On Duality of Behavioural Systems

Harish K. Pillai
ISIS Research Group, Department of Electronics and Computer Science,
University of Southampton, SO17 1BJ Southampton, UK
e-mail:hp@ecs.soton.ac.uk
and
H. Narayanan
Department of Electrical Engineering
Indian Institute of Technology, Bombay
India

Abstract

The notions of controllability and observability are duals in the traditional input-state-output framework of systems theory for 1-D systems. Recently, there has been a study on the duality between controllability and observability in behavioural systems [2] by the construction of a suitable adjoint for 1-D behavioural systems. In this paper we show, among other things, that formally this definition carries over also to n-D behavioural systems. We then explore the nature of the duality between controllability and observability for n-D systems using this definition of adjoint.

Introduction

The input-state-output framework of System theory is well developed and widely accepted. One of the most attractive features of this theory is the duality between the notions of controllability and observability. Behavioural systems theory has as its basic credo, the principle that systems should be studied as far as possible 'as they are' without casting them into artificial canonical forms and without separating their relevant variables into 'input' and 'output'. Definitions of controllability and observability have been already extended to behavioural system theory. However it is not clear how far the duality between these concepts carries over from the input-state-output framework to the behavioural framework. We continue the study of duality begun for 1-D systems in [2] systematically for N-D systems (see also [3] in this connection). We use a defintion of 'adjoint' similar to the one introduced in [2] for this purpose. We also make systematic use of certain 'Implicit Duality Theorems' from [2],[3].

Preliminaries

In this section, we first define some operations on modules over $\mathbb{R}[\xi]$. Here ξ is a shorthand notation for $\xi_1, \ldots, \xi_n$ and thus $\mathbb{R}[\xi]$ is the ring of polynomials in n indeterminates.

Let W be a finite set, say $W = \{1, 2, \ldots \mathtt{w}\}$. A vector f on W over $\mathbb{R}[\xi]$ is a function with domain W and range $\mathbb{R}[\xi]$. We will refer to elements of the ring $\mathbb{R}[\xi]$ as scalars. A *chain group* on W (see for instance [12]) over $\mathbb{R}[\xi]$ is a collection of vectors on W closed under addition and scalar multiplication. So, a chain group on W is in reality a $\mathbb{R}[\xi]$-module.

We consider the free module $\mathbb{R}^{1\times \mathtt{w}}[\xi]$. The *standard dot product* $< f, g >$ of $f = (f_1, \ldots, f_{\mathtt{w}})$, and $g = (g_1, \ldots, g_{\mathtt{w}}) \in \mathbb{R}^{1\times \mathtt{w}}[\xi]$ is given by

$$< f, g > := \sum_{i=1}^{\mathtt{w}} f_i g_i$$

We say f, g are *orthogonal* to each other iff $< f, g >= 0$.

Given a chain group $\mathcal{K}_W$ on W, we denote by $\mathcal{K}_W^*$, the collection of all vectors orthogonal to $\mathcal{K}_W$. Quite clearly, for any given $\mathcal{K}_W$, $\mathcal{K}_W \subseteq (\mathcal{K}_W^*)^*$. We say $\mathcal{K}_W$ is a *reflexive* chain group over the standard dot product if $(\mathcal{K}_W^*)^* = \mathcal{K}_W$.

Lemma 1 *Let $\mathcal{K}_W$ be a chain group over $\mathbb{R}[\xi]$ on W. Then $\mathcal{K}_W^*$ is reflexive over the standard dot product.*

Lemma 2 *Let the rows of a matrix $R(\xi)$ generate a chaingroup $\mathcal{K}_R$ over $\mathbb{R}[\xi]$ on some finite set W and the rows of the matrix $R(-\xi)$ generate another chaingroup $\mathcal{K}_{-R}$ over $\mathbb{R}[\xi]$ on the same finite set W. Then, if $\mathcal{K}_R$ is reflexive over the standard dot product, $\mathcal{K}_{-R}$ is also reflexive over the standard dot product.*

The next definition is found useful in the study of electrical network topology [4] and 1-D systems [2].

Let W and L be two disjoint finite sets. Let $\mathcal{K}_{WL}$, $\mathcal{K}_L$ denote chain groups on $W \uplus L$ and L respectively. The *generalized minor* $\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L$ of $\mathcal{K}_{WL}$ with respect to $\mathcal{K}_L$ is defined by

$$\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L := \{f_W : \exists f_{WL} \in \mathcal{K}_{WL} \text{ s.t. } f_W = f_{WL}/W, \text{and} \, f_{WL}/L \in \mathcal{K}_L\}$$

In words, this means that the generalized minor $\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L$ consists of those vectors on W which arise as restrictions of vectors on $W \uplus L$ that belong to $\mathcal{K}_{WL}$ and whose restrictions onto L are contained in $\mathcal{K}_L$.

We denote by $\mathcal{K}_{WL} + \mathcal{K}_L$, the collection of vectors $f_{WL} + (0 \mid f_L)$ on $W \uplus L$, where $f_{WL} \in \mathcal{K}_{WL}$ and $f_L \in \mathcal{K}_L$.

For any chain group $\mathcal{K}_{WL}$ of vectors on $W \uplus L$, we call

$$\mathcal{K}_{WL} \cdot L = \{f_L : f_L = f/L, \text{ where } f \in \mathcal{K}_{WL}\}$$

the *restriction* of $\mathcal{K}_{WL}$ to L and we call

$$\mathcal{K}_{WL} \times L = \{f_L : f_L = f/L, \text{ where } f \in \mathcal{K}_{WL} \text{ and } f/W = 0\}$$

the *contraction* of $\mathcal{K}_{WL}$ to L.

We obtain the following lemma.

Lemma 3 *Let $\mathcal{K}_{WL}$ be a chain group over $\mathbb{R}[\xi]$ on $W \uplus L$. Then (a) $(\mathcal{K}_{WL} \cdot L)^* = \mathcal{K}^*_{WL} \times L$. (b) In addition, if $\mathcal{K}_{WL}$ and $\mathcal{K}^*_{WL} \cdot L$ are reflexive (with respect to standard dot product), then $(\mathcal{K}_{WL} \times L)^* = \mathcal{K}^*_{WL} \cdot L$.*

We now state the Implicit Duality theorem. Various versions of this theorem for vector spaces, cones, unimodular modules etc. can be found in [2, 4]. For N-D systems, this result is true for reflexive chain groups and may be found in [3]. In order to make the present paper self contained we give a proof of this result in Appendix 1. It is interesting to note that the same skeleton of proof goes through for all the known versions of this 'meta' theorem.

Theorem 4 (Implicit Duality Theorem) *Let $\mathcal{K}_{WL}, \mathcal{K}_L$ be modules on $W \uplus L, L$ respectively, such that $\mathcal{K}_{WL} + \mathcal{K}_L$ is reflexive. Then,*
(a) $(\mathcal{K}^*_{WL} \leftrightarrow \mathcal{K}^*_L)^* = \mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L$
(b) $\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L$ is reflexive.
If, in addition, $\mathcal{K}_{WL}, \mathcal{K}_L$, $\mathcal{K}^*_{WL} + \mathcal{K}^*_L$ are reflexive, then
(c) $(\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L)^* = \mathcal{K}^*_{WL} \leftrightarrow \mathcal{K}^*_L$
(d) $\mathcal{K}^*_{WL} \leftrightarrow \mathcal{K}^*_L$ is reflexive.

Note that we have stated the Implicit Duality theorem using the concept of reflexive chain groups, which in turn was defined using the standard dot product that we defined earlier. We now look at a more general setting.

We define a vector w on a finite set W over $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$ as a function with domain W and range $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$. The ring of scalars will again be $\mathbb{R}[\xi]$. A chain group on W over $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$ is a collection of vectors on W closed under addition and scalar multiplication. Note here, that scalar multiplication is defined as follows : if $p \in \mathbb{R}[\xi]$ and $w \in \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$, then

$$p(\xi_1, \ldots, \xi_n)w := p(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n})w$$

Thus, the scalar multiplication is in reality, differentiation. Again note that such a chain group is also a $\mathbb{R}[\xi]$-module. Also note that we could have chosen some other function space like distributions instead of $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$. But in this paper we will restrict ourselves to $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$.

We now define a new *mixed dot product* between two chain groups on W, one of them over $\mathbb{R}[\xi]$ and the other over $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$. Let $f = (f_1, \ldots, f_{\mathtt{w}}) \in \mathbb{R}^{1 \times \mathtt{w}}[\xi]$ and $w = (w_1, \ldots, w_{\mathtt{w}}) \in (\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}))^{\mathtt{w}}$. Then

$$< f, w > := \sum_{i=1}^{\mathtt{w}} fw$$

We call f and w *orthogonal with respect to the mixed dot product* if $< f, w >= 0$

Let $\mathcal{K}_W$ be a chain group on W over $\mathbb{R}[\xi]$ or $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$. We call the chain group $\mathcal{K}_W^\star$ obtained by collecting together all w which are orthogonal with respect to the mixed dot product to every element of $\mathcal{K}_W$. We can again define reflexive chain groups (now with respect to the mixed dot product) by the condition that a chain group $\mathcal{K}_W$ is reflexive with respect to the mixed dot product iff $(\mathcal{K}_W^\star)^\star = \mathcal{K}_W$. Note that the choice of the function space makes a difference as to which chain groups on W over $\mathbb{R}[\xi]$ are reflexive with respect to the mixed dot product. Such reflexive chain groups have been studied in the control literature, and they go by the name *Willems submodules* (for more details, see [7]). We now quote an important theorem from [7] which was first shown by Oberst [5].

Theorem 5 *(Oberst). Every submodule of the free module $\mathbb{R}^{1\times \mathtt{w}}[\xi]$ is reflexive with respect to the mixed dot product (or equivalently, Willems with respect to $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$).*

Note that the above theorem also holds for other solution spaces, like distributions. In such cases, the appropriate dot product should be defined.

The Implicit Duality theorem can now be stated ([3]) for such reflexive chain groups, defined by the mixed dot product.

