A Unified Theory of Full-Order and Low-Order Observers Based on Singular System Theory

J. Daniel Cobb

Abstract—**The standard construction of the minimal-order state observer for a linear time-invariant plant is shown to be a degenerate form of the ordinary full-order observer. The naive state estimation technique of successively differentiating the plant output is also shown to be a limiting case of the same full-order structure. Various results concerning the set of all degenerate forms of the full-order observer are presented. In particular, conditions are established under which shifting observer eigenvalues far to the left in the complex plane implies proximity to a successively differentiating system. In addition, it is shown that such a scheme can never yield satisfactory estimation. Singular perturbation and singular system theory are used throughout the analysis.**

I. Introduction

The theory of state observers has long been one of the cornerstones of modern system theory, with applications to a variety of control and filtering problems. Given the linear time-invariant plant

$$\dot{x} = Ax, \quad y = Cx \tag{1}$$

probably the simplest structure for estimating the state x is the full-order observer

$$\dot{z} = (A - LC)z + Ly. \tag{2}$$

Here we have ignored the possibility of external input terms in (2), since we assume that their influence on x is known exactly.

In addition to (2), other observer structures have been proposed. For example, reduced-order observers which exploit the direct transmission of state information through the C matrix, are well known. Also, it has been suggested that successive differentiation of y leads to a construction that yields the state variable exactly, although this method is generally dismissed as being "sensitive to noise." For an elementary discussion of these ideas, see [1, Section 7-4].

In this paper our objective is to demonstrate how these concepts can be unified through application of singular perturbation and singular system techniques. Specifically, we show that minimal-order and successive differentiation observers are simply degenerate forms of the full-order observer (2) obtained by letting L diverge. It will be seen that a number of striking results concerning the limiting forms of (2) are made possible by our approach.

II. Preliminaries

To speak in precise terms about the degenerate forms of (2), we will need some elementary concepts from singular system theory. Consider the singular differential equation

$$F\dot{z} = Gz + Hy \tag{3}$$

where F and G are $n \times n$, (F, G) is regular (i.e., $\det(sF - G) \not\equiv 0$), and H is $n \times p$. We could associate with each (3) a point $(F, G, H) \in \mathbb{R}^{n(2n+p)}$, but this would introduce redundancy in the parameterization, since premultiplication of (3) by a nonsingular matrix has no effect on the dynamics of the system. Thus it is convenient to adopt the differentiable manifold viewpoint developed in [2], where it is shown that systems (3) can be naturally associated with points in an $n(n+p)$-dimensional Grassman manifold.

Briefly, let $\Sigma(n, p) = \{(F, G, H) \in \mathbb{R}^{n(2n+p)} \mid (F, G) \text{ is regular}\}$, and define an equivalence on $\Sigma(n, p)$ by identifying points (F_1, G_1, H_1) and (F_2, G_2, H_2) whenever there exists a nonsingular matrix M such that $[F_1\ G_1\ H_1] = M[F_2\ G_2\ H_2]$. It is easily shown that the resulting quotient set $\mathcal{L}(n, p)$ is an open, dense submanifold of the Grassmanian $G_n(\mathbb{R}^{2n+p})$. Also, the ordinary state equations (viz. system (3) with $F = I$) imbed naturally into $\mathcal{L}(n, p)$ as an open, dense submanifold. The equivalence on $\Sigma(n, p)$ can in fact be extended to all of $\mathbb{R}^{n(2n+p)}$. By $[F, G, H]$ we mean the point in $G_n(\mathbb{R}^{2n+p})$ represented by (F, G, H); thus $[F_1, G_1, H_1] = [F_2, G_2, H_2]$ if and only if (F_1, G_1, H_1) and (F_2, G_2, H_2) are equivalent. The imbedding $\mu(F, G, H) = [F, G, H]$ which takes $\mathbb{R}^{n(2n+p)}$ into $G_n(\mathbb{R}^{2n+p})$ is a submersion; hence, if $\xi = [F, G, H]$ and $\xi_k \to \xi$, there exists a sequence $(F_k, G_k, H_k) \to (F, G, H)$ such that $\xi_k = [F_k, G_k, H_k]$ for all k.

It will be useful to invoke the Weierstrass Decomposition Theorem for regular pencils of matrices. (See [3].) According to the theorem, for each regular pencil (F, G), there exist nonsingular M and N such that

$$MFN = \operatorname{diag}(I, A_f), \quad MGN = \operatorname{diag}(A_s, I) \tag{4}$$

where A_f is nilpotent. Let $\Delta(s) = \det(sE - A)$ and $r = \deg \Delta$. Then A_f is $(n-r) \times (n-r)$, and we define $\operatorname{ind} A_f$ to be the smallest integer q such that $A_f^q = 0$. (If $n = r$, we set $\operatorname{ind} A_f = 0$.) For $\xi = [F, G, H] \in \mathcal{L}(n, p)$, we may then define the index and order of ξ by $\operatorname{ind}\xi = \operatorname{ind} A_f$ and $\operatorname{ord}\xi = r$, respectively. Note that the these two definitions depend only on ξ and not on the particular choice of representative (F, G, H). The eigenvalues of ξ are taken to be the eigenvalues of A_s or, equivalently, the roots of Δ. (Note that Δ is uniquely defined up to a scaling constant for a given ξ.)

In addition to the parametric representation of (3), we are also interested in its solutions. To this end, we need some basic concepts from the theory of distributions. (See e.g., [4].) Let K be the space of C^∞ functions $\varphi: \mathbb{R} \to \mathbb{R}$ with bounded support, and let K'_+ denote the distributions (continuous linear functionals on K) with support in $[0, \infty)$. Each locally L^1 function f is considered a distribution, since it determines a functional $\varphi \to \int f\varphi$. The unit impulse δ is defined to be the evaluation functional $\langle \delta, \varphi \rangle = \varphi(0)$. Every distribution has a derivative defined by $\langle \dot{f}, \varphi \rangle = -\langle f, \dot{\varphi} \rangle$; thus, $\langle \delta^{(i)}, \varphi \rangle = (-1)^i \varphi^{(i)}(0)$.

