On Singular Value Functions and Hankel Operators for Nonlinear Systems

Jacquelien M.A. Scherpen
Delft University of Technology
Fac. of Inf. Techn. & Syst., Dept. Electrical Eng.
P.O. Box 5031, 2600 GA Delft
The Netherlands
J.M.A.Scherpen@et.tudelft.nl

W. Steven Gray
Department of Electrical and Computer Engineering
Old Dominion University
Norfolk, Virginia 23529-0246, U.S.A.
gray@ece.odu.edu

In: Proc. ACC'99, San Diego, CA, USA, pp. 2360–2364

Abstract

In linear system theory, the Hankel singular values are often computed in a state space setting using the product of Gramian matrices. They are known, however, to be intrinsically dependent only on the input-output map and not on any choice of state space coordinates. In the nonlinear case, there are well defined notions of singular value functions and a Hankel operator, but the connections between the two have not been established. In this paper we address the problem in two ways, and show that it is possible to establish an explicit connection using a method that is very reminiscent of the way singular values are usually defined for compact linear operators.

1 Introduction

In the theory of continuous-time linear systems, the system Hankel operator plays an important role in a number of realization problems. The compact Hankel operator supplies a set of similarity invariants, the so called Hankel singular values, which can be used to quantify the importance of each state in the corresponding input-output map [8]. The Hankel operator can also be factored into the composition of an observability and controllability operator, from which Gramian matrices can be defined and the notion of balanced realizations follows [5, 8, 11]. The Hankel singular values are most easily computed in a state space setting using the product of the Gramian matrices, though intrinsically they depend only on the given input-output mapping.

In the case of continuous-time nonlinear systems Hankel theory is less developed, but important results do exist. The first result along these lines is due to Fliess [3, 4, 7] who used a system Hankel *mapping* to describe when an affine realization of an input-output map described by a formal power series is minimal. In a quite different setting, the notion of Hankel singular values was generalized to the nonlinear case by Scherpen in [13, 14] and used in model reduction problems. Connections between minimality and these invariants were then introduced in [15]. In [6], a nonlinear analogue of the system Hankel operator was introduced. Its relationship to the Hankel mapping of Fliess was established, and the controllability/observability factorization problem was solved.

Despite this progress, however, there are many important open questions in Hankel theory for nonlinear systems. In this paper we address two questions. We first extend the development of the theory in [6] concerning the connection between the Hankel operator factorization and the energy functions via the notion of a nonlinear Hilbert adjoint mapping. Some explicit expressions for the operators are presented, and a nonlinear extension of the duality between observability and controllability Gramians for linear state space systems is given. We then develop an explicit connection to the singular functions using a method that is very reminiscent of the way singular values are usually defined in the linear case, except that a state space model is still required.

In Section 2, we review the existing theory concerning Hankel operators and Hankel singular values for linear and nonlinear input-output systems. Then in Section 3, we present the new developments in the theory for the nonlinear case in two separate subsections. The first subsection addresses the problem of relating energy functions to factorizations of the Hankel operator. The final subsection presents our connection between singular value functions and the Hankel operator. The paper concludes with some observations about future research.

The mathematical notation used throughout is fairly standard. The inner product on $\mathbb{R}^n$ is represented as $\langle x, y\rangle = x^T y$. $L_2^i(a,b)$ represents the set of Lebesgue measurable functions, i-component vector-valued, with finite L_2 norm, $\|\cdot\|_{L_2}$. The inner product on $L_2^i(a,b)$ is denoted by

$$\langle f, g\rangle_{L_2} = \int_a^b f(t)^T g(t)\, dt.$$

We abbreviate $L_2^i(-\infty,\infty)$ as L_2^i. If h is a differentiable function, and g is a vector field then $L_g h$ denotes the Lie derivative of h with respect to g.

2 Preliminaries

In this section we review some existing theory concerning Hankel operators, Hankel singular values and system Gramians. The linear time-invariant case is first outlined as a kind of paradigm for the nonlinear theory, which is covered subsequently.