Theorem 6 *Let $\mathcal{K}_{WL}, \mathcal{K}_L$ be chain groups (or modules) over $\mathbb{R}[\xi]$ on $W \uplus L, L$ respectively, let $\mathfrak{B}_{WL} \equiv (\mathcal{K}_{WL})^\star$ and let $\mathfrak{B}_L \equiv (\mathcal{K}_L)^\star$. Then (a) $(\mathcal{K}_{WL} \longleftrightarrow \mathcal{K}_L)^\star = \mathcal{K}_{WL}^\star \longleftrightarrow \mathcal{K}_L^\star$. (b) $(\mathfrak{B}_{WL} \longleftrightarrow \mathfrak{B}_L)^\star = \mathfrak{B}_{WL}^\star \longleftrightarrow \mathfrak{B}_L^\star$.*

For completeness we repeat the proof given in [3], in Appendix 1.

The Implicit Duality theorem and its variations (for example, the above theorem) have several ramifications in behavioural theory of systems. So we now introduce some concepts from the behavioural theory.

Multidimensional Systems

We view a system as a family of trajectories mapping a set of "independent" variables into a set of "dependent" variables. See [11] for an elaboration of this with examples. Thus a *system* Σ is defined as a triple $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ where $\mathbb{T}$ is the indexing set, the set of independent variables, $\mathbb{W}$ is the signal space, the set of dependent variables, and $\mathfrak{B} \subset \mathbb{W}^{\mathbb{T}}$ is the behaviour. In the this paper we consider systems with $\mathbb{T} = \mathbb{R}$ (we call these *lumped dynamical* systems or 1-D systems) and systems with $\mathbb{T} = \mathbb{R}^n$ (we call these *distributed* systems or n-D systems). Also, we assume throughout that $\mathbb{W}$ is a finite dimensional real vector space, $\mathbb{W} = \mathbb{R}^{\mathtt{w}}$.

We consider systems $\Sigma = (\mathbb{R}^n, \mathbb{R}^{\mathtt{w}}, \mathfrak{B})$, where $\mathfrak{B}$ is the solution set of a system of constant coefficient partial differential equations. More precisely, if there exists a real polynomial matrix $R \in \mathbb{R}^{\bullet \times \mathtt{w}}[\xi]$ in n indeterminates, $\xi = (\xi_1, \ldots, \xi_n)$, such that $\mathfrak{B}$ consists of the $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$-solutions of

$$R(\frac{d}{d\mathbf{x}})w = 0 \tag{1}$$

where $\frac{d}{d\mathbf{x}} = (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n})$. We shall only consider $\mathfrak{C}^\infty$-solutions, but the results obtained are equally valid for other solution concepts, for example, for distributions. We denote the family of systems $\Sigma = (\mathbb{R}^n, \mathbb{R}^{\mathtt{w}}, \mathfrak{B})$ as $\mathfrak{L}_n^{\mathtt{w}}$. We also denote $(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}}, \mathfrak{B}) \in \mathfrak{L}_n^{\mathtt{w}}$ as $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$, since the indexing set and the signal space are then obvious from the context. Note that $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ is a chain group on the finite set $W = \{1, 2, \ldots, \mathtt{w}\}$ over $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$.

Lemma 7 *A system $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ is uniquely specified by its annhilators, defined by*

$$\mathfrak{B}^\star \quad = \quad \left\{ p \in \mathbb{R}^{1 \times \mathtt{w}}[\xi] \mid p(\frac{d}{d\mathbf{x}})\mathfrak{B} = 0 \right\}$$

$\mathfrak{B}^\star$ is a submodule of $\mathbb{R}^{1 \times \mathtt{w}}[\xi]$ viewed as a module over $\mathbb{R}[\xi]$ and thus a chain group. Conversely, given a submodule of $\mathbb{R}^{1 \times \mathtt{w}}[\xi]$, we can uniquely identify a behaviour to this submodule.

The above lemma is just a restatement of Theorem 5. So, there is a one-to-one relation between $\mathfrak{L}_n^{\mathtt{w}}$ and the submodules of $\mathbb{R}^{1 \times \mathtt{w}}[\xi]$. Thus, whereas $R \in \mathbb{R}^{\bullet \times \mathtt{w}}[\xi]$ uniquely specifies a behaviour $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ through (1) with $\mathfrak{B}^\star$ the module generated by the rows of R, any other polynomial matrix whose rows generate the same submodule define the same behaviour.

The family of systems $\mathfrak{L}_n^{\mathtt{w}}$ enjoys many convenient properties and this has been studied in detail in [7] using the *fundamental principle*. The fundamental principle is actually a classical result from the theory of partial differential equations, which can be found in [1, 6].

Theorem 8 *(Ehrenpreis-Palamodov). Let A be a $\mathtt{g} \times \mathtt{w}$ matrix with entries from $\mathbb{R}[\xi]$ such that the columns of A generate a chain group $\mathcal{K}_A$ on the finite set $G = \{1, 2, \ldots, \mathtt{g}\}$. Let b be in $(\mathfrak{C}^\infty)^{\mathtt{g}}$. Then there exists a w in $(\mathfrak{C}^\infty)^{\mathtt{w}}$ such that $A(w) = b$ iff $q(f) = 0$ for all q in $\mathcal{K}_A^*$.*

Whereas we have *defined* the behaviour of a system in $\mathfrak{L}_n^{\mathtt{w}}$ as the set of solutions of a system of PDE's in the system variables, often, in practical applications, the specification of the behaviour involves other, auxillary variables, which we call *latent variables.* Specifically, consider the system of PDE's

$$R(\frac{d}{dx})w = M(\frac{d}{dx})\ell \tag{2}$$

with $w \in \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$ and $\ell \in \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\ell})$ and with $R \in \mathbb{R}^{\bullet \times \mathtt{w}}[\xi]$ and $M \in \mathbb{R}^{\bullet \times \ell}[\xi]$ polynomial matrices with the same number of rows. The set

$$\mathfrak{B}_f = \{(w, \ell) \in \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}+\ell}) \mid (2) \text{ holds}\} \tag{3}$$

obviously belongs to $\mathfrak{L}_n^{\mathtt{w}+\ell}$. It immediately follows from the elimination theorem that the set

$$\{w \in \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}}) \mid \exists \ell \in \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^\ell) : (w, \ell) \in \mathfrak{B}_f\} \tag{4}$$

belongs to $\mathfrak{L}_n^{\mathtt{w}}$. We call (2) a *latent variable representation* (with manifest variables w and latent variables ℓ) of the system with *full behaviour* (3) and *manifest behaviour* (4). Correspondingly, we call (1) a kernel representation of the system with the behaviour $\ker(R(\frac{d}{d\mathbf{x}}))$. We shall soon meet another sort of representation, the image representations, in the context of controllability.

Consider the equation (2), which can be rewritten as

$$R(\frac{d}{d\mathbf{x}})w - M(\frac{d}{d\mathbf{x}})\ell = 0$$

Let $\mathcal{K}_{WL}$ be the module generated by the rows of the matrix $[R(\xi) \ \ -M(\xi)]$. Here W is the finite set corresponding to the manifest variables and L is the one corresponding to the latent variables. Theorem 9 ('the elimination theorem'), below, is an important consequence of the fundamental principle. It gives an algorithm for the kernel representation of the manifest behaviour.

Theorem 9 *Let the module $\mathcal{K}_{WL}$ define the behaviour $\mathfrak{B}_{WL}$ as in (3). Then $\mathcal{K}_{WL} \times W$ forms a module whose generators give a kernel representation of the manifest behaviour.*

Proof Let the module $\mathcal{K}_L$ be the zero module which acts on the latent variables. The corresponding behaviour $\mathfrak{B}_l = \mathcal{K}_L^\star$ are all possible trajectories of the latent variables. We now apply Theorem 6 and obtain $\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L = (\mathfrak{B}_f \leftrightarrow \mathfrak{B}_l)^\star$. But $\mathfrak{B}_f \leftrightarrow \mathfrak{B}_l$ is equal to the manifest behaviour and $\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L = \mathcal{K}_{WL} \times W$. □

As a result of the elimination theorem, we obtain several interesting results. Let $F \in \mathbb{R}^{\mathtt{w}_1 \times \mathtt{w}_2}[\xi]$. Then $\mathfrak{B}_2 \in \mathfrak{L}_n^{\mathtt{w}_2}$ implies $F(\frac{d}{d\mathbf{x}})\mathfrak{B}_2 \in \mathfrak{L}_n^{\mathtt{w}_1}$ and $\mathfrak{B}_1 \in \mathfrak{L}_n^{\mathtt{w}_1}$ implies $(F(\frac{d}{d\mathbf{x}}))^{-1}\mathfrak{B}_1 \in \mathfrak{L}_n^{\mathtt{w}_2}$. This in particular implies that if $\mathfrak{B}_1, \mathfrak{B}_2 \in \mathfrak{L}_n^{\mathtt{w}}$, then $\mathfrak{B}_1 \cap \mathfrak{B}_2 \in \mathfrak{L}_n^{\mathtt{w}}$ and $\mathfrak{B}_1 + \mathfrak{B}_2 \in \mathfrak{L}_n^{\mathtt{w}}$. Another implication of the elimination theorem states that for any $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}_1+\mathtt{w}_2}$, the set

$$\{w_1 \in \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}_1}) \mid \exists w_2 \in \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}_2}) : (w_1, w_2) \in \mathfrak{B}\}$$

is itself an element of $\mathfrak{L}_n^{\mathtt{w}_1}$.

Controllability and Observability

Controllability and observability of dynamical systems are two of the most influential properties defined in the classical setting. These properties have been rephrased in the behavioural setting in [13] for lumped dynamical systems and in [7, 8] generalizations to distributed systems have been introduced.

