i) \( \Rightarrow \) (ii): We have

\[
\dot{y} = \tilde{A}y, \ y(0) = 0
\]
is given by \( y \equiv 0 \). Since \( x_0 \) is a non-zero vector such that \( y(t) = Ce^{At}x_0 \equiv 0 \), it must belong to the unobservable subspace. This shows that \( \ker C \) is contained in the unobservable subspace, and since the unobservable subspace is trivially contained in \( \ker C \), we have shown the claim.

(ii) \( \Rightarrow \) (iii): Consider a non-zero element \( x_0 \in \ker C \). The solution to

\[
\dot{y} = \tilde{A}y, \ y(0) = 0
\]
is exactly \( y \equiv 0 \). Since \( x_0 \) is a non-zero vector such that \( y(t) = Ce^{At}x_0 \equiv 0 \), it must belong to the unobservable subspace. This shows that \( \ker C \) is contained in the unobservable subspace, and since the unobservable subspace is trivially contained in \( \ker C \), we have shown the claim.

(iii) \( \Rightarrow \) (iv): Choose a basis of \( \mathbb{R}^n \) adapted to \( \ker C \), i.e. the first \( (n-m) \) basis elements form a basis of \( \ker C \). Since \( \ker C \) is \( A \)-invariant, the description of \( A \) in the new basis is given by the claimed upper block triangular form. Moreover, the first \( (n-m) \) basis elements are mapped to zero by \( C \), which, together with the fact that \( C \) has full row rank, yields the claimed structure for the transformed system. (iv) \( \Rightarrow \) (i): Given a system \( (A, C) \), transform it into the form described in (iv). In this canonical form, the result (i) we would like to show clearly holds. Bringing everything back into the original coordinates shows that \( CA = A_{22}C \). □

**IV. AGGREGATION FOR LINEAR SYSTEMS**

In the case of linear systems \( \dot{x} = Ax + Bu \), the question of linear aggregation, i.e. finding a linear output \( y = Cx \) so that its dynamics closes, can be completely answered in terms of the (left)invariant subspaces of \( A \). These in turn can be particularly easily identified by having \( A \) in (real) Jordan form. As a simple example, consider the system

\[
\dot{x} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x
\]

Besides the trivial choices \( C = 0 \) and \( C = I \), we can choose the output matrix \( C = \begin{pmatrix} 0 & 1 \end{pmatrix} \), which is the only left-eigenvector. We have

\[
CA = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 0 \begin{pmatrix} 0 & 1 \end{pmatrix}.
\]

Thus, with the definition of the aggregated variable \( y = Cx \), we have the reduced dynamics \( \dot{y} = 0 \), i.e. \( y \) is a conserved quantity. This is expected as the considered model is a double-integrator that describes the motion of an unforced mass travelling with a constant speed \( y \).
An example that very clearly highlights the added value of taking quadratic outputs into account is given by considering the system
\[ \dot{x} = \begin{pmatrix} \lambda & \omega \\ -\omega & \lambda \end{pmatrix} x. \] (4)

This is an oscillatory system, which has complex eigenvalues \( \lambda \pm i\omega \). It does not admit a (real) scalar linear output \( y = c_1 x_1 + c_2 x_2 \) so that dynamic exactness can be achieved. However, it is a well-known fact that with the choice 
\[ y = x_1^2 + x_2^2, \]
we obtain the closed dynamics 
\[ \dot{y} = 2 \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} \lambda & \omega \\ -\omega & \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2\lambda y. \]

A complete characterization of aggregation for linear systems with polynomial outputs can be achieved through considering the notion of linear tensor systems describing the evolution of monomials of the state of an underlying linear system. To this end, it is useful to recall the notion of tensor systems, cf. [8]. For a multi-index \( \alpha \in \mathbb{N}_0^n \), we define the (multivariate) monomial of the vector \( x \in \mathbb{R}^n \) associated to the multi-index \( \alpha \) by 
\[ x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}. \]

Furthermore, \( x^{[p]} \) shall denote the vector with components \( x^\alpha \) of the same order
\[ |\alpha| = \alpha_1 + \cdots + \alpha_n = p, \]
listed in a lexicographical order. Then, if the evolution of the original state \( x \) is governed by a linear system \( \dot{x} = Ax \), the dynamics of \( x^{[p]} \) can be straightforwardly computed by considering for all \( \alpha \) with order \( |\alpha| = p \)
\[ \frac{d}{dt} x^\alpha = \sum_{i=1}^{n} \alpha_i x_i^{\alpha_i-1} \left( \prod_{j \neq i} x_j^{\alpha_j} \right) \dot{x}_i. \] (5)

This shows that the dynamics of \( x^{[p]} \) can in fact be written in terms of a closed linear system
\[ \frac{d}{dt} x^{[p]} = A_{[p]} x^{[p]}. \] (6)

**Proposition 2:** Given a matrix \( A \) with (not necessarily distinct) eigenvalues \( \lambda_i \), the eigenvalues of \( A_{[p]} \) are of the form \( \sum_{i=1}^{n} \alpha_i \lambda_i \), where \( \alpha \) is a multi-index of order \( p \).