A sequence of distributions f_k is said to converge weak* to f if $\langle f_k, \varphi \rangle \to \langle f, \varphi \rangle$ for every $\varphi \in K$. Besides weak* convergence, we will sometimes refer to uniform convergence $f_k \to f$ on an interval $T \subset \mathbb{R}$. This simply means that there exist locally L_1 functions g_k, g defined on T such that $g_k \to g$ uniformly, $\langle f_k, \varphi \rangle = \langle g_k, \varphi \rangle$, and $\langle f, \varphi \rangle = \langle g, \varphi \rangle$ for all φ with support in T.

Let z_0 be the initial condition in (3). The response of singular systems to arbitrary initial conditions is a fairly subtle matter and is discussed at length in [5] and [6, Chapter 22]. Letting

$$[H_s^T, H_f^T]^T = MH, \quad [z_s^T\ z_f^T]^T = N^{-1}z, \quad [z_{0s}^T\ z_{0f}^T]^T = N^{-1}z_0 \tag{5}$$

we have

$$z_s = \exp(A_s)z_{0s}, \quad z_f = -\sum_{i=0}^{q-1} \delta^{(i-1)} A_f^i z_{0f} \tag{6}$$

Manuscript received June 24, 1993; revised January 28, 1994. This work was supported by NSF Grant ECS-8929981.

The author is with the Department of Electrical and Computer Engineering, University of Wisconsin-Madison, Madison, WI 53706-1691 USA.

IEEE Log Number 9405669.

where $\exp(X)(t) = e^{tX}\theta(t)$, where θ is the unit step function. Based on (6), we say that the system ξ is stable, if all its eigenvalues λ satisfy $\operatorname{Re}\lambda < 0$ and $\operatorname{ind}\xi \leq 1$. Finally, we define the natural response matrix corresponding to (3) as

$$\Phi = N \operatorname{diag}\left(\exp(A_s), -\sum_{i=0}^{q-1} \delta^{(i-1)} A_f^i\right) N^{-1}. \tag{7}$$

It is easily verified that Φ is the unique distribution with support in $[0, \infty)$ satisfying $F\Phi = G\Phi + \delta F$ and $z = \Phi z_0$. (See [6, Chapter 22].) Φ may be viewed as an ordinary function on $(0, \infty)$ or, if $\operatorname{ind}\xi \leq 1$, on $[0, \infty)$. Also, the mapping $(F, G) \rightarrow \Phi$ is invariant under multiplication of (3) by M, so the map $\xi \rightarrow \Phi$ from $\mathcal{L}(n, p)$ into $(K'_+)^{n\times n}$ is well defined.

III. The Manifold of Observers

The full-order observers (2) for a given plant (1) imbed naturally into $\mathcal{L}(n, p)$ via the map $L \rightarrow [I, A - LC, L]$. We denote the image of $\mathbb{R}^{n\times p}$ under this map by $\mathcal{O}_r$. Let $\mathcal{O}$ denote the closure of $\mathcal{O}_r$ in $\mathcal{L}(n, p)$, and set $\mathcal{O}_s = \mathcal{O} - \mathcal{O}_r$. $\mathcal{O}$ is the set of observers corresponding to (1), points in $\mathcal{O}_r$ are the regular observers, and elements of $\mathcal{O}_s$ are the singular observers.

Another way of defining the sets $\mathcal{O}$, $\mathcal{O}_r$, and $\mathcal{O}_s$ is through use of the submersion μ. Let $\Omega_r = \{(M, M(A - LC), ML) \mid \det M \neq 0, L \in \mathbb{R}^{n\times p}\}$, Ω be the closure of Ω_r in $\Sigma(n, p)$, and $\Omega_s = \Omega - \Omega_r$. It is routine to verify that $\mu(\Omega) = \mathcal{O}$, $\mu(\Omega_r) = \mathcal{O}_r$, and $\mu(\Omega_s) = \mathcal{O}_s$.

A great deal more can be said about the structure of the set of observers $\mathcal{O}$.

Theorem 3.1:

1) $\mathcal{O} = \{[X, XA - YC, Y] \in \mathcal{L}(n, p) \mid X \in \mathbb{R}^{n\times n}, Y \in \mathbb{R}^{n\times p}\}$
2) $\mathcal{O}$ is a regular submanifold of $\mathcal{L}(n, p)$ with dimension np.
3) $\mathcal{O}_r$ is a (relatively) open, dense submanifold of $\mathcal{O}$.
4) If $[X, XA - YC, Y] \in \mathcal{O}_s$, then X is a singular matrix.

Proof:

1) Let X and Y be given. Then there exists a nonsingular sequence $M_k \rightarrow X$; setting $L_k = M_k^{-1}Y$ yields $M_k L_k \rightarrow Y$ and $M_k(A - L_k C) \rightarrow XA - LC$. Thus the closure of Ω_r contains $\Omega = \{(X, XA - YC, Y) \in \Sigma(n, p) \mid X \in \mathbb{R}^{n\times n}, Y \in \mathbb{R}^{n\times p}\}$. But Ω is itself closed, so the desired result follows from $\mu(\Omega) = \mathcal{O}$.
2) It is routine to verify that the map $[X, Y] \rightarrow [X, XA - YC, Y]$ from $G_n(\mathbb{R}^{n+p})$ into $G_n(\mathbb{R}^{2n+p})$ has full rank; hence, $G_n(\mathbb{R}^{n+p})$ may be viewed as a submanifold with dimension np. The proof of regularity follows along the same lines as the proof of Theorem 2, part (2) in [2].
3) Density of $\mathcal{O}_r$ is obvious. Since Ω_r is open in Ω and μ is a submersion (and therefore an open map), $\mu(\Omega_r)$ is open in $\mu(\Omega)$.
4) Let $\xi = [X, XA - YC, Y]$. If X is nonsingular, $\xi = [I, A - X^{-1}YC, X^{-1}Y]$. Setting $L = X^{-1}Y$ implies that $\xi \in \mathcal{O}_s$. □

According to Theorem 3.1, part 4), the set $\mathcal{O}_s$ of singular observers corresponds to all degenerate forms of full-order observers (2). This reduction in order can only occur as a result of letting L diverge in such a way that some or all eigenvalues of (2) tend to infinity.