Definition 10 *A system $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ is said to be* controllable *if for all $w_1, w_2 \in \mathfrak{B}$ and for all bounded sets $U_1, U_2 \subset \mathbb{R}^n$ with disjoint closure, there exists a $w \in \mathfrak{B}$ such that $w \mid_{U_1} = w_1 \mid_{U_1}$ and $w \mid_{U_2} = w_2 \mid_{U_2}$*

Thus controllable PDE's are those in which the solutions can be 'patched up' from solutions on subsets. There are a several characterizations of controllability. An important chracterization of controllability is the existence of an image representation [7]. Consider the following special latent variable representation

$$w = M(\frac{d}{d\mathbf{x}})\ell \tag{5}$$

with $M \in \mathbb{R}^{\mathtt{w}\times\ell}[\xi]$. Obviously, by the elimination theorem, its manifest behaviour $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$. Here $\mathfrak{B} = \mathrm{im}(M(\frac{d}{d\mathbf{x}}))$ with $M(\frac{d}{d\mathbf{x}})$ viewed as a map from $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^\ell)$ to $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$. For this reason, we call (5) an *image representation* of its manifest behaviour. Whereas every $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ allows (by definition) a kernel representation and hence trivially a latent variable representation, not every $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ allows an image representation. In fact:

Theorem 11 *$\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ admits an image representation if and only if it is controllable.*

We denote the set of controllable systems in $\mathfrak{L}_n^{\mathtt{w}}$ by $\mathfrak{L}_{n,\mathrm{cont}}^{\mathtt{w}}$. Interestingly, there is another characterization of controllable systems shown in [7].

Lemma 12 *Let $\mathcal{K}_W$ be a submodule of $\mathbb{R}^{1\times\mathtt{w}}[\xi]$ such that $\mathcal{K}_W^\star = \mathfrak{B}$. Then $\mathfrak{B}$ is controllable iff $\mathcal{K}_W$ is reflexive with respect to the standard dot product (i.e., $\mathcal{K}_W^{**} = \mathcal{K}_W$).*

Proof Assume that one is given an image representation of a behaviour $\mathfrak{B}$, i.e., $w \in \mathfrak{B}$ iff $w = M(\frac{d}{d\mathbf{x}})\ell$. Let the columns of M generate a chain group $\mathcal{K}_M$. Then by Theorem 8, $w \in \mathfrak{B}$ iff $q(w) = 0$ for all $q \in \mathcal{K}_M^*$. But by Lemma 1, $\mathcal{K}_M^*$ is reflexive with respect to the standard dot product. $\square$

Given any behaviour $\mathfrak{B} = \mathcal{K}_W^\star$, it contains a unique largest controllable sub-behaviour $\mathfrak{B}_c$. This sub-behaviour $\mathfrak{B}_c$ is called the controllable part of the behaviour $\mathfrak{B}$. It has been shown in [7] that $\mathfrak{B}_c = (\mathcal{K}_W^{**})^\star$.

Observability is a property of systems that have two kinds of variables - the first set are the observed set of variables and the second set are the variables which are to be deduced from the observed variables. Thus, observability is not an intrinsic property of a behaviour. One needs to be given a partition of variables into two classes before one can comment on whether one set of variables are observable from the other set of variables.

Definition 13 *Let $\mathfrak{B}$ be a behaviour and let $w = (w_1, w_2)$ be a partition of variables in $\mathfrak{B}$. Then w_2 is said to be* observable *from w_1 in $\mathfrak{B}$ if given any two trajectories $(w_1', w_2'), (w_1'', w_2'') \in \mathfrak{B}$ such that $w_1' = w_1''$, then $w_2' = w_2''$.*

In other words, we can uniquely identify the w_2 part of the trajectory $w \in \mathfrak{B}$, once we know w_1. A natural situation where the variables are partitioned into two types, is when one has a latent variable representation. In this situation, it is natural to ask whether the latent variables are observable from the manifest variables. If this is the case, then we call the latent variable representation *observable.*

Given a latent variable representation $Rw = M\ell$ (with R, M being matrices of appropriate dimensions and with entries from $\mathbb{R}[\xi]$), the latent variables ℓ are observable from the manifest variables w if and only if the matrix M has full column rank for all $\xi \in \mathbb{C}^n$ [8].

It has been mentioned earlier that every controllable behaviour has an image representation. A natural question that springs to mind is - does every controllable behaviour have an observable image representation. It turns out that for 1-D systems, every controllable behaviour has an observable image representation. Sadly, in n-D systems, this is no longer true. There are several characterizations of controllable behaviours that have an observable image representation [8]. One of the characterizations is the following: a controllable behaviour that has an observable image representation, also has a kernel representation $\mathfrak{B} = \ker R(\xi)$, where $R(\xi)$ has *constant* row rank for all $\xi \in \mathbb{C}^n$. We will denote the set of controllable systems in $\mathfrak{L}_n^{\mathtt{w}}$ that have an observable image representation by $\mathfrak{L}_{n,\text{obs}}^{\mathtt{w}}$. Then the following is true :

$$\mathfrak{L}_{n,\text{obs}}^{\mathtt{w}} \;\subseteq\; \mathfrak{L}_{n,\text{cont}}^{\mathtt{w}} \;\subseteq\; \mathfrak{L}_n^{\mathtt{w}}$$

Moreover, in 1-D case, we have

$$\mathfrak{L}_{1,\text{obs}}^{\mathtt{w}} \;=\; \mathfrak{L}_{1,\text{cont}}^{\mathtt{w}} \;\subseteq\; \mathfrak{L}_1^{\mathtt{w}}$$

In classical input-state-output setting of 1-D control theory, the concepts of controllability and observability are duals of each other. Thus, given a dynamical system which is controllable and observable, one can construct an "adjoint" system which is also controllable and observable. We shall now use the Implicit Duality theorem to construct an "adjoint" for a n-D behaviour which generalizes the concept of adjoints for 1-D systems.

Adjoints

First, we take a look at what is the meaning of an adjoint. In matrix theory, the classical terminology refers to the transpose of the matrix of co-factors associated with a square matrix A as the adjoint of A. Thus, $A^{-1} = (\text{adjoint } A)/\det A$. On the other hand, in the theory of linear transformations on an inner product space, the adjoint transformation associated with a given transformation L is defined as that transformation L^{adj} such that $< y, L(x) >=< L^{adj}(y), x >$. It is this latter concept of adjoints that we will be interested in, since it is this definition of adjoints that is used in differential equation theory.

Note that we are looking at behaviours that are subsets of $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$. Though we cannot define an inner product on $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$, we can define an inner product on compactly supported trajectories in $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$ (denoted by $\mathfrak{D}(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$), by

$$< u, v >= \int_{\mathbb{R}^n} u^T v d\mathbf{x}$$

This integral is not well defined over all of $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$ and hence one has to be satisfied with this restricted sense of inner product.

Let us consider 1-D behaviours $\mathfrak{B}_1$ and $\mathfrak{B}_2 \in \mathfrak{L}_1^{\mathtt{w}}$. We call $\mathfrak{B}_1$ and $\mathfrak{B}_2$ *complementary* if $\mathfrak{B}_1 \oplus \mathfrak{B}_2 = \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^{\mathtt{w}})$. It can be shown that both $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are controllable. We call these behaviours *dual* if they are complementary and if $< w_1, w_2 >= 0$ for all $w_1 \in \mathfrak{B}_1 \cap \mathfrak{D}(\mathbb{R}, \mathbb{R}^{\mathtt{w}})$ and $w_2 \in \mathfrak{B}_2 \cap \mathfrak{D}(\mathbb{R}, \mathbb{R}^{\mathtt{w}})$. In this case, we denote $\mathfrak{B}_2 = \mathfrak{B}_1^\perp$.

For the n-D case, given a controllable behaviour $\mathfrak{B}_1$, it is not possible to always find a behaviour $\mathfrak{B}_2$ such that $\mathfrak{B}_1 \oplus \mathfrak{B}_2 = \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$. But we can still look at the compactly supported trajectories in the behaviours and ask if a restricted sense of dual exists, i.e., given $\mathfrak{B}_1 \in \mathfrak{L}^{\mathtt{w}}_{n,\mathrm{cont}}$, is it possible to find $\mathfrak{B}_2 \in \mathfrak{L}^{\mathtt{w}}_{n,\mathrm{cont}}$ such that $< w_1, w_2 >= 0$ for all $w_1 \in \mathfrak{B}_1 \cap \mathfrak{D}(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$ and $w_2 \in \mathfrak{B}_2 \cap \mathfrak{D}(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$ and $(\mathfrak{B}_1 \cap \mathfrak{D}(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})) \oplus (\mathfrak{B}_2 \cap \mathfrak{D}(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})) = \mathfrak{D}(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$. It turns out, that this is indeed possible. In this case, we again denote $\mathfrak{B}_2 = \mathfrak{B}_1^\perp$. In this case, we call $\mathfrak{B}_2$ the dual of $\mathfrak{B}_1$.

In both the above cases, 1-D and n-D, we have $(\mathfrak{B}_1^\perp)^\perp = \mathfrak{B}_1$. Moreover, it can also be shown that if $R(\frac{d}{d\mathbf{x}})w = 0$ is a kernel representation of the controllable $\mathfrak{B}$, then $w = R^T(-\frac{d}{d\mathbf{x}})\ell$ is an image representation of $\mathfrak{B}^\perp$. (Note that the operator $R^T(-\frac{d}{d\mathbf{x}})$ is the adjoint of the operator $R(\frac{d}{d\mathbf{x}})$). For the 1-D case, this duality extends to state representations [14], in the following sense. If

$$E\frac{dx}{dt} + Fx + Gw = 0 \tag{6}$$

is a minimal state representation of $\mathfrak{B}$ (with number of states equal to m), then $\mathfrak{B}^\perp$ also admits a minimal state representation

$$E'\frac{dz}{dt} + F'z + G'v = 0 \tag{7}$$

with the same number of states, such that for all $(w, x) \in \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^{\mathtt{w}+m})$ satisfying (6), and for all $(v, z) \in \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^{\mathtt{w}+m})$ satisfying (7), there holds the following kind of duality involving the state

$$\frac{d}{dt}z^T x = v^T w$$

Thus any integral of the form $\int_a^b v^T w dt$ for $v \in \mathfrak{B}^\perp$ and $w \in \mathfrak{B}$ depends only on the state of the systems at the instances a and b and is independent of the trajectory. Such a pair of minimal state maps (X, Z) of $\mathfrak{B}$ and $\mathfrak{B}^{\perp}$ is called a matched pair. For more details, see [14].