**Proof:** Assume without loss of generality that \( A \) is given in Jordan normal form. Then its diagonal entries \( \lambda_i \) are the (not necessarily distinct) eigenvalues of \( A \) (including multiplicity). The superdiagonal elements of \( A \) are either 0 or 1. More explicitly, if \( A \) is given in Jordan normal form, the autonomous system \( \dot{x} = Ax \) could be described via
\[ \dot{x}_i = \lambda_i x_i + \mu_i x_{i+1} \] (7)
where \( \mu_i \) is either 0 or 1.

Plugging (7) into (5), we obtain
\[ \frac{d}{dt} x^\alpha = \sum_{i=1}^{n} \alpha_i x_i^{\alpha_i-1} \left( \prod_{j \neq i} x_j^{\alpha_j} \right) \left( \lambda_i x_i + \mu_i x_{i+1} \right) \]
\[ = \left( \sum_{i=1}^{n} \alpha_i \lambda_i \right) x^\alpha + \sum_{i=1}^{n} \mu_i \alpha_i x_i^{\alpha_i-1} \left( \prod_{j \neq i} x_j^{\alpha_j} \right) x_{i+1}. \]

Since the terms in the second sum are lexicographic successors of the term \( x^\alpha \), we can conclude that \( A_{[p]} \) has an upper triangular structure with diagonal entries \( \sum_{i=1}^{n} \alpha_i \lambda_i \), from which the result follows.

**Example 3:** We reconsider the oscillatory system (4) in order to shed light into the how the dynamically closed, quadratic output \( y = x_1^2 + x_2^2 \) can be systematically recovered. The key here will be the insights about tensor systems established in this section. Notice that while the original system matrix \( A \) has complex eigenvalues with no real eigenvector corresponding to these, the second order tensor matrix \( A_{[2]} \) has eigenvalues \( 2\lambda \pm 2i\omega \) and \( 2\lambda \). The real eigenvalue \( 2\lambda \) is obtained as a result of the addition of the two complex conjugated eigenvalues \( \lambda + \lambda = 2\Re(\lambda) \). More explicitly, we have
\[ x^{[2]} = \begin{pmatrix} 2\lambda & \omega & 0 \\ -\omega & 2\lambda & \omega \\ 0 & -\omega & 2\lambda \end{pmatrix} x^{[2]}, \]
and it is straightforward to check that
\[ \left( \begin{array}{ccc} 1 & 0 & 1 \\ -\omega & 2\lambda & \omega \\ 0 & -\omega & 2\lambda \end{array} \right) = 2\lambda \left( \begin{array}{ccc} 1 & 0 & 1 \end{array} \right). \]

Clearly, there is a close connection between this linear algebraic insight and the dynamic exactness of the quadratic output \( y = x_1^2 + x_2^2 \). Indeed, we can see that
\[ y = \left( \begin{array}{ccc} 1 & 0 & 1 \end{array} \right) x^{[2]} = \left( \begin{array}{ccc} 1 & 0 & 1 \end{array} \right) \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_2^2 \end{pmatrix} = x_1^2 + x_2^2, \]
as well as
\[ \dot{y} = \left( \begin{array}{ccc} 1 & 0 & 1 \end{array} \right) A_{[2]} x^{[2]} = 2\lambda \left( \begin{array}{ccc} 1 & 0 & 1 \end{array} \right) x^{[2]} = 2\lambda y. \]

To summarize, in the linear case, the Jordan normal form of \( A \) provides complete insight into the invariant subspaces and as such into the problem of finding polynomial outputs so that these are dynamically closed, cf. [9], [10]. An interesting observation was that in the case of complex conjugated eigenvalues, quadratic outputs open up a new possibility of finding dynamically closed outputs with minimal dimension. This essentially covers all the possible exceptional cases for linear systems and polynomial outputs. In the next section, we move towards the study of aggregation in the case of nonlinear systems.
V. Aggregation for Nonlinear Systems