Another important subset of $\mathcal{O}$ is described as follows. Let $\mathcal{V} = \{(X, Y) \in \mathbb{R}^{n(n+p)} \mid XA - YC = I\}$. Clearly, $\mathcal{V}$ is nonempty iff $\operatorname{rank}[A^T\ C^T]^T = n$ (or, equivalently, zero is not an unobservable mode of (A, C)). In fact, Theorem 3.1, part 1) implies that $\xi = [X, I, Y] \in \mathcal{O}$ iff $(X, Y) \in \mathcal{V}$. Thus the map $(X, Y) \rightarrow [X, I, Y]$ puts $\mathcal{V}$ (diffeomorphically) into one-to-one correspondence with the set of observers $[F, G, H]$ with G nonsingular. Consequently, $\mathcal{V}$ may be viewed both as a submanifold of $\mathcal{O}$ or an affine subset of $\mathbb{R}^{n(n+p)}$. From (4), $\mathcal{V}$ is precisely the set of observers with no eigenvalue at the origin and therefore includes all stable and all zeroth-order observers. The stable and zeroth-order observers are the subject of the next three sections.

IV. Minimal Order Stable Observers

We begin this section by noting that there is no loss of generality in assuming $\operatorname{rank} C = p$, since otherwise we could redefine the output as $w = Ty$, where TC has independent rows. With this in mind, we are in a position to prove a variety of results concerning the stable points in $\mathcal{O}_s$.

Lemma 4.1: If (A, C) has an unobservable eigenvalue λ, then λ is an eigenvalue of every point in $\mathcal{O}$.

Proof: Since λ is an unobservable eigenvalue of (A, C), λ is also an eigenvalue of $A - LC$ for every matrix L and, hence, an eigenvalue of every point in $\mathcal{O}_r$. Let $\xi = [F, G, H] \in \mathcal{O}_s$, and choose $\xi_k \in \mathcal{O}_r$ such that $\xi_k \rightarrow \xi$. Then there exists $(F_k, G_k, H_k) \rightarrow (F, G, H)$ such that $\xi_k = [F_k, G_k, H_k]$. Since λ is a root of $\Delta_k(s) = \det(sF_k - A_k)$ for every k and since the roots of a polynomial of degree n are continuous relative to its coefficients, λ is also a root of Δ. Hence λ is an eigenvalue of ξ. □

Theorem 4.2:

1) If $\mathcal{O}_s$ contains a stable point, then (A, C) is detectable.
2) If (A, C) is detectable, then $\mathcal{O}_s$ contains a stable point ξ with $\operatorname{ord}\xi = n - p$.

Proof:

1) Suppose (A, C) is not detectable. Then it has an unobservable eigenvalue λ with $\operatorname{Re}\lambda \geq 0$. From Lemma 4.1, every $\xi \in \mathcal{O}$ has the same eigenvalue λ, so every ξ is unstable.
2) Choose a nonsingular matrix N such that $CN = [I\ 0]$, and let

$$\begin{bmatrix} P & Q \\ R & S \end{bmatrix} = N^{-1}AN. \tag{8}$$

Since (A, C) is detectable

$$\operatorname{rank}\begin{bmatrix} \lambda I - P & -Q \\ -R & \lambda I - S \\ I & 0 \end{bmatrix} = n$$

for every λ with $\operatorname{Re}\lambda \geq 0$. Hence, $\operatorname{rank}[\lambda I - S^T\ Q^T]^T = n - p$; i.e., (S, Q) is detectable. Choose Λ such that $S - \Lambda Q$ is stable, and let

$$\xi = \left[N\begin{bmatrix} 0 & 0 \\ -(S - \Lambda Q)^{-1}\Lambda & (S - \Lambda Q)^{-1} \end{bmatrix} N^{-1}, \right.$$
$$\left. I, N\begin{bmatrix} -I \\ (S - \Lambda Q)^{-1}(R - \Lambda P) \end{bmatrix}\right]. \tag{9}$$

Clearly, $\xi \in \mathcal{L}(n, p)$, and ξ is stable, since it has unit index and $(S - \Lambda Q)^{-1}$ is stable. That $\xi \in \mathcal{O}$ follows from

$$N\begin{bmatrix} 0 & 0 \\ -(S - \Lambda Q)^{-1}\Lambda & (S - \Lambda Q)^{-1} \end{bmatrix} N^{-1}A$$
$$- N\begin{bmatrix} -I \\ (S - \Lambda Q)^{-1}(R - \Lambda P) \end{bmatrix} C = I. \quad □$$

Theorem 4.3:

1) If $\xi \in \mathcal{O}_s$ is stable, then $\operatorname{ord}\xi \geq n - p$.
2) $\xi \in \mathcal{O}_s$ is stable with $\operatorname{ord}\xi = n - p$ iff (S, Q) is detectable and (9) holds.

Proof:

1) If ξ is stable, then $\xi = [X, I, Y]$ with $(X, Y) \in \mathcal{V}$. Since $XA - YC = I$, rank$[X \; -Y] = n$; hence, rank$X \geq n - p$. Appealing to (4), stability implies $A_f = 0$, so ordξ = rank$X$ and the result follows.
2) Premultiplication of (9) by $\overline{M} = \begin{bmatrix} -(S-\Lambda Q)\Lambda & S-\Lambda Q \\ I & \Lambda \end{bmatrix} N^{-1}$ and postmultiplication of the first two entries by $\overline{N} = N\begin{bmatrix} 0 & I \\ I & \Lambda \end{bmatrix}$ yields a decomposition (4), (5) $\overline{\sigma} = \left(\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} S-\Lambda Q & 0 \\ 0 & I \end{bmatrix}, \begin{bmatrix} (S-\Lambda Q)\Lambda + R - \Lambda P \\ -I \end{bmatrix}\right)$. Sufficiency follows immediately.