For n-D systems, the best we can do is the following: given a controllable behaviour $\mathfrak{B} = \ker R(\frac{d}{d\mathbf{x}}) = \operatorname{im} M(\frac{d}{d\mathbf{x}})$ and its dual $\mathfrak{B}^{\perp} = \operatorname{im} R^T(-\frac{d}{d\mathbf{x}})$, it is possible to find operators in 2n variables $\Psi_1, \ldots, \Psi_n \in \mathbb{R}^{m_1 \times m_2}[\zeta, \eta]$ which induce bilinear differential forms on the latent variables ℓ and ℓ' that appear in the image representations of $\mathfrak{B}$ and $\mathfrak{B}^{\perp}$ respectively, such that

$$\frac{\partial}{\partial x_1} L_{\Psi_1}(\ell, \ell') + \cdots + \frac{\partial}{\partial x_n} L_{\Psi_n}(\ell, \ell') = v^T w$$

where $\mathfrak{B} \ni v = M(\frac{d}{d\mathbf{x}})\ell$ and $\mathfrak{B}^{\perp} \ni w = R^T(-\frac{d}{d\mathbf{x}})\ell'$. Thus, we obtain the product $v^T w$ as the divergence of a set of bilinear differential forms that act on the latent variables. Note that in the 1-D case, divergence equals $\frac{d}{dt}$. Moreover, if we consider $\int_\Omega v^T w d\mathbf{x}$ over some closed bounded region $\Omega \subset \mathbb{R}^n$, then this integral depends only on the values of the latent variables ℓ and ℓ' on the boundary $\partial\Omega$ of Ω. Thus, the properties of the integral $\int_\Omega v^T w d\mathbf{x}$ are completely analogous to the properties of the integral mentioned earlier in relation to 1-D systems. For more details on this result and information about such bilinear differential forms, see [9].

Thus, we would like the "adjoint" to be defined in such way that given a controllable behaviour $\mathfrak{B}$, its adjoint is its dual $\mathfrak{B}^{\perp}$. Moreover, we would like some condition similar to the divergence condition to hold. Finally, we want the construction of the adjoint to be computationally a very easy procedure.

So first, let us consider a controllable behaviour $\mathfrak{B}$ given by $\mathfrak{B} = \ker R(\frac{d}{d\mathbf{x}})$. Let the rows of $R(\xi)$ generate a chain group $\mathcal{K}_W$ over $\mathbb{R}[\xi]$ on W. Here W is a finite set having the same cardinality as the number of manifest variables of the behaviour. Consider the chain group $\mathcal{K}_{WL}$ over $\mathbb{R}[\xi]$ on $W \uplus L$, where the cardinality of L is equal to the number of rows of the matrix R. Let $\mathcal{K}_{WL}$ be generated by the rows of the matrix $\begin{bmatrix} R(\xi) & I \end{bmatrix}$ where I is the identity matrix. Consider the chain group $\mathcal{K}_L$ on L generated by the rows of the matrix I. By Theorem 5, all these chain groups are reflexive with respect to the mixed dot product. Thus using Theorem 6, we can say that the $(\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L) = \mathfrak{B}^{\star}$. Clearly, the chain groups $\mathcal{K}_{WL}, \mathcal{K}_L$ are also reflexive with respect to the standard dot product. Moreover, since $\mathcal{K}_W$ is reflexive with respect to the standard dot product (because $\mathfrak{B}$ is controllable, lemma 12), the chain group $\mathcal{K}_{WL} + \mathcal{K}_L$ is also reflexive with respect to the standard dot product. Using Implicit Duality Theorem (Theorem 4), we can conclude that $\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L = (\mathcal{K}_{WL}^* \leftrightarrow \mathcal{K}_L^*)^*$. Now $\mathcal{K}_{WL}^*$ is generated by the rows of the matrix $\begin{bmatrix} I & -R^T(\xi) \end{bmatrix}$ and $\mathcal{K}_L^*$ is generated by the rows of the matrix 0. Thus, $\mathcal{K}_{WL}^* \leftrightarrow \mathcal{K}_L^*$ gives rise to a behaviour given by

$$\mathfrak{B}_1 = \{w : \exists \ell \text{ such that } w = R^T(\frac{d}{d\mathbf{x}})\ell\}$$

But this behaviour $\mathfrak{B}_1$ is not the equal to $\mathfrak{B}^\perp$ and we are interested in the behaviour $\mathfrak{B}^\perp$. But notice that

$$\mathfrak{B}^\perp = \{w : \exists \ell \text{ such that } w = R^T(-\frac{d}{d\mathbf{x}})\ell\}$$

The similarity of the equations describing $\mathfrak{B}_1$ and $\mathfrak{B}^\perp$ gives us an inkling as to how one can calculate the module that corresponds to $\mathfrak{B}^\perp$. The idea is to replace ξ in the concerned equations by $-\xi$. This can be done in several ways. For example, we could consider the ring $\mathbb{R}[\imath\xi_1, \imath\xi_2, \ldots, \imath\xi_n]$ instead of $\mathbb{R}[\xi_1, \ldots, \xi_n]$, where $\imath = \sqrt{-1}$ and redefine the dot product accordingly (by taking complex conjugates). But that would lead to a whole babel of different dot products. So, we will just stick to substituting ξ by $-\xi$ in the resulting equations.

Let $\mathfrak{B}$ have a latent variable representation given by

$$\left[\begin{array}{cc} R(\xi) & M(\xi) \end{array} \right] \left(\begin{array}{c} w \\ \ell \end{array} \right) = 0 \tag{8}$$

We define the adjoint $\mathfrak{B}^{adj}$ as having a latent variable representation

$$\left[\begin{array}{cc} I & -R^T(-\xi) \\ 0 & -M^T(-\xi) \end{array} \right] \left(\begin{array}{c} \hat{w} \\ \hat{\ell} \end{array} \right) = 0 \tag{9}$$

A similar construction was obtained in [2]. In the above equations, w and $\hat{w}$ are the manifest variables whereas ℓ and $\hat{\ell}$ are the latent variables. We denote the full behaviours in the two cases by $\mathfrak{B}_f$ and $\mathfrak{B}_f^{adj}$ respectively. Note that both w and $\hat{w}$ have the same number of entries, whereas the same cannot be said about the cardinality of ℓ and $\hat{\ell}$.

A few comments are in order here. Clearly, if $\mathfrak{B} = \ker R(\frac{d}{d\mathbf{x}})$ is controllable, then $\mathfrak{B}^{adj} = \operatorname{im} R^T(-\frac{d}{d\mathbf{x}}) = \mathfrak{B}^\perp$. Thus, this construction does give us the dual behaviour as the adjoint behaviour, when the original behaviour is controllable and expressed in kernel representation. In fact, the adjoint of any behaviour expressed in the kernel representation is controllable. As a result of [7, Proposition 1], we can say more about the adjoint of a controllable behaviour. If the controllable behaviour $\mathfrak{B}$ is expressed as the kernel of some other operator R_1, then it can be shown that the adjoint $\mathfrak{B}^{adj}$ obtained using R_1 is the same as the adjoint obtained using R.

Lemma 14 *Let $\mathfrak{B} = \ker R = \ker R_1$. Then the adjoint behaviour $\mathfrak{B}^{adj}$ obtained by the prescribed method of adjoint construction is the same in both cases.*

Proof Let $\mathfrak{B}^{adj} = \operatorname{im} R^T(-\xi)$ and $\mathfrak{B}_1^{adj} = \operatorname{im} R_1^T(-\xi)$. Since the rows of $R(\xi)$ and the rows of $R_1(\xi)$ generate the same chain group, we conclude that the columns of $R^T(-\xi)$ and the columns of $R_1^T(-\xi)$ generate the same chain group. From [7, Proposition 1], we have that every image representation is equivalent to the image of a corresponding right prime morphism. Since the two chain groups generated by the columns of $R^T(-\xi)$ and $R_1^T(-\xi)$ are the same, the corresponding right prime morphism is the same and thus, $\mathfrak{B}^{adj} = \mathfrak{B}_1^{adj}$. $\square$

Thus, using the adjoint construction procedure stated above, we find that any controllable behaviour expressed in kernel representation produces the dual behaviour as the adjoint behaviour. Let us now consider the more general, latent variable representation of behaviours and study how the method of adjoint construction that we have stated before, behaves in these situations. Note that if we obtain a new latent variable representation for the same behaviour by a unimodular transformation of the equations involved, then the adjoint of the system defined in this fashion is the same as before. We can see this by considering a new latent variable representation for $\mathfrak{B}$ obtained from (8) by unimodular transformation of the equations. Thus

$$U(\xi)\begin{bmatrix} R(\xi) & M(\xi) \end{bmatrix}\begin{pmatrix} w \\ \ell \end{pmatrix} = 0$$

where $U(\xi)$ is a unimodular matrix. (By unimodular matrix, we mean a square matrix, that has a non-zero constant determinant.) Then the adjoint is given by

$$\begin{aligned}\begin{bmatrix} I & -R^T(-\xi)U^T(-\xi) \\ 0 & -M^T(-\xi)U^T(-\xi) \end{bmatrix}\begin{pmatrix} \hat{w} \\ \hat{\ell} \end{pmatrix} &= \begin{bmatrix} I & -R^T(-\xi) \\ 0 & -M^T(-\xi) \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & U^T(-\xi) \end{bmatrix}\begin{pmatrix} \hat{w} \\ \hat{\ell} \end{pmatrix} \\ &= \begin{bmatrix} I & -R^T(-\xi) \\ 0 & -M^T(-\xi) \end{bmatrix}\begin{pmatrix} \hat{w} \\ \hat{\ell_1} \end{pmatrix} = 0\end{aligned}$$

where $\hat{\ell_1} = U^T(-\xi)\hat{\ell}$ is a unimodular transformation of the latent variables. Clearly, $\mathfrak{B}^{adj}$ remain unchanged by a unimodular transformation of the equations involved.