The importance of the viewpoints of observability, as well as the corresponding decompositions discussed in the previous section is reflected in the fact that precisely these concepts were chosen as the fundamental basis for establishing a theory of nonlinear control systems in [6], see also [11]. Indeed, one of the very first introductory discussions regarding a nonlinear system \( \dot{x} = f(x) \), are concerned with the system decomposition

\[
\begin{align*}
\dot{x}_1 &= f_1(x_1, x_2), \\
\dot{x}_2 &= f_2(x_2),
\end{align*}
\]  

(8)

as well as the geometric relevance. If we choose the output

\[ y = h(x) = x_2, \]

we obtain a nonlinear analogue of the observability canonical form in the linear case. Moreover, the output mapping naturally induces a partitioning \( f \) foliation of the state space, where the fibers of the foliation are the level sets

\[ h^{-1}(\{y\}) = \{x \in \mathbb{R}^n : h(x) = y\}. \]

This foliation has the special property that fibers (referred to as “slices” in [6]) of the foliation are transported by the flow into fibers of the same foliation. Figure 2 provides an illustration of the foliation, as well as the geometric implication of the decomposition (8), which is that for any two states \( x', x'' \) on the same slice, i.e. \( h(x') = h(x'') \), any forward propagation of \( x', x'' \), denoted by \( \Phi_t x', \Phi_t x'' \) will satisfy \( h(\Phi_t x') = h(\Phi_t x'') \) for all times \( t \geq 0 \). This would yield a well-defined flow on fibers.

Fig. 2. Left: Illustration of a foliation of a state space into manifolds. Right: Illustration of the property of a flow that carries slices of a given foliation into other slices of the same foliation. Adapted from [6].

In the context of aggregation one is often interested only in finding some quantity \( y = h(x) \) for which the dynamics is closed and not in the full decomposition of the system into the form (8). To this end, we need to find \( h : \mathbb{R}^n \to \mathbb{R}^m \) so that the Lie derivative \( L_f h \) of \( h \) along the vector field \( f \) satisfies

\[ L_f h = w \circ h \iff \frac{\partial h(x)}{\partial x} f(x) = w(h(x)), \]

for some function \( w \) the equation. Then we would obtain the aggregate model \( \dot{y} = w(y) \). This can again be geometrically visualized along the lines of the linear case. The starting point is to view the Lie derivative \( (L_f h)(x) = \frac{\partial h(x)}{\partial x} f(x) \) as a measure for the strength of the “effective” component of the vector field \( f \) at a point \( x \in \mathbb{R}^n \) that is related to a change in the value of \( h(x) \). Now because of the equation \( L_f h = w \circ h \), this means that for all points on the same slice \( S_h(y_0) := \{x \in \mathbb{R}^n : h(x) = y_0\} \), we have the same strength of the “effective” component of the vector field \( f \) given by \( (L_f h)(x) = w(h(x)) = w(y_0) \).

The above presented approach in fact provides a more economic and perhaps more tangible description of specific matters related to the system decomposition (8) that completely circumvents the introduction and discussion of involutive and \( f \)-invariant distributions in an intermediate step, as is commonly found in other related presentations, cf. [6]. The two different approaches discussed here are, however, fundamentally equivalent. Recall that for an integrable distribution \( \Delta \) that is invariant under the vector field \( f \), it holds that for all \( \tau \in \Delta \) we have \( [f, \tau] \in \Delta \). This can be geometrically thought of as follows. Suppose we start at a point \( x \) and move along the vector field \( \tau \) on the integral manifold to another point on the same integral manifold and then with the flow of the vector field \( f \) at that point bringing us to a different integral manifold of \( \Delta \). There we move along the vector field \( -\tau \in \Delta \) and then with the vector field \( -f \).

Because of the property \( [f, \tau] \in \Delta \), we will end up on the same slice of the integrable distribution \( \Delta \) that we started on. This is another more elaborate description of the property that slices are carried into slices.

Finding a pair comprised of a vector field \( f \), an output mapping \( h \) and a reduced mapping \( w \) so that \( L_f h = w \circ h \) is in general not easy and not really suitable for practical approaches. Since the operator \( L_f : h \mapsto L_f h \) is linear, one could e.g. focus instead on the simpler eigenvalue problem \( L_f h = \lambda h \), which one could attempt to solve by parametrizing \( h \) in terms of polynomials and finding the corresponding coefficients. With such an approach, however, there are still significant challenges to be overcome, so that a solution is not immediate. Another idea is to study the situation on the level of flows through the semi-conjugacy (see e.g. [12])

\[ h \circ \Phi_t = \Psi_t \circ h, \]

(9)

in which case we would have

\[ y(t) = h(\Phi_t x_0) = \Psi_t(h(x_0)) = \Psi_t y_0, \]

that is, a well-defined flow for the output variable.

