On Observers and Behaviors

Paul A. Fuhrmann
Department of Mathematics
Ben Gurion University
Beer Sheva, Israel

Abstract

In this paper, we explore the connection between the classic, input output based theory of observers for linear functions of the state and the theory of behavior based observers as developed in the paper Valcher and Willems [1999].

1 Introduction

Lately, see Rosenthal, Schumacher and York [1996] and Rosenthal [2000], the applicabilty of system theoretic ideas and in particular behavioral theory to the study of convolutional codes has been pointed out. Since error correction decoding is closely related to tracking and filtering, however in an algebraic context and with a nonstandard, Hamming, metric. These topics are directly related to the theory of observers, whether in the state space or behavioral context. It is hoped that a better understanding of this area of system theory will lead itself to application in the area of convolutional codes.

The theory of observers for linear systems dates back to the early years of the development of modern control theory, see Luenberger [1966]. Althought it can be rightly argued that the process of observation is the basis of any control system, it has generally received less attention than the control part of the theory of linear systems. This is changing lately with new additions to the literature on this subject. In this connection we point out Fuhrmann and Helmke [2001] and the thesis by J. Trumpf [2002]. In these papers a comprehensive analysis is undertaken and observers are studied via the geometry of conditioned invariant and related subspaces. In parallel with this, and as part of the development of behavioral theory as initiated by J.C. Willems, see Willems [1989,1991], an important work on observers in the behavioral context is presented in Valcher and Willems [2002]. While much has been said about the advantages of the behavioral framework for the study of interconnections of systems, it seems to this author that it has also distinct advantages as far as observers are concerned. This is due to the fact that elimination of latent variables is a basic technique in behaviors and this allows the reduction of a complicated system to a potentially simpler one, involving only the variables that are observed or are to be estimated.

This paper will outline how starting from classical observer theory, we can focus on a natural behavior homomorphism, introduced in Fuhrmann [2002], from the full state behavior to the full observer behavior. We will use elimination theory to recast the problem in behavioral terms. We make contact with the important work of Antoulas [1983] that gives a parametrization of the set of asymtotic observers and this in turn leads to the results of Valcher and Willems.

Observers come in different forms and contexts. They depend also on the assumptions on the system in terms of what the time set is (discrete or continuous, full or semi-axis), what are the input and output function spaces, and the degree of observability, detectability we assume. To simplify matters, we will deal solely with asymptotic observers for discrete time systems. For a full exposition, we refer to a forthcoming paper, Fuhrmann [2003].

2 Observers

Definition 2.1. *Given a linear system*

$$\begin{cases} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \\ z(t) &= Kx(t) \end{cases} \tag{2.1}$$

with A, B, C, K *in* $\mathbf{R}^{n\times n}, \mathbf{R}^{n\times m}, \mathbf{R}^{p\times n}, \mathbf{R}^{l\times n}$ *respectively. We shall assume that* C, K *are of full row rank. The system*

$$\begin{cases} \xi(t+1) &= F\xi(t) + Gy(t) + Hu(t) \\ \zeta(t) &= J\xi(t) \end{cases} \tag{2.2}$$

with F, G, H, J *in* $\mathbf{R}^{q\times q}, \mathbf{R}^{q\times p}, \mathbf{R}^{q\times m}, \mathbf{R}^{k\times q}$*, respectively, with* J *of full row rank, and driven by the input* u *and output* y *of (2.1), will be called a an* **asymptotic observer** *for* K *if for all initial conditions of the state,* $\lim_{t\to\infty}(z(t) - \zeta(t)) = 0$*.*

Clearly, the existence of an observer is invariant under a state space similarity. Also, an output injection map would change appropriately the observer. Thus we may assume, without loss of generality, that the pair (C, A) is in dual Brunovsky form. The map K leads directly to a nice sequence of $k \times \lambda_\nu$ matrices $\{K^{(\nu)}\}_{\nu=1}^{\mu_1}$. The derivation of this sequence is somewhat technical and the reader is referred to Fuhrmann and Helmke [2001] for the details. We will denote by Γ_K the set of all strictly proper rational functions $\sum_{\nu=1}^{\infty} L^{(\nu)} z^{-\nu}$ satisfying

$$L_{ij}^{(\nu)} = K_{ij}^{(\nu)}, \quad i = 1, \ldots, k;\ j = 1, \ldots, \lambda_\nu. \tag{2.3}$$

The following theorem, summarizing a lot of previous work on observers, is taken from Fuhrmann and Helmke [2001] where more references can be found.