Given a controllable behaviour $\mathfrak{B}$, the adjoint (constructed in the fashion described above) is not unique. It depends essentially on the representation of the behaviour and this is to be expected, since the construction of the adjoint depends on the representation of the behaviour. We now show this by means of an example.

Example 15 Consider the controllable behaviour given by the latent variable representation $\begin{bmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} \end{bmatrix}\begin{pmatrix} w \\ \ell \end{pmatrix} = 0$. Clearly, the manifest behaviour $\mathfrak{B}$ is all of $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$ and so a kernel representation of $\mathfrak{B} = \ker R(\frac{d}{d\mathbf{x}})$ is given by the matrix $R = 0$. Constructing the adjoint with the initial latent variable representation, we get $\mathfrak{B}^{adj}_{latent} = \ker \frac{\partial}{\partial x_2}$. On the other hand, constructing the adjoint using the kernel representation of $\mathfrak{B}$, one obtains $\mathfrak{B}^{adj}_{kernel} = \ker 1 = \mathfrak{B}^\perp$. Clearly, the adjoint is not unique since $\mathfrak{B}^{adj}_{latent} \neq \mathfrak{B}^{adj}_{kernel}$ and it is the adjoint $\mathfrak{B}^{adj}_{kernel}$ that we ideally require to be the adjoint behaviour. $\square$

As exhibited earlier, the adjoint is invariant under various kernel representations of a controllable behaviour, but in the case of latent variable representations, the adjoint need not be invariant under different representations. This is clear from the earlier example. We can conclude that the latent variables are responsible for spoiling the invariance of the adjoint. So, we should explore the possibility of putting some constraints on the latent variables, which should (hopefully) plug this inadequacy. We propose the following restriction on the latent variables – the columns of the matrix $M(\xi)$ should generate a reflexive chain group with respect to the standard dot product. Note that this restriction on the latent variables is very similar in flavour to the requirement in the classical 1-D state space theory, that the number of states of the system should be minimal and they must be observable. Moreover, in the case of kernel representations of controllable behaviours, since the matrix $M(\xi) = 0$, this restriction is trivially satisfied.

We now use the Implicit Duality Theorem. We have

Theorem 16 *Let the behaviour $\mathfrak{B}$ and its adjoint $\mathfrak{B}^{adj}$ be defined in latent variable representations as in (8) and (9) respectively. Let $\mathfrak{B} = \ker R_w(\xi)$ and $\mathfrak{B}^{adj} = \ker R^{adj}(\xi)$ be kernel representations of the manifest behaviours. Let the rows of $R_w(\xi)$ define a chain group $\mathcal{K}_{R_w}$ over $\mathbb{R}[\xi]$ on the set W (which has the cardinality equal to the number of manifest variables of $\mathfrak{B}$) and let the rows of $R^{adj}(-\xi)$ generate another chain group $\mathcal{K}_{R_a}$ over $\mathbb{R}[\xi]$ on W. If the rows of the matrix $\begin{bmatrix} R(\xi) & M(\xi) \end{bmatrix}$ generate a reflexive chain group with respect to the standard dot product, then*

$$\mathcal{K}_{R_w} = \mathcal{K}^*_{R_a}$$

If the columns of the matrix $M(\xi)$ generate a reflexive chain group with respect to the standard dot product, then

$$\mathcal{K}^*_{R_w} = \mathcal{K}_{R_a}.$$

Proof Let $\mathfrak{B}_f$ be the full behaviour for the latent variable representation (8) of the behaviour $\mathfrak{B}$ and let $\mathfrak{B}^{adj}_f$ be the full behaviour for the latent variable representation (9) of the adjoint behaviour $\mathfrak{B}^{adj}$. Consider the discrete set W to have the cardinality of w (and hence $\hat{w}$). Let L have the cardinality of the latent variables ℓ appearing in (8). Furthermore, let L' be another discrete set with cardinality equal to the number of equations in (8). Consider the chain group $\mathcal{K}_{WLL'}$ over $\mathbb{R}[\xi]$ on $W \uplus L \uplus L'$ generated by the rows of the matrix $\begin{bmatrix} R(\xi) & 0 & I \end{bmatrix}$. Similarly, consider the chain group $\mathcal{K}_{LL'}$ over $\mathbb{R}[\xi]$ on $L \uplus L'$ generated by the rows of the matrix $\begin{bmatrix} M(\xi) & -I \end{bmatrix}$. By using the results of Theorem 5 and Theorem 6, we obtain $\mathcal{K}_{WLL'} \leftrightarrow \mathcal{K}_{LL'} = \mathfrak{B}^{\star}$. Thus, $\mathcal{K}_{WLL'} \leftrightarrow \mathcal{K}_{LL'} \simeq \mathcal{K}_{R_w}$ defined in the statement of this theorem.

Using the standard dot product, we obtain the following: generating matrix for $\mathcal{K}^*_{WLL'}$ is $\begin{bmatrix} 0 & I & 0 \\ I & 0 & -R^T(\xi) \end{bmatrix}$ and the generating matrix for $\mathcal{K}^*_{LL'}$ is given by $\begin{bmatrix} I & M^T(\xi) \end{bmatrix}$. Note that the generating matrix for $\mathcal{K}^*_{WLL'} \leftrightarrow \mathcal{K}^*_{LL'}$ (after simplification) is the matrix obtained from

$$\begin{bmatrix} I & -R^T(\xi) \end{bmatrix} \leftrightarrow \begin{bmatrix} -M^T(\xi) \end{bmatrix}$$

Moreover, $\mathcal{K}^*_{WLL'} \leftrightarrow \mathcal{K}^*_{LL'}$ is a chain group defined on the set W. Note that $\mathcal{K}^*_{WLL'} \leftrightarrow \mathcal{K}^*_{LL'} \simeq \mathcal{K}_{R_a}$, where $\mathcal{K}_{R_a}$ is as defined in the statement of the theorem.

Observe that both $\mathcal{K}_{WLL'}$ and $\mathcal{K}_{LL'}$ are reflexive with respect to the standard dot product. In addition, $\mathcal{K}_{WLL'} + \mathcal{K}_{LL'}$ is also reflexive with respect to the standard dot product if the chain group generated by the rows of $\begin{bmatrix} R(\xi) & M(\xi) \end{bmatrix}$ is reflexive with respect to the standard dot product. So invoking the Implicit Duality Theorem (Theorem 4) we can conclude that $\mathcal{K}_{R_w} \simeq \mathcal{K}_{WLL'} \leftrightarrow \mathcal{K}_{LL'} = (\mathcal{K}^*_{WLL'} \leftrightarrow \mathcal{K}^*_{LL'})^* \simeq \mathcal{K}^*_{R_a}$.

Note that the chain groups $\mathcal{K}^*_{WLL'}$ and $\mathcal{K}^*_{LL'}$ are reflexive with respect to the standard dot product. If the columns of $M(\xi)$ generate a reflexive chain group with repect to the standard dot product, then this translates to the fact that $\mathcal{K}^*_{WLL'} + \mathcal{K}^*_{LL'}$ is reflexive with respect to the standard dot product. Applying the Implicit Duality Theorem (Theorem 4) once more, we conclude that $\mathcal{K}^*_{R_w} = (\mathcal{K}_{WLL'} \leftrightarrow \mathcal{K}_{LL'})^* = \mathcal{K}^*_{WLL'} \leftrightarrow \mathcal{K}^*_{LL'} = \mathcal{K}_{R_a}$. $\square$

The condition that the rows of $\begin{bmatrix} R(\xi) & M(\xi) \end{bmatrix}$ generate a reflexive chain group with respect to the standard dot product implies two things – (a) the full behaviour $\mathfrak{B}_f$ is controllable and (b) from the above theorem, since $\mathcal{K}_{R_w} = \mathcal{K}^*_{R_a}$, using lemma 1, we conclude that the manifest behaviour $\mathfrak{B}$ is also controllable. This statement is always true (for 1-D and n-D cases), without any conditions on the number of rows in the representation matrix.

On the other hand, the condition that the columns of the matrix $M(\xi)$ generates a reflexive chain group with respect to the standard dot product is not enough to guarantee that the latent variables ℓ are observable from the manifest variables w in the full behaviour $\mathfrak{B}_f$ (for both 1-D and n-D cases). But it is enough to guarantee that the adjoint of the behaviour $\mathfrak{B}$ is controllable, since $\mathcal{K}_{R_a} = \mathcal{K}^*_{R_w}$ (lemma 1). In fact, one can guarantee that for any representation of given behaviour $\mathfrak{B}$ that satisfies the above two conditions (i.e., the rows of $\begin{bmatrix} R(\xi) & M(\xi) \end{bmatrix}$ and the columns of $M(\xi)$ generate reflexive chain groups with respect to the standard dot product), both $\mathfrak{B}$ and the adjoint $\mathfrak{B}^{adj}$ are controllable and $\mathfrak{B}^{adj} = \mathfrak{B}^{\perp}$. We therefore submit that Theorem 16 is the most basic duality theorem as far as behavioural theory of systems is concerned.