To prove necessity, we again note that $\xi = [X, I, Y]$ with $XA - YC = I$. Let $\begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix} = N^{-1}XN$ and $\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = N^{-1}Y$. This yields $\begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix}\begin{bmatrix} P & Q \\ R & S \end{bmatrix} - \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}[I \; 0] = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$; hence,

$$[X_3 \quad X_4][Q^T \quad S^T]^T = I \tag{10}$$

so rank $[X_3 \; X_4] = n - p$. As in part 1), rank X = ord$\xi = n - p$. Hence, there exists a matrix M such that $[X_1 \; X_2] = M[X_3 \; X_4]$. But $M = M[X_3 \; X_4]\begin{bmatrix} Q \\ S \end{bmatrix} = [X_1 \; X_2]\begin{bmatrix} Q \\ S \end{bmatrix} = 0$, $X = N\begin{bmatrix} 0 & 0 \\ X_3 & X_4 \end{bmatrix}N^{-1}$, and $\det X_4 \neq 0$. Let $\Lambda = -X_4^{-1}X_3$. Then, from (10), $X_4^{-1} = S - \Lambda Q$ and, since ξ is stable, X_4 is stable. Finally, we note that (10) gives $Y_1 = -I$ and $Y_2 = X_3P + X_4R = X_4(R - \Lambda P)$. □

We have thus established that the minimal order stable observers are parameterized by (9), where Λ ranges over all matrices that make $S - \Lambda Q$ stable. These systems are closely related to the conventional definition of minimal order dynamic observers found in elementary references such as [1, Section 7.4]. The exact connection can be made as follows: Every point $\xi \in \mathcal{O}_s$ characterized by (9) determines a differential equation

$$N\begin{bmatrix} 0 & 0 \\ -(S-\Lambda Q)^{-1}\Lambda & (S-\Lambda Q)^{-1} \end{bmatrix} N^{-1}\dot{z} = z + N\begin{bmatrix} -I \\ (S-\Lambda Q)^{-1}(R - \Lambda P) \end{bmatrix} y.$$

Let $\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} -\Lambda & I \\ I & 0 \end{bmatrix} N^{-1}z$. Solving for z, direct substitution, and premultiplying by $M = \begin{bmatrix} -(S-\Lambda Q)L & S-\Lambda Q \\ I & 0 \end{bmatrix} N^{-1}$ yields $\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \dot{w}_1 \\ \dot{w}_2 \end{bmatrix} = \begin{bmatrix} S-\Lambda Q & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \begin{bmatrix} (S-\Lambda Q)\Lambda + R - \Lambda P \\ -I \end{bmatrix} y$, which can be rewritten $\dot{w}_1 = (S - \Lambda Q)w_1 + ((S - \Lambda Q)\Lambda + R - \Lambda P)y$, $z = N[0 \; I]^T w_1 + N[I \; \Lambda^T]^T y$. The last two equations are easily recognized as the conventional form for the minimal order dynamic observer. (See [1, p. 362].) Thus all stable minimal order observers are merely limiting forms of the full-order observer (2).

V. Zeroth Order Observers

As indicated in (4), zeroth-order systems in $\mathcal{L}(n, p)$ are those of the form $\xi = [X, I, Y]$, where X is nilpotent. If such a system is an observer, it must lie in $\mathcal{O}_s$; indeed, a zeroth-order observer is simply a limiting form of the full-order structure (2), where we let L diverge in such a way that all eigenvalues of (2) tend to infinity. Zeroth-order observers must belong to $\mathcal{V}$. Denote the set of all such points by $\mathcal{Z}$.

If (A, C) is observable, a great deal of insight can be gained by exploiting a certain affine parameterization of $\mathcal{V}$. The parameterization is based on the Brunovsky Canonical Form of (A, C). (See [7, Section 7.1.3] for details.)

Let T, L, and G be matrices with T and G nonsingular such that $\overline{A} = T^{-1}(A - LGC)T = \text{diag}(A_1, \cdots, A_p)$ and $\overline{C} = GCT = \text{diag}(C_1, \cdots, C_p)$, where

$$A_i = \begin{bmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & & \\ & & 1 & 0 \end{bmatrix}, \quad C_i = [0 \quad \cdots \quad 0 \quad 1].$$

Assume A_i is $n_i \times n_i$. If

$$\overline{Y} = [Y_1^T \cdots Y_p^T], \quad Y_i = [y_{qr}^i] \tag{11}$$

then it follows by direct calculation that $\overline{XA} - \overline{YC} = I$ implies $\overline{X} = [X_{ij}]$, where

$$X_{ii} = \begin{bmatrix} y_{1i}^i & 1 & & \\ & 0 & \ddots & \\ \vdots & & \ddots & 1 \\ y_{n_i i}^i & & & 0 \end{bmatrix}; \quad X_{ij} = \begin{bmatrix} y_{1j}^i & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ y_{n_i j}^i & 0 & \cdots & 0 \end{bmatrix}, \quad i \neq j. \tag{12}$$

The map $\overline{Y} \to \overline{X}$ determined by (11) and (12) is affine and one-to-one. An affine parameterization $\overline{Y} \to (X, Y)$ of $\mathcal{V}$ is thus generated by reversing the Brunovsky transformation $X = T\overline{X}T^{-1}$, $Y = T(\overline{Y} + \overline{X}T^{-1}L)G$. We can now prove several results concerning the zeroth-order observers $\mathcal{Z}$.

Theorem 5.1:

1) $\mathcal{Z}$ is nonempty iff (A, C) is observable.
2) If (A, C) is observable and $p = 1$, $\mathcal{Z}$ is a singleton.
3) If (A, C) is observable and $p > 1$, $\mathcal{Z}$ is uncountable and unbounded (as a subset of $\mathbb{R}^{n(n+p)}$).