Theorem 2.1. *Given the linear system (2.1), then The following conditions are equivalent:*

1. *There exists an order* q *asymptotic observer for* K*.*

2. *There exist linear transformations, Z, F, G, H, J, with Z surjective of rank q and F stable, such that*

$$\begin{cases} ZA - FZ &= GC \\ H &= ZB \\ K &= JZ \end{cases} \tag{2.4}$$

holds.

3. *There exist proper stable rational functions M, N that solve*

$$\begin{pmatrix} M & N \end{pmatrix} \begin{pmatrix} zI - A \\ C \end{pmatrix} = K. \tag{2.5}$$

4. *Define $Z_K(z) = K(zI - A)^{-1}B$ and $Z_C(z) = C(zI - A)^{-1}B$. There exist strictly proper stable rational functions Z_1, Z_2, with $\delta \begin{pmatrix} Z_1 & Z_2 \end{pmatrix} = q$, that solve*

$$Z_K = Z_1 Z_C + Z_2. \tag{2.6}$$

5. *There exists a codimension q, outer detectable subspace $\mathcal{V} \subset \mathcal{X}$ satisfying*

$$\mathcal{V} \subset \operatorname{Ker} K. \tag{2.7}$$

6. *The nice sequence of matrices, $K^{(1)}, \dots, K^{(\mu_1)}$, has a McMillan degree q stable partial realization.*

We note that solvability of (2.6) by strictly proper stable rational functions is equivalent to the inclusion

$$\operatorname{Ker} H_{Z_K} \supset \operatorname{Ker} H_{Z_C}, \tag{2.8}$$

where $H_{Z_K} : H^2_+ \longrightarrow H^2_-$ and $H_{Z_C} : H^2_+ \longrightarrow H^2_-$, are appropriate Hankel operators. In the sequel we will assume that the pair $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A \right)$ is observable. That this entails no loss of generality is pointed out in Antoulas [1983].

In the context of this paper, a behavior $\mathcal{B}$ is a linear, shift invariant and complete subspace of $z^{-1}F^m[[z^{-1}]]$. Behaviors have been introduced by J.C. Willems, see Willems [1989,1991] and the further references therein. Fuhrmann [2002] contains a derivation in the spirit of the present paper. Using the prevailing notation in behavioral theory for the backward shift operator, namely σ, the system equations (2.1) can be written in behavioral AR form as

$$\mathcal{B}_{fsys} = \operatorname{Ker} \begin{pmatrix} \sigma I - A & -B & 0 & 0 \\ C & 0 & -I & 0 \\ K & 0 & 0 & -I \end{pmatrix} \tag{2.9}$$

Similarly, the full behavior of the observer (2.2) is given by

$$\mathcal{B}_{fob} = \text{Ker} \begin{pmatrix} \sigma I - F & -G & -H & 0 \\ J & 0 & 0 & -I \end{pmatrix} \tag{2.10}$$