Let us now revisit Example 15. Note that the matrix $\begin{bmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} \end{bmatrix}$ does generate a reflexive chain group with respect to the standard dot product and so $\mathfrak{B} \in \mathfrak{L}^1_{2,\text{cont}}$. On the other hand, the columns of $M(\xi) = \begin{bmatrix} \frac{\partial}{\partial x_2} \end{bmatrix}$ do not generate a reflexive chain group (w.r.t. the standard dot product) and hence, $\mathfrak{B}^{adj}_{latent} \notin \mathfrak{L}^1_{2,\text{cont}}$. Also observe that the controllable part of $\mathfrak{B}^{adj}_{latent}$ is equal to $\mathfrak{B}^{\perp}$. This observation prompts the following –

Corollary 17 *Consider $\mathfrak{B}$ and $\mathfrak{B}^{adj}$ and the corresponding $\mathcal{K}_{R_w}$ and $\mathcal{K}_{R_a}$ as in Theorem 16. Then if the rows of the matrix $\begin{bmatrix} R(\xi) & M(\xi) \end{bmatrix}$ generate a reflexive chain group with respect to the standard dot product, then $(\mathcal{K}^*_{R_w})^\star$ is equal to the controllable part of $\mathcal{K}^\star_{R_a}$. Similarly, if the columns of $M(\xi)$ generate a reflexive chain group with respect to the standard dot product, then $(\mathcal{K}^*_{R_a})^\star$ is equal to the the controllable part of $\mathcal{K}^\star_{R_w}$.*

Proof From Theorem 16, we know that if the rows of the matrix $\begin{bmatrix} R(\xi) & M(\xi) \end{bmatrix}$ generate a reflexive chain group with respect to the standard dot product, then $\mathcal{K}_{R_w} = \mathcal{K}^*_{R_a}$. Clearly, if $\mathcal{K}_{R_w} = \mathcal{K}^*_{R_a}$, then $\mathcal{K}^*_{R_w} = \mathcal{K}^{**}_{R_a}$. Moreover, $(\mathcal{K}^{**}_{R_a})^\star$ is the controllable part of $\mathcal{K}^\star_{R_a}$. Hence the result. Similar arguments prove the second part of the statement of this corollary. □

The above corollary gives us a very interesting result. It states that if some behaviour $\mathfrak{B} \in \mathfrak{L}^{\mathtt{w}}_{n,\mathrm{cont}}$ is described in a latent variable representation, such that the full behaviour $\mathfrak{B}_f \in \mathfrak{L}^{\mathtt{w}+\ell}_{n,\mathrm{cont}}$, then the adjoint that we obtain $\mathfrak{B}^{adj}$ has the following very interesting property – the controllable part of $\mathfrak{B}^{adj}$ is in reality $\mathfrak{B}^\perp$. So even though, the adjoint behaviour is not unique (since it depends upon the representation), the controllable part of the adjoint is uniquely defined, provided the full behaviour is controllable. Similarly, if the columns of $M(\xi)$ generate a reflexive chain group with respect to the standard dot product, then the controllable part of the behaviour $\mathfrak{B}$ is the dual of $\mathfrak{B}^{adj}$. Note that $\mathfrak{B}^{adj}$ in this case is controllable.

In classical 1-D state space framework, the notions of controllability and observability are duals of one another, i.e., given a state space representation of a system which is controllable and observable, one obtains a dual system which is controllable and observable. Note that if a state representation of a system is controllable and observable, then it is a minimal representation of the system. As stated earlier, this duality of minimal state representations of 1-D controllable behaviours was demonstrated in [14]. Quite clearly, one cannot expect controllability and observability (as defined in the behavioural setting) to be duals, since controllability is an intrinsic property of the behaviour, whereas observability is not. So, it is better to look for some intrinsic property of a behaviour, that acts as a dual to controllability. One such intrinsic property of a behaviour is the property of admitting an observable image representation. As stated previously, in 1-D systems, every controllable behaviour has an observable image representation, whereas this is not true for the n-D case ($n > 1$).

We revisit the characterization of controllable behaviours. Apart from having an image representation (Theorem 11) and the related chain group being reflexive with respect to the standard dot product (Lemma 12), there are several other characterizations of controllability. In 1-D systems, there is another additional characterization of controllable behaviours, which comes from the fact that every controllable 1-D behaviour has an observable image representation. If $\mathfrak{B} = \ker R$, the matrix $R(\xi)$ having constant rank for all $\xi \in \mathbb{C}^n$ gives a characterization of behaviours that have an observable image representation. Therefore $R(\xi)$ having constant rank for all $\xi \in \mathbb{C}$ is an additional characterization of controllable 1-D behaviours. (Note that this condition only requires that $R(\xi)$ should have constant rank.) Also note the similarity with condition for the latent variables ℓ to be observable from the manifest variables w that we have mentioned in the previous section (the condition being that the matrix $M(\xi)$ must have full column rank for all $\xi \in \mathbb{C}^n$). As a result of all this, we get the following:

Proposition 18 *Let the behaviour $\mathfrak{B}$ and its adjoint $\mathfrak{B}^{adj}$ be defined in latent variable representations as in (8) and (9) respectively. Let $\mathfrak{B} = \ker R_w(\xi)$ and $\mathfrak{B}^{adj} = \ker R^{adj}(\xi)$ be kernel representations of the manifest behaviours. Let the rows of $R_w(\xi)$ define a chain group $\mathcal{K}_{R_w}$ over $\mathbb{R}[\xi]$ on the set W (which has the cardinality equal to the number of manifest variables of $\mathfrak{B}$) and let the rows of $R^{adj}(-\xi)$ generate another chain group $\mathcal{K}_{R_a}$ over $\mathbb{R}[\xi]$ on W. If the matrix $\begin{bmatrix} R(\xi) & M(\xi) \end{bmatrix}$ has full row rank for all $\xi \in \mathbb{C}^n$, then*

$$\mathcal{K}_{R_w} = \mathcal{K}^*_{R_a}$$

If the matrix $M(\xi)$ has full column rank for all $\xi \in \mathbb{C}^n$, then

$$\mathcal{K}^*_{R_w} = \mathcal{K}_{R_a}.$$

Proof Clearly, if the matrix $\begin{bmatrix} R(\xi) & M(\xi) \end{bmatrix}$ has full row rank for all $\xi \in \mathbb{C}^n$, then its rows generate a reflexive chain group with respect to the standard dot product. Using Theorem 16, we have the first part of the result. Similarly, if $M(\xi)$ has full column rank for all $\xi \in \mathbb{C}^n$, then the columns of M generate a reflexive chain group with respect to the standard dot product. Hence, the second part of the conclusion (Theorem 16). □

Let $\mathfrak{B}_f$ be the full behaviour of the representation (8) of $\mathfrak{B}$ and let $\mathfrak{B}_f^{adj}$ be the full behaviour of the latent variable representation of the adjoint system. Then the condition in the last proposition that $\begin{bmatrix} R(\xi) & M(\xi) \end{bmatrix}$ has full row rank for all $\xi \in \mathbb{C}^n$ implies that $\mathfrak{B}_f$ is controllable and it has an observable image representation. The other condition that $M(\xi)$ has full column rank for all $\xi \in \mathbb{C}^n$ implies that the latent variables ℓ are observable from the manifest variables w in $\mathfrak{B}_f$. The above proposition then implies that the adjoint system (9) has a full behaviour $\mathfrak{B}_f^{adj}$ which is controllable and its latent variables $\hat{\ell}$ are observable from its manifest variables $\hat{w}$. Moreover, $\mathfrak{B}_f^{adj}$ has an observable image representation. Thus we conclude that if $\mathfrak{B} \in \mathfrak{L}^{\mathtt{w}}_{n,\mathrm{obs}}$, then $\mathfrak{B}^{adj} \in \mathfrak{L}^{\mathtt{w}}_{n,\mathrm{obs}}$. We had demostrated earlier that the adjoint behaviour of a controllable behaviour is another controllable behaviour. In addition, if the controllable behaviour has an observable image representation, then the adjoint behaviour also has an observable image representation. It is now clear why controllability and observability are dual concepts in the 1-D theory. Using the characterization of controllability and observability for 1-D systems, we get the following familiar corollary:

Corollary 19 *Let $R(\frac{d}{dt})w + M(\frac{d}{dt})\ell = 0$ be a latent variable representation of a behaviour $\mathfrak{B}$. Let $\mathfrak{B}_f$ be the full behaviour of the above representation. If $\mathfrak{B}_f$ is controllable and its latent variables ℓ are observable from the manifest variables w, then the full behaviour of the adjoint system given by*

$$\begin{bmatrix} I & -R^T(-\frac{d}{dt}) \\ 0 & -M^T(-\frac{d}{dt}) \end{bmatrix} \begin{pmatrix} \hat{w} \\ \hat{\ell} \end{pmatrix} = 0$$

is also controllable and its latent variables $(\hat{\ell})$ are observable from its manifest variables $(\hat{w})$ (provided the number of equations defining $\mathfrak{B}_f$ is minimal). Moreover, the manifest behaviour of the adjoint system is the dual of $\mathfrak{B}$.

Proof Straightforward from Proposition 18. Note that the minimality condition appears in this corollary, since in Proposition 18 we have a *full* row rank condition on the matrix $\begin{bmatrix} R(\xi) & M(\xi) \end{bmatrix}$, whereas the condition for $\mathfrak{B}_f$ to be controllable and have an observable image representation is a *constant* row rank condition. Thus, we cannot have redundant equations in the kernel representation of the full behaviour $\mathfrak{B}_f$, since it affects the observability of the latent variables in the adjoint system. $\square$

This last corollary makes it amply clear why "minimality" of the state representation is awfully important when one considers the duality of controllability and observability in the classical 1-D state space framework.

Applications

We will now apply our method of constructing adjoints to problems well known in control literature. This exercise will help us check if the adjoints that we construct are compatible with the well known adjoints in the literature.

First we will consider the input/state/output model in 1-D systems, given by

$$\begin{aligned} \frac{dx}{dt} &= Ax + Bu \\ y &= Cx + Du \end{aligned}$$

Here A, B, C, D are constant matrices of appropriate dimensions, x is the state variable (latent variable in behavioural terms), u is the input and y is the output. We rewrite the equations as

$$\begin{bmatrix} -B & 0 & \xi I - A \\ -D & I & -C \end{bmatrix} \begin{pmatrix} u \\ y \\ x \end{pmatrix} = 0$$

Here $w^T = (u^T, y^T)$ and $\ell = x$. The adjoint will then be

$$\begin{bmatrix} I & 0 & B^T & D^T \\ 0 & I & 0 & -I \\ 0 & 0 & \xi I + A^T & C^T \end{bmatrix} \begin{pmatrix} \hat{w}_1 \\ \hat{w}_2 \\ \hat{\ell}_1 \\ \hat{\ell}_2 \end{pmatrix} = 0$$

Noting that $\hat{w}_2 = \hat{\ell}_2$, we can now rewrite these equations as

$$\begin{aligned} \frac{dz}{dt} &= -A^T z - C^T \hat{w}_2 \\ \hat{w}_1 &= -B^T z - D^T \hat{w}_2 \end{aligned}$$

Here $z = \hat{\ell}_1$. We know from control literature that the above equations describe the adjoint system, provided the original system is state minimal, controllable and observable. From Corollary 19, we know that if the full behaviour is controllable and its latent variables are observable, then the full behaviour of the adjoint system is also controllable and its latent variables are observable. For the full behaviour to be controllable, the condition to be satisfied is $\begin{bmatrix} \xi I - A & -B \end{bmatrix}$ should have full row rank for all $\xi \in \mathbb{C}$ (which is classical state-space controllability). On the other hand, the latent variables in the full behaviour being observable, translates to $\begin{bmatrix} \xi I - A \\ -C \end{bmatrix}$ having full column rank for all $\xi \in \mathbb{C}$ (classical state-space observability). Moreover, if the original system is minimal, then so is the adjoint. In addition,

$$-\frac{d}{dt} z^T x = \hat{w_2}^T y + \hat{w_1}^T u$$

Setting $z = -z$ gets rid of that negative sign. Thus, the divergence condition that we refered to earlier, is also satisfied.