Proof:

1) If (A, C) is not observable, it has an unobservable eigenvalue λ. From Lemma 4.1, λ is an eigenvalue of every point in $\mathcal{O}$. But a zeroth-order point must have constant Δ. On the other hand, if (A, C) is observable, the parameterization (11), (12) yields nilpotent X for $\overline{Y} = 0$.
2) For $p = 1$, (11), (12) reduce to

$$\overline{X} = \begin{bmatrix} y_1 & 1 & & \\ & 0 & \ddots & \\ \vdots & & \ddots & 1 \\ y_n & & & 0 \end{bmatrix}, \quad \overline{Y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}. \tag{13}$$

If $\overline{X}$ is nilpotent, $\overline{Y} = 0$.

3) For $p > 1$, nilpotent X may be achieved in a variety of ways. For example

$$\overline{Y} = \begin{bmatrix} 0 & y_{12} & \cdots & y_{1p} \\ \vdots & \ddots & \ddots & \vdots \\ & & & y_{p-1,p} \\ 0 & \cdots & & 0 \end{bmatrix}$$

yields nilpotent X regardless of the values $y_{ij} \in \mathbb{R}^{n_i}$. The desired result follows from the Brunovsky transformation. □

Next we investigate the dynamic properties of zeroth-order points.

Theorem 5.2: Every $\xi \in \mathcal{Z}$ satisfies ind$\xi \geq (n/p)$.

Proof: As noted at the end of Section III, $\xi = [X, I, Y]$, where $XA - YC = I$. Thus rank$X \geq n - p$; also ord$\xi = 0$ implies that X is nilpotent. The Jordan form of X must therefore consist of no more than p blocks. Clearly, ind$X$ is equal to the dimension of the largest block, which must be at least (n/p). □

Theorem 5.3: If $p = 1$, $\xi_k \in \mathcal{O}_r$, and all eigenvalues of ξ_k diverge, then $\xi_k \to \xi$, where ξ is the unique zeroth-order point in $\mathcal{O}_s$.

Proof: Since the eigenvalues of ξ_k diverge, $\xi_k \in \mathcal{V}$ and (A, C) must be observable. We may therefore apply the parameterization (11), (12) yielding $\overline{X}_k$, $\overline{Y}_k$ of the form (13). The eigenvalues of $\overline{X}_k$ must tend to zero, so $\overline{Y}_k \to 0$ and $\overline{X}_k$ converges to a nilpotent. Applying the Brunovsky transformation, $\xi_k \to \xi$, where $\operatorname{ord}\xi = 0$. □

Unfortunately, Theorem 5.3 does not generalize to cases where $p > 1$. From Theorem 5.1, part 3), we may construct a sequence $(X_k, Y_k) \in \mathcal{V}$ where X_k is nonsingular and has eigenvalues converging to zero, but is unbounded. Hence $\xi_k = [X_k, I, Y_k] \in \mathcal{O}_r$ and has diverging eigenvalues, but does not converge in $\mathcal{O}$.

We are now in a position to compare the zeroth-order observers described within our framework to the "successive differentiation" scheme which is often discussed in elementary system theory texts. (see, e.g., [1, Problem 5.17].) Through repeated differentiation of (1) and substitution for $\dot{x}$ from the state equation, for $t > 0$ we obtain

$$[y^T(t) \quad \dot{y}^T(t) \quad \cdots \quad y^{(n-1)T}(t)]^T = Vx(t) \tag{14}$$

where V is the observability matrix of (1). Thus the left side of (14) is guaranteed to lie in the image space of V for all $t > 0$. Observability of (1) implies that V has full rank, so we may premultiply (14) by any left inverse W of V to obtain $x(t)$ explicitly. Hence, the proposed "observer" is

$$z(t) = W[y^T(t) \quad \dot{y}^T(t) \quad \cdots \quad y^{(n-1)T}(t)]^T. \tag{15}$$

This scheme appears ideal in that an exact copy of $x(t)$ is apparently obtained; however, differentiation is known to be "sensitive to noise," so the scheme is not considered viable by most system engineers and theorists. Interpreting (15) within our framework, the essential problem is that (14) is incomplete; indeed, the derivatives of y in (14) are not well defined about the origin, unless some information about y is given for $t < 0$. For example, one might imagine a small disturbance added to the plant output y just prior to $t = 0$, vanishing discontinuously at $t = 0$, and thus producing impulses in $\dot{y}$, $\ddot{y}$, etc. Such behavior is not accounted for by (14). In general, if output disturbances (or even plant input disturbances) are present, (14) does not give a complete picture of the behavior of the state x.

Our singular system framework does, however, enable us to achieve an understanding of transient phenomena about $t = 0$ in the successive differentiation approach. We first note that a zeroth-order observer $\xi = [X, I, Y]$ corresponds to a differential equation $X\dot{z} = z + Yy$, where X is nilpotent. Equation (6) indicates that the state estimate z is given by

$$z = -\sum_{i=0}^{n-1} X^i Y y^{(i)} - \sum_{i=1}^{n-1} \delta^{(i-1)} X^i z_0. \tag{16}$$

Note that (16) does indeed involve derivatives of y. The initial condition may be the result of disturbances prior to $t = 0$, so the impulsive terms in (16) may be viewed as the "noisy" part of z. For $t > 0$ we have

$$z(t) = U[y^T(t) \quad \dot{y}^T(t) \quad \cdots \quad y^{(n-1)T}(t)]^T \tag{17}$$

where $U = -[Y \ XY \ \cdots \ X^{n-1}Y]$. Comparing (15) and (17), it remains to show that U is a left inverse of V.

Lemma 5.4: If X is nilpotent and $XA - YC = I$, then $UV = I$.

Proof:

$$\begin{aligned} UV &= -\sum_{i=0}^{n-1} X^i Y C A^i = \sum_{i=0}^{n-1} X^i (I - XA) A^i \\ &= \sum_{i=0}^{n-1} X^i A^i - \sum_{=1}^{n} X^i A^i = I - X^n A^n = I. \end{aligned}$$ □

We have thus established that every zeroth-order point in $\mathcal{O}$ is an implementation of the successive differentiation technique. Indeed, since the left side of (14) is guaranteed to lie in the image space of V, the exact choice of left inverse W in (15) has no effect on the state estimate for $t > 0$. Hence, setting $W = U$ yields the same estimate for $t > 0$, independent of X and Y. In fact, the only change in (16) caused by varying (X, Y) over $\mathcal{Z}$ is the structure of the impulsive part of the response. Unfortunately, in view of Theorem 5.2, no choice of X and Y can remove the impulses altogether, except when $p = n$.