In order to go into a polynomial setting, the data on the system needs to be appropriately encoded in polynomial terms. The key to this is the following left coprime factorization

$$\begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix}^{-1} \begin{pmatrix} \Theta_1 \\ \Theta_2 \end{pmatrix} = \begin{pmatrix} C \\ K \end{pmatrix} (zI - A)^{-1} \tag{2.11}$$

which has the following additional properties:
(i) $D_{11}(z)^{-1}\Theta_1(z)$ is a left coprime factorization of $C(zI - A)^{-1}$.
(ii) D_{11} is row proper.
(iii) $D_{21}D_{11}^{-1}$ is strictly proper.
The special form of the left coprime factorization in (2.11) is derived using the fact that a left coprime factorization is only determined up to a common left unimodular factor and this freedom, used judiciously, leads to this form. Moreover, we have $n = \deg\det(zI - A) = \deg\det D_{11} + \deg\det D_{22}$ and that D_{22} is a nonsingular polynomial matrix. It is a nonsingular constant matrix if and only if the pair (C, A) is observable. It is a stable matrix if and only if the pair (C, A) is detectable. We can use the left coprime factorization to eliminate the state variable from (2.9), which leads to the manifest behavior representation

$$\mathcal{B}_{sys} = \text{Ker} \begin{pmatrix} -\Theta_1 B & D_{11}(\sigma) & 0 \\ -\Theta_2 B & D_{21}(\sigma) & D_{22}(\sigma) \end{pmatrix} \tag{2.12}$$

In a completely analogous manner, with $Q^{-1}\Pi$ a left coprime factorization of $J(zI - F)^{-1}$, and defining $P = \Pi G$ and $R = \Pi H$, then the manifest observer behavior, after elimination of the state variable ξ, is given by

$$\mathcal{B}_{ob} = \text{Ker} \begin{pmatrix} -\Pi(\sigma)G & -\Pi(\sigma)H & Q(\sigma) \end{pmatrix} = \text{Ker} \begin{pmatrix} -P(\sigma) & -R(\sigma) & Q(\sigma) \end{pmatrix}. \tag{2.13}$$

Using the left coprime factorization (2.11), equation (2.6) is easily seen to be solvable with the, not necessarily proper, rational functions Z_1, Z_2 given by

$$\begin{aligned} Z_1 &= -D_{22}^{-1} D_{21} \\ Z_2 &= \quad D_{22}^{-1}\Theta_2. \end{aligned} \tag{2.14}$$

If (C, A) is observable then $D_{22} = I$ and a polynomial solution of equation (2.6) is given by $Z_K = -D_{21}Z_C + \Theta_2$. The set of all rational solutions of the system (2.1), under the assumption that (C, A) is an observable pair, is given by

$$\begin{aligned} Z_1 &= -D_{21} - WD_{11} \\ Z_2 &= \quad \Theta_2 + W\Theta_1 \end{aligned} \tag{2.15}$$

with W an arbitrary rational function. However, for the construction of a causal observer, we need to find strictly proper solutions. It turns out that under the assumption that (C, A) is an observable pair, then Z_1, Z_2 is a strictly proper rational solution of equation (2.6) if and only if (2.15) holds with $W \in \Gamma_K$. We will refer to equation (2.15) as the **Antoulas parametrization**, see Antoulas [1983].

Next, we note that equation (2.4), characterizing observers can be rewritten in matrix form as

$$\begin{aligned} &\begin{pmatrix} Z & G & 0 \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} zI - A & -B & 0 & 0 \\ C & 0 & -I & 0 \\ K & 0 & 0 & -I \end{pmatrix} \\ &= \begin{pmatrix} zI - F & -H & -G & 0 \\ J & 0 & 0 & -I \end{pmatrix} \begin{pmatrix} Z & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \end{aligned} \tag{2.16}$$

In view of the results in Fuhrmann [2002], this means that the map $\tau : \mathcal{B}_{fsys} \longrightarrow \mathcal{B}_{fob}$, defined by

$$\tau \begin{pmatrix} x \\ u \\ y \\ z \end{pmatrix} = \begin{pmatrix} Zx \\ u \\ y \\ z \end{pmatrix} \tag{2.17}$$

is a behavior homomorphism, which by our assumption of the observability of the pair $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$, is necessarily injective. This means that $\operatorname{Im} \tau$ is a subbehavior of $\mathcal{B}_{fob}$, and after applying the elimination procedure to both behaviors, we obtain

$$\mathcal{B}_{sys} \subset \mathcal{B}_{ob}. \tag{2.18}$$

It is worth pointing out that the above can be interpreted as a manifestation of the internal model principle. In fact, it seems that the behaviorial context and in particular the use of behavior homomorphisms provide the right language in which to formulate general internal model principles.