We now look at another set of well known duals that occur in control theory. Let us consider a *driving variable representation* of a system

$$\begin{aligned} \frac{dx}{dt} &= Ax + Bd \\ w &= Cx + Dd \end{aligned}$$

Here both x (the state variables) and d (the driving variables) are latent variables. Rewriting the equations, we get

$$\begin{bmatrix} 0 & -B & \xi I - A \\ I & -D & -C \end{bmatrix} \begin{pmatrix} w \\ d \\ x \end{pmatrix} = 0$$

The adjoint is given by

$$\begin{bmatrix} I & 0 & -I \\ 0 & B^T & D^T \\ 0 & \xi I + A^T & C^T \end{bmatrix} \begin{pmatrix} \hat{w} \\ \hat{\ell}_1 \\ \hat{\ell}_2 \end{pmatrix} = 0$$

Noting that $\hat{w} = \hat{\ell}_2$ and sustituting $z = -\hat{\ell}_1$, we obtain

$$\begin{aligned} \frac{dz}{dt} &= -A^T z + C^T \hat{w} \\ 0 &= B^T z - D^T \hat{w} \end{aligned}$$

From Theorem 16 and Corollary 19 we conclude that the full behaviour $\mathfrak{B}_f$ is controllable if the matrix $\begin{bmatrix} \xi I - A & -B \end{bmatrix}$ has full row rank for all $\xi \in \mathbb{C}$. Moreover, the full behaviour of the adjoint is controllable if the matrix $\begin{bmatrix} -B & \xi I - A \\ -D & -C \end{bmatrix}$ has full column rank. In fact, if these rank conditions are fulfilled, then the representation above gives a *output nulling representation* of the dual behaviour. Again, we note that

$$\frac{d}{dt} z^T x = \hat{w}^T w$$

Of course, if we start with a system in output nulling representation, then we obtain the dual in driving variable representation, provided some matrix conditions are satisfied.

We now look at another application for this method of constructing adjoints. This time we use the general n-D setting and the result goes through for the 1-D case too. We had defined earlier the dual of a controllable behaviour. Note that this concept of dual may itself be generalized. For this we need to use the theory of bilinear differential forms. Please see [9, 10] for definitions of bilinear and quadratic differential forms. We also add an appendix at the end of this paper, for an introduction to this theory.

Given a $\Phi \in \mathbb{R}^{\mathtt{w} \times \mathtt{w}}[\zeta, \eta]$ which is symmetric, that generates a BLDF L_Φ, one can define the following: Given a controllable behaviour $\mathfrak{B}$, define a controllable behaviour $\mathfrak{B}^{\perp_\Phi}$ as one such that $L_\Phi(v, w) = 0$ for all $v \in (\mathfrak{B} \cap \mathfrak{D}(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}}))$ and $w \in (\mathfrak{B}^{\perp_\Phi} \cap \mathfrak{D}(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}}))$. We call such a behaviour $\mathfrak{B}^{\perp_\Phi}$ as the dual of $\mathfrak{B}$ with respect to the BLDF induced by Φ. Note that $\mathfrak{B}^\perp = \mathfrak{B}^{\perp_I}$, i.e., $\mathfrak{B}^\perp$ is the dual of $\mathfrak{B}$ with respect to the BLDF induced by I (where I is the identity matrix). Again in this case, one would first like to know if such a dual exists, and if it does, one would like some sort of construction procedure that guarantees the correct construction of a dual with respect to a BLDF.

What we do in this case is the following: Let $\mathfrak{B} = \ker R(\frac{d}{d\mathbf{x}})$. Consider the set of equations given by

$$\begin{aligned} R(\frac{d}{d\mathbf{x}})\ell &= 0 \\ \Phi(-\frac{d}{d\mathbf{x}}, \frac{d}{d\mathbf{x}})\ell &= w \end{aligned}$$

Construct the dual of this system, considering w to be the manifest variables and ℓ to be the latent variables. Note that if $\mathfrak{B}$ is controllable, then the first part of the conditions of Theorem 16 are satisfied. On the other hand, if the rows of the matrix $\begin{bmatrix} -\Phi^T(-\xi,\xi) & R^T(\xi) \end{bmatrix}$ generate a reflexive chain group with respect to the standard dot product, then the adjoint is controllable (Theorem 16). Thus, we just need to check that the rows of this matrix generate a reflexive chain group and that guarantees that the manifest behaviour of the adjoint system

$$\Phi(-\frac{d}{d\mathbf{x}},\frac{d}{d\mathbf{x}})\hat{w} \quad = \quad R^T(-\frac{d}{d\mathbf{x}})\hat{\ell}$$

is indeed $\mathfrak{B}^{\perp_\Phi}$. Note that in the the case of taking Φ to be the identity matrix, the last equation actually defines the image representation of the dual behaviour.

In this case, it is not clear that if we start with $\mathfrak{B} \in \mathfrak{L}^{\mathtt{w}}_{n,\mathrm{obs}}$ (i.e., $\mathfrak{B}$ is controllable and has an observable image representation), and the rows of the matrix $\begin{bmatrix} -\Phi^T(-\xi,\xi) & R^T(\xi) \end{bmatrix}$ generate a reflexive chain group with respect to the standard dot product, then $\mathfrak{B}^{\perp_\Phi} \in \mathfrak{L}^{\mathtt{w}}_{n,\mathrm{obs}}$ (i.e., $\mathfrak{B}^{\perp_\Phi}$ is also controllable and has an observable image representation). But, invoking Proposition 18, we can add that if the matrix $\begin{bmatrix} -\Phi^T(-\xi,\xi) & R^T(\xi) \end{bmatrix}$ has full row rank for all $\xi \in \mathbb{C}^n$, then the behaviour $\mathfrak{B}^{\perp_\Phi} \in \mathfrak{L}^{\mathtt{w}}_{n,\mathrm{obs}}$. The simplification for 1-D case is obvious. Moreover, in the case of calculating duals, we observe that the matrix concerned is $\begin{bmatrix} -I & R^T(\xi) \end{bmatrix}$ which has full row rank for all $\xi \in \mathbb{C}^n$ and hence we have the fact that if $\mathfrak{B} \in \mathfrak{L}^{\mathtt{w}}_{n,\mathrm{obs}}$, then $\mathfrak{B}^{\perp} \in \mathfrak{L}^{\mathtt{w}}_{n,\mathrm{obs}}$.

Thus, in this paper we have proposed a method for constructing adjoints of behavioural systems. We have shown that if we start with a controllable behaviour, we obtain a controllable behaviour as the adjoint behaviour (for both 1-D and n-D systems) under certain conditions (Theorem 16). We also showed that under more stricter conditions (Proposition 18), we obtain that the adjoint of a system which is controllable and observable is itself controllable and observable. These results are the most general kind of duality results. We pointed out the differences in the 1-D and n-D ($n > 1$) cases and explained why one cannot expect controllability and observability to be duals in the n-D behavioural case. We also gave several examples of applications for this construction of adjoints and explained why some well known conditions are attached to known problems of obtaining duals in 1-D systems theory.

References

[1] L. Hörmander. *The Analysis of Linear Partial Differential Operators, I & II.* Grundlehren der mathematischen Wissenschaften 256 & 257. Springer Verlag, 1983.

[2] H. Narayanan. On the duality between controllability and observability in behavioural systems theory. In *Proc. International Conference on Communications Control and Signal Proc. CCSP 2000*, Bangalore, July 25-July 28,2000, pp 183-186.

[3] H. Narayanan. Matroids representable over Modules,Electrical Network Topology and Behavioural Systems Theory In *Research Report EE/HN/2/2000, EE Dept. IIT Bombay*, October 2000.

[4] H. Narayanan. *Submodular functions and electrical networks*, volume 54 of *Annals of Discrete Maths.* North Holland, 1997.

[5] U. Oberst. Multidimensional constant linear systems. *Acta Applicandae Mathematicae*, 20:1–175, 1990.

[6] V. P. Palamodov. *Linear Differential Operators with Constant Coefficients*, volume 168 of *Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen.* Springer Verlag, 1970.

[7] H. K. Pillai and Shiva Shankar. A behavioural approach to control of distributed systems. *SIAM Journal of Control and Optimization*, 37:388–408, 1999.

[8] H. K. Pillai and J. C. Willems. The behavioural approach to distributed systems. In *CDC*, 1999.

[9] Harish K. Pillai and Eric Rogers. On bilinear and quadratic differential forms for n-d systems. Preprint.

[10] Harish K. Pillai and J. C. Willems. Lossless and dissipative distributed systems. submitted to SIAM J. of Control and Optimization.

[11] J. W. Polderman and J. C. Willems. *Introduction to Mathematical Systems Theory: A Behavioural Approach.* Springer Verlag, 1998.

[12] W.T.Tutte. Lectures on matroids. *Journal of Research of the National Bureau of Standards*, B69:1–48,1965.

[13] J. C. Willems. Paradigms and puzzles in the theory of dynamical systems. *IEEE Transactions on Automatic Control*, 36:259–294, 1991.