VI. Regularization and Error Convergence

One advantage of our theory is that the error vector $e = z - x$ is governed by a single equation, regardless of which of the observer structures studied above is chosen. To derive this equation, recall Theorem 3.1, part 1), and note that $X\dot{x} = XAx$, $X\dot{z} = (XA - YC)z + YCx$. Hence, $X\dot{e} = (XA - YC)e$. In particular, if $(X, Y) \in \mathcal{V}$, this equation reduces further to

$$X\dot{e} = e. \tag{18}$$

Our construction of $\mathcal{O}$ in Section III guarantees that $\mathcal{O}_r$ is dense in $\mathcal{O}$. This means that, for any $\xi \in \mathcal{O}$, a sequence L_k can be found such that $[I, A - L_kC, L_k] \to \xi$. This statement provides very little information, however, about the behavior of the corresponding error functions e_k as k becomes large. The results of this section address precisely this issue.

Let E_k be the natural response matrix of the kth error system; i.e., let $E_k = \exp(A - L_kC)$. Also let E be the natural response matrix of the limiting error system determined by $\xi = [X, I, Y]$. This means that E is the unique distribution with support in $[0, \infty)$ such that $X\dot{E} = E + \delta X$. (See [6, Chapter 22].)

Lemma 6.1: The unique solution of $X\dot{E} = E + \delta X$ also satisfies $\dot{E}X = E + \delta X$.

Proof: There exists a nonsingular N such that $N^{-1}XN = \operatorname{diag}(A_s^{-1}, A_f)$ for some nonsingular A_s and nilpotent A_f. (This is equivalent to the Weierstrass decomposition (4) for $G = I$.) Hence the solution of $X\dot{E} = E + \delta X$ is $E = N\left(\exp(A_s), -\sum_{i=1}^{n-1}\delta^{(i-1)}A_f^i\right)N^{-1}$. The result then follows by direct substitution. □

Theorem 6.2: Let $\xi \in \mathcal{O}_s$ be stable with $\operatorname{ord}\xi = n - p$ as in (9), and let $L_k = N[(P + kI)^T(R + k\Lambda)^T]^T$ and $\xi_k = [I, A - L_kC, L_k]$. Then $\xi_k \to \xi$, the sequence E_k is uniformly bounded, and $E_k \to E$ uniformly on $[\epsilon, \infty)$ for every $\epsilon > 0$.

Proof: Let $\Gamma = S - \Lambda Q$, and recall the Γ is stable. Then $\det(A - L_kC) = \det\begin{bmatrix} -kI & Q \\ -k\Lambda & S\end{bmatrix} = (-k)^p \det\Gamma \neq 0$. Therefore, we may write $\xi_k = [(A - L_kC)^{-1}, I, (A - L_kC)^{-1}L_k]$. Next observe that

$$\begin{aligned} &N^{-1}(A - L_kC)^{-1}N \\ &= \begin{bmatrix} -kI & Q \\ -k\Lambda & S \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{1}{k}(I + Q\Gamma^{-1}\Lambda) & \frac{1}{k}Q\Gamma^{-1} \\ -\Gamma^{-1}\Lambda & \Gamma^{-1} \end{bmatrix} \\ &\to \begin{bmatrix} 0 & 0 \\ -\Gamma^{-1}\Lambda & \Gamma^{-1} \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} &N^{-1}(A - L_kC)^{-1}L_k \\ &= \begin{bmatrix} -kI & Q \\ -k\Lambda & S \end{bmatrix}^{-1} \begin{bmatrix} P + kI \\ R + k\Lambda \end{bmatrix} \\ &= \begin{bmatrix} -I - \frac{1}{k}(P - Q\Gamma^{-1}(R - \Lambda P)) \\ \Gamma^{-1}(R - \Lambda P) \end{bmatrix} \to \begin{bmatrix} -I \\ \Gamma^{-1}(R - \Lambda P) \end{bmatrix}. \end{aligned}$$

Thus $\xi_k \to \xi$.

We have $E_k = N \exp\left(\left[\begin{smallmatrix} -kI & Q \\ -k\Lambda & S \end{smallmatrix}\right]\right) N^{-1}$. Let

$$\Psi_k = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} N^T E_k^T N^{-T} \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} = \exp\left(\begin{bmatrix} S^T & Q^T \\ -k\Lambda^T & -kI \end{bmatrix}\right).$$

By substitution it follows that Ψ_k is the solution of

$$\begin{bmatrix} I & 0 \\ 0 & (1/k)I \end{bmatrix} \dot{\Psi}_k = \begin{bmatrix} S^T & Q^T \\ -\Lambda^T & -I \end{bmatrix} \Psi_k + \delta \begin{bmatrix} I & 0 \\ 0 & (1/k)I \end{bmatrix}. \tag{19}$$

Standard singular perturbation theory shows that Ψ_k is uniformly bounded and $\Psi_k \to \Psi$ uniformly on each $[\epsilon, \infty)$, where Ψ is the solution of

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \dot{\Psi} = \begin{bmatrix} S^T & Q^T \\ -\Lambda^T & -I \end{bmatrix} \Psi + \delta \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}. \tag{20}$$