While (9) is reminiscent of the perhaps more well-known concept of conjugacy, such as encountered in the Hartman-Grobman theorem, it is to be stressed that for a semi-conjugacy the mapping \( h : X \to Y \) need only be surjective (i.e. the dimension of \( Y \) is less or equal to the dimension of \( X \)) and not a diffeomorphism or homeomorphism. This forms an important distinction. In Figure 3, the semi-conjugacy (9) is illustrated in terms of a commutative diagram.

In the context of semi-conjugacies, the map \( \Psi_t \) is often called a (topological) factor of \( \Phi_t \) (cf. the terminology for quotient / factor spaces), and, conversely, the map \( \Phi_t \) is called an extension of \( \Psi_t \). This abstract notion of semi-conjugacy provides a rigorous description of the slightly loosely defined idea of having a reduced system emulating
parts of the more complicated system, or, more generally, that the dynamics $\Psi_t$ is contained in the dynamics $\Phi_t$. In fact many systems theoretic concepts and works implicitly refer to this idea, though without any reference to the very interesting general mathematical theory.

The idea of considering the semi-conjugacy $h \circ \Phi_t = \Psi_t \circ h$ rather than the infinitesimal description is what will yield a complete solution for solving the problem of finding aggregate models for nonlinear systems by means of computational methods. This solution is established through a direct link with Koopman operators.

VI. CONNECTION WITH KOOPMAN OPERATORS

Recently, there has been a renewed and ever growing interest in Koopman operators, which are increasingly being used to solve challenging applied problems described by nonlinear differential equations. Koopman operators were originally introduced in 1931 by B. O. Koopman in his study of Hamiltonian dynamical systems in Hilbert spaces [13]. For a fixed time $t \geq 0$, the Koopman operator $U_t$ is acting on “observables”, which, in systems theoretic terms, are simply output functions $h : \mathbb{R}^n \to \mathbb{R}^m$, via a simple composition

$$(U_t h)(x) = h(\Phi_t x) = (h \circ \Phi_t)(x).$$

Because of this simple action, Koopman operators are also referred to as composition operators. There is a direct connection between the Koopman operator theoretic viewpoint and the more systems theoretic viewpoint involving Lie derivatives. The evolution of the function $H(t, \cdot) = U_t h$ is governed by the partial differential equation

$$\frac{\partial}{\partial t} H(t, x) = \frac{\partial}{\partial x} H(t, x) f(x),$$

with the initial condition $H(0, x) = h(x)$.

We recognize that the right-hand side is the Lie derivative $L_f H(t, \cdot) = L_f[U_t h]$. In view of the fact that the equation $L_f h = w \circ h$ is the infinitesimal version of (9), this insight should not come as a surprise. Nevertheless, these cross-connections have not, to the best knowledge of the author, been pointed out in the literature before.

Now as simple as the Koopman operator may appear, the spectrum of the Koopman operator, which is characterized by the eigenvalue equation

$$U_t h = e^{\lambda t} h \iff h(\Phi_t x) = e^{\lambda t} h(x),$$

also referred to as Schröder’s equation, turns out to contain crucial, highly nontrivial, information about different dynamical features of the system, see e.g. [14] for a survey.

In view of our initial study of the concept of aggregation, the problem of finding (output) functions $h : \mathbb{R}^n \to \mathbb{R}^m$ so that for $y = h(x)$ we would have $h \circ \Phi_t = \Psi_t \circ h$ for some flow $\Psi_t$ is clearly solved via the spectral analysis of Koopman operators. Indeed, if $h$ were an eigenvector of $U_t$ with eigenvalue $\lambda$, we would have

$$(U_t h)(x) := (h \circ \Phi_t)(x) = e^{\lambda t} h(x) = (\Psi_t \circ h)(x),$$

with the definition $\Psi_t : y \mapsto e^{\lambda t} y$. Again, passing over to the infinitesimal description, we would have $\dot{y} = \lambda y$, which directly corresponds to the eigenvalue problem $L_f h = \lambda h$ mentioned earlier. This connection also very clearly demonstrates why the eigenvalue equation of the Koopman operator is naturally formulated in terms of $U_t h = e^{\lambda t} h$ rather than $U_t h = \lambda h$. With the increase of computational power over the recent years, Koopman operator-based computational approaches have turned out to be extremely fruitful, and may be directly applied to the systems theoretic problem of identifying aggregate models by virtue of the connections established in this paper.