[14] J. C. Willems and H. L. Trentelman. On quadratic differential forms. *SIAM Journal of Control and Optimization*, 36(5):1703–1749, 1998.

Appendix 1

The proof of Theorem 4 requires the following lemma.

Lemma 20 *Let A be an $m \times n$ matrix over $R_\varepsilon[\xi_1, \cdots, \xi_n]$, whose column module is reflexive, b an $m \times 1$ vector over $R_\varepsilon[\xi_1, \cdots, \xi_n]$. Then $Ax = b$ has a solution (with x a vector over $R_\varepsilon[\xi_1, \cdots, \xi_n]$) iff for every $(1 \times m)$ vector λ^T over $R_\varepsilon[\xi_1, \cdots, \xi_n]$ we have $\lambda^T A = 0 \Rightarrow \lambda^T b = 0$.*

Proof of Theorem 4

Let $\mathcal{K}_{WL}$ have the representative matrix $[A_W \quad A_L]$ and let $\mathcal{K}_L$ have the representative matrix $[\hat{A}_L]$. Then $\mathcal{K}_{WL} + \mathcal{K}_L$ is generated by the rows of the representative matrix

$$\begin{bmatrix} A_W & A_L \\ 0 & \hat{A}_L \end{bmatrix}.$$

Further we have $\mathcal{K}_{WL} + \mathcal{K}_L$ reflexive by hypothesis. Now $x_W \in \mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L$ iff $\exists \lambda_1, \lambda_2$ over $R_\varepsilon[\xi_1, \cdots, \xi_n]$ s.t.

$$\begin{bmatrix} {A_W}^T & 0 \\ {A_L}^T & \hat{A}_L^T \end{bmatrix} \begin{matrix} \lambda_1 \\ \lambda_2 \end{matrix} = \begin{pmatrix} x_W \\ 0 \end{pmatrix}$$

i.e., (by Lemma 20, since column module of the above matrix is reflexive unimodular) iff whenever, for y_W, y_L over $R_\varepsilon[\xi_1, \cdots, \xi_n]$ we have

$$\begin{bmatrix} y_W^T & y_L^T \end{bmatrix} \begin{bmatrix} A_W^T & 0 \\ A_L^T & \hat{A}_L^T \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

we also have

$$y_W^T x_W = 0;$$

i.e., iff whenever $(y_W^T \; y_p^T) \in \mathcal{K}_{WL}^*, \;\; y_L^T \in \mathcal{K}_L^*$ we also have $y_W^T x_W = 0$;
i.e., iff whenever $y_W \in \mathcal{K}_{WL}^* \leftrightarrow \mathcal{K}_L^*$ we have ${y_W}^T x_W = 0$.
Thus, $x_W \in \mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L$ iff $x_W \in (\mathcal{K}_{WL}^* \leftrightarrow \mathcal{K}_L^*)^*$. This proves (a).

The statement in (b) follows from (a) and the definition of reflexivity .

(c) Suppose $\mathcal{K}_{WL}^* + \mathcal{K}_L^*$ is reflexive. We can now use (a) above and conclude that

$$(\mathcal{K}_{WL}^{**} \leftrightarrow \mathcal{K}_L^{**})^* = \mathcal{K}_{WL}^* \leftrightarrow \mathcal{K}_L^*$$

But $\mathcal{K}_{WL}, \mathcal{K}_L$ are also given to be reflexive and so $\mathcal{K}_{WL}^{**} = \mathcal{K}_{WL}$ and $\mathcal{K}_L^{**} = \mathcal{K}_L$. Hence $(\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L)^* = \mathcal{K}_{WL}^* \leftrightarrow {\mathcal{K}_L}^*$.

Statement (d) follows from (c) and the definition of reflexivity.

□

Proof of Theorem 6

We first restate Theorem 8 as the following lemma.

Lemma 21 *The equation*

$$[A[\xi]]w = f(\cdot), \tag{10}$$

$f(\cdot)$ a vector with $\mathcal{C}^\infty$ functions as entries, has a solution which is a vector with $\mathcal{C}^\infty$ functions as entries iff whenever

$$\lambda^T[\xi]A[\xi] = 0,$$

we also have

$$[\lambda^T[\xi], f(\cdot)] = 0. \tag{11}$$

Proof of Theorem 6

We use the following notation: If $\mathfrak{B} = (\mathcal{K})^\star$, we write $\mathfrak{B} = \mathfrak{B}(\mathcal{K})$ or say that $\mathfrak{B}$ is associated with chain group (or module) $\mathcal{K}$.

(a) Let $\mathcal{K}_{WL}$ have the generator matrix $[A_W \quad A_L]$ and let $\mathcal{K}_L$ have the generator matrix $[\hat{A}_L]$. Then $\mathcal{K}_{WL} + \mathcal{K}_L$ has the generator matrix

$$\begin{bmatrix} A_W & A_L \\ 0 & \hat{A}_L \end{bmatrix}.$$

Let $\mathfrak{B}_+$ be associated with the module $\mathcal{K}_{WL} + \mathcal{K}_L$. Clearly $\mathfrak{B}(\mathcal{K}_{WL}) \leftrightarrow \mathfrak{B}(\mathcal{K}_L) = \mathfrak{B}_+/w_W$. In other words, $w'_W \in \mathfrak{B}(\mathcal{K}_{WL}) \leftrightarrow \mathfrak{B}(\mathcal{K}_L)$ iff $\exists w'_L$, s.t.

$$\begin{bmatrix} A_W & A_L \\ 0 & \hat{A}_L \end{bmatrix} \begin{matrix} w'_W \\ w'_L \end{matrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

i.e., iff $\exists w'_L$, s.t.

$$\begin{bmatrix} A_L \\ \hat{A}_L \end{bmatrix} w'_L = \begin{pmatrix} -A_W w'_W \\ 0 \end{pmatrix}$$

i.e., (by Lemma 21) iff whenever

$$\begin{bmatrix} \lambda_1[\xi] & \lambda_2[\xi] \end{bmatrix} \begin{bmatrix} A_L \\ \hat{A}_L \end{bmatrix} = \begin{pmatrix} 0_L \end{pmatrix}$$

(equivalently whenever $\lambda_1[\xi]A_W \in \mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L$),
we also have $[\lambda_1[\xi], A_W w'_W] = 0$ (equivalently $[\lambda_1[\xi]A_W, w'_W] = 0$),
i.e., iff whenever $\lambda_1[\xi]A_W \in \mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L$, we also have $[\lambda_1[\xi]A_W, w'_W] = 0$,
i.e., iff $f_W \in \mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L$ implies $[f_W, w'_W] = 0$ (since every vector in $\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L$ is a linear combination of the rows of A_W). Thus $\mathfrak{B}(\mathcal{K}_{WL}) \leftrightarrow \mathfrak{B}(\mathcal{K}_L) = \mathfrak{B}(\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L)$.

(b)We use Theorem 5. We have
$(\mathfrak{B}(\mathcal{K}_{WL}) \leftrightarrow \mathfrak{B}(\mathcal{K}_L))^\star = (\mathfrak{B}(\mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L))^\star = \mathcal{K}_{WL} \leftrightarrow \mathcal{K}_L$, as required.

□

Appendix 2

For the sake of completeness, we give a brief introduction to the theory of bilinear and quadratic differential forms. For more details, please see [9]. For convenience, let ζ denote $(\zeta_1, \ldots, \zeta_n)$ and η denote $(\eta_1, \ldots, \eta_n)$. Let $\mathbb{R}^{q_1 \times q_2}[\zeta, \eta]$ denote the set of real polynomial matrices in the (commutating) $2n$ indeterminates ζ and η. We will consider bilinear forms of the type $\Phi \in \mathbb{R}^{q_1 \times q_2}[\zeta, \eta]$. Explicitly,

$$\Phi(\zeta, \eta) = \sum_{\mathbf{k},\mathbf{l}} \Phi_{\mathbf{k},\mathbf{l}} \zeta^{\mathbf{k}} \eta^{\mathbf{l}}$$

The sum above ranges over all non-negative multi-indices $\mathbf{k}, \mathbf{l} \in \mathbb{N}^n$ and is assumed to be finite. Moreover, $\Phi_{\mathbf{k},\mathbf{l}} \in \mathbb{R}^{q_1 \times q_2}$. Note that ζ denotes differentiation of terms to the left and η refers to differentiation of the terms to the right. Hence Φ induces a *bilinear differential form* (BLDF), that is, the map

$$L_\Phi : \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{q_1}) \times \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{q_2}) \to \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$$

defined by

$$L_\Phi(v, w)(\mathbf{x}) := \sum_{\mathbf{k},\mathbf{l}} \left(\frac{d^{\mathbf{k}} v}{d\mathbf{x}^{\mathbf{k}}}(\mathbf{x}) \right)^T \Phi_{\mathbf{k},\mathbf{l}} \left(\frac{d^{\mathbf{l}} w}{d\mathbf{x}^{\mathbf{l}}}(\mathbf{x}) \right)$$

If $q_1 = q_2 = q$ then Φ induces a *quadratic differential form* (QDF)

$$Q_\Phi : \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^q) \to \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R})$$

defined by

$$Q_\Phi(w) := L_\Phi(w, w)$$

Define the ^{F} operator

$$^F : \mathbb{R}^{q_1 \times q_2}[\zeta, \eta] \to \mathbb{R}^{q_2 \times q_1}[\zeta, \eta]$$

by

$$\Phi^F(\zeta, \eta) := \Phi^T(\eta, \zeta)$$

where ^{T} denotes matrix transposition. Clearly $L_\Phi(v, w) = L_{\Phi^F}(w, v)$. If $\Phi \in \mathbb{R}^{q \times q}[\zeta, \eta]$ satisfies $\Phi = \Phi^F$, then Φ is called *symmetric*. For the purposes of QDF's induced by polynomial matrices, it is enough to only consider the symmetric quadratic differential forms, since $Q_\Phi = Q_{\Phi^F} = Q_{\frac{1}{2}(\Phi + \Phi^F)}$.