(See, e.g., [9].) From (19), E_k is uniformly bounded, and $E_k \to \overline{E}$ on every $[\epsilon, \infty)$, where $\overline{E} = N\left[\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right]\Psi^T\left[\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right]N^{-1}$. It remains to show that $\overline{E} = E$. Note that $\left[\begin{smallmatrix} S^T & Q^T \\ -\Lambda^T & -I \end{smallmatrix}\right]^{-1} = \left[\begin{smallmatrix} \Gamma^{-T} & \Gamma^{-T}Q^T \\ -\Lambda^T\Gamma^{-T} & -I-\Lambda^T\Gamma^{-T}Q^T \end{smallmatrix}\right]$. Premultiplication of (20) by the right side of the last equation yields $\left[\begin{smallmatrix} \Gamma^{-T} & 0 \\ -\Lambda^T\Gamma^{-T} & 0 \end{smallmatrix}\right]\dot{\Psi} = \Psi + \delta\left[\begin{smallmatrix} \Gamma^{-T} & 0 \\ -\Lambda^T\Gamma^{-T} & 0 \end{smallmatrix}\right]$. Thus the expression for $\overline{E}$ and Lemma 6.1 imply

$$\begin{aligned} N&\begin{bmatrix} 0 & 0 \\ -\Gamma^{-1}\Lambda & \Gamma^{-1} \end{bmatrix} N^{-1}\dot{\overline{E}} - \overline{E} - \delta N \begin{bmatrix} 0 & 0 \\ -\Gamma^{-1}\Lambda & \Gamma^{-1} \end{bmatrix} N^{-1} \\ &= N \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \left(\dot{\Psi} \begin{bmatrix} \Gamma^{-T} & 0 \\ -\Lambda\Gamma^{-T} & 0 \end{bmatrix} - \Psi \right. \\ &\quad \left. - \delta \begin{bmatrix} \Gamma^{-T} & 0 \\ -\Lambda\Gamma^{-T} & 0 \end{bmatrix}\right)^T [0\ \ 1] N^{-1} = 0. \end{aligned}$$

□

A weaker result than Theorem 6.2 can be proven with regard to zeroth-order observers. Since (A, C) is observable, $(A^T C^T)$ is controllable, so from [1, pp. 342–343] there exist $K_1 \in \mathbb{R}^{p\times n}$ and $v \in \mathbb{R}^p$ such that $(A^T + C^T K_1, C^T v)$ is controllable with $A^T + C^T K_1$ nilpotent. Thus there exists a nonsingular $N \in \mathbb{R}^{n\times n}$ such that

$$N^{-1}(A^T + C^T K_1)N = \begin{bmatrix} 0 & 1 & & \\ & & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix},$$

$$N^{-1}C^T v = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Let $\beta_{ik} = \binom{n}{i} k^{n-i}$, $k_{2k} = [\beta_{0k} \cdots \beta_{n-1,k}]$, $L_k = -(K_1 + v k_{2k} N^{-1})^T$.

Theorem 6.3: Let $\xi_k = [I, A - L_k C, L_k]$. Then $\xi_k \to [X, I, Y]$ for some X, Y with X nilpotent and $E_k \to E$ weak* and uniformly on $[\epsilon, \infty)$ for every $\epsilon > 0$.

Proof: Define

$$X = N^{-T} \begin{bmatrix} 0 & 1 & & \\ & & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} N^T,$$

$$Y = -XK_1^T - N^{-T} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} v^T.$$

Then

$$\begin{aligned} &(A - L_k C)^{-1} \\ &= N^{-T}(N^{-1}(A^T + C^T K_1)N + (N^{-1}C^T v)K_{2k})^{-T} N^T \\ &= N^{-T} \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ \beta_{0k} & \cdots & & \beta_{n-1,k} \end{bmatrix}^{-T} N^T \\ &= N^{-T} \begin{bmatrix} -\frac{\beta_{1k}}{\beta_{0k}} & 1 & & \\ \vdots & 0 & \ddots & \\ -\frac{\beta_{n-1,k}}{\beta_{0k}} & & \ddots & 1 \\ \frac{1}{\beta_{0k}} & & & 0 \end{bmatrix} N^T \to X \end{aligned} \tag{21}$$

and $(A - L_k C)^{-1} L_k = -(A - L_k C)^{-1} K_1^T - N^{-T}[0 \cdots 0\ \ 1]^T v^T \to Y$. Hence, $\xi_k \to [X, I, Y]$.

It follows from (21) that ξ_k has $\Delta_k(s) = (s + k)^n$; hence, from [5, Theorem 3], $E_k \to E$ weak*. For $t > 0$, $E(t) = 0$ and $E_k = \exp(A - L_k C)$. By direct calculation it follows that each entry of E_k is of the form $e_{ijk}(t) = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \gamma_{ij} k^j t^i e^{-kt}$, where the γ_{ij} are independent of k. From elementary analysis, $e_{ijk} \to 0$ uniformly on $[\epsilon, \infty)$ for every $\epsilon > 0$ and every i, j. □

In spite of the fact that Theorems 6.2 and 6.3 make similar statements about error convergence for certain approximations to both minimal-order stable and zeroth-order observers, there is a critical difference between the two results: In Theorem 6.2 uniform boundedness of E_k is guaranteed, while in Theorem 6.3 uniform boundedness is replaced by weak* convergence. This change constitutes a drastic weakening in the kind of convergence that one can expect when approximating the two types of observers. Indeed, upon closer examination, it can be seen that the construction used in Theorem 6.3 yields an error sequence E_k which exhibits large peaking behavior in the "boundary layer" $[0, \epsilon)$. Such peaking occurs typically when approximating a system exhibiting impulsive behavior as in (6).

Another difference between Theorems 6.2 and 6.3 is that Theorem 6.2 provides an explicit construction for approximating any minimal-order stable observer, while Theorem 6.3 merely gives an approximating sequence for a single zeroth-order observer. Given an arbitrary zeroth-order observer, finding a full-order regularization which is well behaved in both the parametric and error sense is an open problem.

Taking the issue of boundary layer peaking one step further, our next result shows that such behavior of the estimation error proves disastrous in a large class of optimal estimation problems. Note that, for any stable $\xi \in \mathcal{O}$, (6) implies that $E \in L_\mu^{n\times n}$ for $1 \le \mu \le \infty$.

Theorem 6.4: Let $p < n$, $1 < \mu \le \infty$, and $\xi_k \in \mathcal{O}$ be stable for all k. If, for every $\sigma < \infty$, there exists a $k_0 < \infty$ such that $k > k_0$ implies that each eigenvalue λ_{ik} of ξ_k satisfies $|\lambda_{ik}| > \sigma$, then $\|E_k\|_\mu \to \infty$ as $k \to \infty$.