To further illustrate the connection between our viewpoint of aggregation and the Koopman operator theoretic framework, consider the following illustrative example.

**Example 4:** Reconsider the linear oscillator (4). Define the two complex-valued observation functions

$h_1(x_1, x_2) = x_1 + ix_2, \quad h_2(x_1, x_2) = x_1 - ix_2,$

associated to the two (complex) left-eigenvectors of the system matrix. Translated into the language of Koopman operators, these observation functions are eigenvectors of the Koopman operator, i.e.

$$U_t h_1 = e^{(\lambda + i\omega)t} h_1, \quad U_t h_2 = e^{(\lambda - i\omega)t} h_2.$$  

Due to the basic property $U_t[h_1 h_2] = (U_t h_1)(U_t h_2)$, the product of the two observation functions

$$h(x_1, x_2) = (x_1 + ix_2)(x_1 - ix_2) = x_1^2 + x_2^2$$

satisfies the eigenvalue equation

$$U_t h = e^{(\lambda + i\omega) - (\lambda - i\omega)t} h = e^{2\lambda t} h.$$

This is in perfect accordance with our insights derived from the actual system dynamics using the viewpoint of aggregation.

$$(U_t h)(x) = h(e^{\lambda t} x) = e^{2\lambda t}(x_1^2 + x_2^2).$$

The two different but fundamentally connected viewpoints are summarized in Figure 4. Because the dynamics of the output $y = h(x)$ is closed, the forward mapping of $h$ to $U_t h$ has the property that all forward mappings $U_t h$ share the same shape of level sets. In more detail, due to relation $(U_t h)(x) = \Psi_t (h(x))$, a level set of $U_t h$ associated to a value $y^*$ is a level set of $h$ associated to a value $\Psi_t^{-1}y^*$. Conversely, only when all forward mappings $U_t h$ share the same shape of level sets, then the dynamics of the output can be described as an autonomous system that only depends on the output itself.
To illustrate how the insights from the connection to the concept of aggregation in fact significantly add to our understanding of the related problems involving the more analytic framework of Koopman operators, in the following we show how two important results from [15] rather trivially follow from the systems theoretic viewpoint via systems with outputs. More specifically, the following result summarizes Theorem 1 and Corollary 1 in [15] rather succinctly.

**Proposition 5:** Consider a nonlinear system $\dot{x} = f(x)$ and an output mapping $h : \mathbb{R}^n \to \mathbb{R}^m$ so that with the definition $y = h(x)$, we have $\dot{y} = Ay$ with a Hurwitz matrix $A \in \mathbb{R}^{m \times m}$. Then the manifold $M = \{ x \in \mathbb{R}^n : h(x) = 0 \}$ is invariant under $\dot{x} = f(x)$ and globally attracting.

**Proof:** A displacement of a point $x_0 \in \mathbb{R}^n$ from the manifold $M = \{ x \in \mathbb{R}^n : h(x) = 0 \}$, which is simply a level set of $h$, can be characterized in terms of the value $y_0 := h(x_0) \neq 0$. The solution of $\dot{y} = Ay$, $y(0) = y_0$ is converging to the origin $y = 0$ in an exponential fashion. This means for the original system that the $x$-trajectory is approaching the subspace exponentially.

VII. CONCLUSIONS

In this paper, we pointed out a connection between the important concepts of aggregation, observability, and Koopman operators. These insights seem promising in eventually establishing a unified theoretical framework. Indeed, there are countless instances scattered across the field of linear and nonlinear systems theory where the fact that the dynamics of an output can be written as an autonomous (often linear) differential equation forms a crucial ingredient in the derivation of many important results. A key contribution of this paper was to bridge the systems theoretic concept of aggregation with the framework of Koopman operators, which provides new viewpoints on known results for the Koopman operator that originally had been formulated solely in terms of a functional analytic formalism. This yielded novel structural insights based on the underlying dynamical system, which were illustrated by means of simple examples.

We pointed out that the dynamic exactness of an output can be geometrically illustrated by a special alignment of the vector field and output mapping, a particular kind of (abstract) symmetry. Thus, from a more general point of view, our treatment emphasizes how the fundamental concepts of symmetry, invariance, continuity, and reduction of dynamics all belong to the same circle of ideas.

REFERENCES