Paraphrasing Valcher and Willems, intuitively it is clear that the effect of the input or control variable can be removed without affecting the solution of the observer problem. Thus we can without any loss of generality disregard the input variable.

Of course, behavior inclusion is related to a factorization. This means that a factorization exists, with X, Y polynomial matrices, of the form

$$\begin{pmatrix} -P(z) & Q(z) \end{pmatrix} = \begin{pmatrix} -X(z) & Y(z) \end{pmatrix} \begin{pmatrix} D_{11}(z) & 0 \\ D_{21}(z) & D_{22}(z) \end{pmatrix}. \tag{2.19}$$

At this point it remains to clarify the relation of this factorization to the standard theory. Let us assume $W \in \Gamma_K$ has the representation

$$W = Y^{-1}X = \overline{X}\overline{Y}^{-1} \tag{2.20}$$

with both factorizations coprime. Then we have the representations

$$\begin{array}{rcl} Z_1 & = & Y^{-1}(-YD_{21} + XD_{11}) \\ Z_2 & = & Y^{-1}(Y\Theta_2 - X\Theta_1) \end{array} \tag{2.21}$$

i.e. we have, for $Z_1 = Q^{-1}P$, with the polynomial matrices P, Q defined by the factorization (2.19). This representation of the behavioral observer matches Theorem 3.4 in Valcher and Willems [1999].

Theorem 2.2. *For the plant whose behavior is described by (2.12) with D_{22} stable, then (2.13) is an asymptotic observer for Σ if and only if there exists a nonsingular polynomial matrix Y and a polynomial matrix X such that (2.19) holds.*

Actually, our analysis goes a bit beyond this theorem in establishing the connection to classical observer theory and in particular the role of the partial realization problem in observer theory.

We have started with a state space based observer construction and by elimination have reduced it to a behavior based observer construction. Of course, this line of reasoning can be reversed. If one starts with a behavior based observer construction problem, one can use realization theory to transform the problem into the standard state space form. Once an observer is constructed, an elimination argument would lead to a behavior based solution.

Acknowledgement: Research supported in part by GIF under Grant No. I-526-034 and by Grant 235/01 of the Israel Science Foundation.

References

[1983] A.C. Antoulas, "New results on the algebraic theory of linear systems: the solution of the cover problem", *Lin. Alg. Appl.*, 50, 1-43.

[2002] P.A. Fuhrmann, A study of behaviors, to appear *Lin. Alg. Appl.*

[2003] P.A. Fuhrmann, "On behaviors and observer theory", in preparation.

[2001] P.A. Fuhrmann and U. Helmke, "On the parametrization of conditioned invariant subspaces and observer theory", *Lin. Alg. Appl.*, vols. 332-334, 265-353.

[1966] D.G. Luenberger, "Observers for multivariable systems", *IEEE Trans. Aut. Control*, AC-11, 190-197.

[2000] J. Rosenthal, "Connections between linear systems and convolutional coding theory", preprint.

[2000] J. Rosenthal, J.M. Schumacher and E.V. York, "On behaviors and convolutional codes", *IEEE Trans. Inf. Th.*, 42, 1881-1891.

[2002] J. Trumpf, "On the geometry and parametrization of almost invariant subspaces", Ph.D. thesis, University of Würzburg.

[1999] M.E. Valcher and J.C. Willems, "Observer synthesis in the behavioral framework", *IEEE Trans. Aut. Contr.*, 44, 2297-2307.

[1989] J.C. Willems, "Models for dynamics", *Dynamics Reported*, vol. 2, 171-269, U. Kirchgraber and H.O. Walther (eds.), Wiley-Teubner.

[1991] J.C. Willems, "Paradigms and puzzles in the theory of dynamical systems", *IEEE Trans. Autom. Contr.* AC-36, 259-294.