Proof: In the first part of the proof, we make use of two topologies on L_μ. Let $\mathcal{T}_1$ be the weak* topology on K' relativized to $L_\mu \subset K'$. $\mathcal{T}_1$ uses K as the space of test functions. Let $\mathcal{T}_2$ be the weak* topology on L_μ, viewed as the dual space of L_ν, where μ and ν are conjugate exponents. Thus $\mathcal{T}_2$ uses L_ν as the test function space. Since $K \subset L_\nu$, $\mathcal{T}_1 \subset \mathcal{T}_2$.

Suppose $\|E_k\|_\mu$ has a bounded subsequence $\|E_{k_\alpha}\|_\mu$. The Banach–Alaoglu Theorem [10, Theorem 3.5.16] implies that $\{E_{k_\alpha}\}$ lies in a $\mathcal{T}_2$ compact subset of $L_\mu^{n\times n}$. Thus there exists a $\mathcal{T}_2$ convergent subsequence $E_{k_{\alpha_\beta}}$. The same subsequence must also converge relative to $\mathcal{T}_1$.

Let $\xi_k = [X_k, I, Y_k]$. From [2, pp. 341–344], $\mathcal{T}_1$ convergence of $E_{k_{\alpha_\beta}}$ guarantees that $X_{k_{\alpha_\beta}}$ converges to some matrix X. The eigenvalues of X_k are simply the eigenvalues of ξ_k along with some additional zeros; consequently, X is nilpotent, and $E_{k_{\alpha_\beta}} \to E = -\sum_{i=0}^{n-1} \delta^{(i-1)} X^i$. Since $p < n$, Theorem 5.2 applied to the sequence of error systems (18) implies $X \neq 0$, so E has an entry $e_{ij} \notin L_\mu$.

Consider the corresponding entry e_{ijk} in E_k. We will show that $\|e_{ijk_{\alpha_\beta}}\|_\mu \to \infty$, contradicting boundedness of $\|E_{k_\alpha}\|$. Impose the L_ν norm on K; this generates the dual space K^* (distinct from K'). In fact, K is dense in L_ν, so $K^* = L_\mu$. Since $e_{ij} \notin L_\mu$, e_{ij} determines an unbounded linear functional on K (i.e., $\sup_{\substack{\varphi \in K \\ \|\varphi\|_\nu = 1}} |\langle e_{ij}, \varphi\rangle| = \infty$). Let $\rho > 0$ be given. There exists a $\psi \in K$ such that $\|\psi\|_\nu = 1$ and $|\langle e_{ij}, \psi\rangle| > \rho$. It follows from $\langle e_{ijk_{\alpha_\beta}}, \psi\rangle \to \langle e_{ij}, \psi\rangle$ that for large β we have $\|e_{ijk_{\alpha_\beta}}\|_\mu = \sup_{\substack{\varphi \in K \\ \|\varphi\|_\nu = 1}} |\langle e_{ijk_{\alpha_\beta}}, \varphi\rangle| \geq |\langle e_{ijk_{\alpha_\beta}}, \psi\rangle| > \rho$. Since ρ is arbitrary, $\|e_{ijk_{\alpha_\beta}}\|_\mu \to \infty$. □

The condition $p < n$ cannot be weakened in Theorem 6.4. Indeed, when $p = n$, Q, and S in (8) are zero dimensional, so (9) reduces to $\xi = [0, I, -C^{-1}]$. Theorem 6.2 shows that ξ can then be approximated by full-order observers in such a way that $\|E_k\|_\mu \to 0$ for $1 \leq \mu < \infty$ (but not $\mu = \infty$). But $p = n$ is the trivial case where all plant states are directly measurable at the output.

Theorem 6.4 is in direct conflict with what is arguably a common piece of folklore related to observer design. According to one school of thought, the performance of full-order observers or minimal order observers can always be improved by choosing L or Λ to move the observer eigenvalues further to the left in the complex plane. This has the effect of making the decay rate in the error system (18) larger, thus making the error "smaller." The same idea is reflected in Theorems 6.2 and 6.3 under the guise of uniform convergence on $[\epsilon, \infty)$. However, Theorem 6.4 shows that the act of moving all eigenvalues arbitrarily far to the left necessarily carries with it the undesirable side effect of boundary layer peaking (or worse) and consequent divergence of the L_μ norm.

Returning to our discussion of noise, we conclude that boundary layer peaking, impulsive behavior, infinite L_μ cost, etc. are all manifestations of the effect of disturbances on the successive differentiation scheme or any approximation to it. Proximity to a zeroth-order observer always carries with it the undesirable behavior present in the zeroth-order observer itself.

References

[1] C.-T. Chen, *Linear System Theory and Design.* New York: Holt, Rinehart, and Winston, 1984.
[2] J. D. Cobb, "Fundamental properties of the manifold of singular and regular systems," *J. Math. Anal. Appl.*, vol. 20, no. 1, pp. 328–353, Nov. 1986.
[3] F. R. Gantmacher, *The Theory of Matrices.* New York: Chelsea, 1960.
[4] I. M. Gelfand and G. E. Shilov, *Generalized Functions.* New York: Academic, 1964.
[5] J. D. Cobb, "On the solutions of linear differential equations with singular coefficients," *J. Differential Equations*, vol. 46, no. 3, pp. 310–323, Dec. 1982.
[6] G. Doetsch, *Introduction to the Theory and Application of the Laplace Transformation.* New York: Springer-Verlag, 1974.
[7] T. Kailath, *Linear Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1980.
[8] N. Safari-Shad and J. D. Cobb, "Optimal LTI least squares estimation for systems with noise-free measurements," submitted.
[9] S. L. Campbell and N. J. Rose, "Singular perturbation of autonomous linear systems," *SIAM J. Mathe. Anal.*, vol. 10, pp. 542–551, 1979.
[10] R. B. Ash, *Real Analysis and Probability.* New York: Academic, 1972.