2.1 The linear case

Consider a continuous-time, causal linear input-output system $S : u \to y$ with impulse response $H(t)$. If S is also BIBO stable then the system Hankel operator is the well defined mapping

$$\begin{aligned}\hat{\mathcal{H}} &: L_2^m[0,+\infty) \to L_2^p[0,+\infty) \\ &: \hat{u} \to \hat{y}(t) = \int_0^\infty H(t+\tau)\hat{u}(\tau)\,d\tau.\end{aligned}$$

If we define the *time flipping* operator as

$$\begin{aligned}\mathcal{F} &: L_2^m[0,+\infty) \to L_2^m(-\infty,0] \\ &: \hat{u} \to u(t) = \begin{cases} \hat{u}(-t) & : \quad t<0 \\ 0 & : \quad t \geq 0 \end{cases}\end{aligned} \tag{1}$$

then clearly $\hat{\mathcal{H}} = S\mathcal{F}$. When $\hat{\mathcal{H}}$ is known to be a compact operator, then its (Hilbert) adjoint operator, $\hat{\mathcal{H}}^*$, is also compact, and the composition $\hat{\mathcal{H}}^*\hat{\mathcal{H}}$, is a self-adjoint compact operator with a well defined spectral decomposition:

$$\hat{\mathcal{H}}^*\hat{\mathcal{H}} = \sum_{i=1}^{\infty} \sigma_i^2 \, \langle \cdot, \psi_i \rangle_{L_2} \, \psi_i, \quad \sigma_i \geq 0, \tag{2}$$

$$\langle \psi_i, \psi_j \rangle_{L_2} = \delta_{ij}, \quad \langle \psi_i, (\hat{\mathcal{H}}^*\mathcal{H})(\psi_i) \rangle_{L_2} = \sigma_i^2. \tag{3}$$

where σ_i^2 is an eigenvalue of $\hat{\mathcal{H}}^*\hat{\mathcal{H}}$ with corresponding eigenvector ψ_i [9], ordered as $\sigma_1 \geq \cdots \geq \sigma_n > 0$, and called the *Hankel singular values* for the input-output system S.

Let (A,B,C) be a state space realization of S with dimension n. If the realization is asymptotically stable then the Hankel operator can be written as the composition of uniquely determined observability and controllability operators; that is, $\hat{\mathcal{H}} = \hat{\mathcal{O}}\hat{\mathcal{C}}$, where the controllability and observability operators are defined as

$$\begin{aligned}\hat{\mathcal{C}} &: L_2^m[0,+\infty) \to \mathbb{R}^n : \hat{u} \to \int_0^\infty e^{At}B\hat{u}(t)\,dt \\ \hat{\mathcal{O}} &: \mathbb{R}^n \to L_2^p[0,+\infty) : x \to \hat{y}(t) = Ce^{At}x.\end{aligned}$$

Since $\hat{\mathcal{C}}$ and $\hat{\mathcal{O}}$ have a finite dimensional range and domain, respectively, they are compact operators; and the composition $\hat{\mathcal{O}}\hat{\mathcal{C}}$ is also a compact operator [9]. From the definition of the adjoint operator, it is easily shown that $\hat{\mathcal{C}}$ and $\hat{\mathcal{O}}$ have corresponding adjoints

$$\begin{aligned}\hat{\mathcal{C}}^* &: \mathbb{R}^n \to L_2^m[0,+\infty) : x \to B^T e^{A^T t} x \\ \hat{\mathcal{O}}^* &: L_2^p[0,+\infty) \to \mathbb{R}^n : y \to \int_0^\infty e^{A^T t} C^T y(t)\,dt.\end{aligned}$$

For any $x_1, x_2 \in \mathbb{R}^n$:

$$\begin{aligned}< x_1, \hat{\mathcal{C}}\hat{\mathcal{C}}^* x_2 > &= x_1^T \int_0^\infty e^{At}BB^T e^{A^T t} dt \; x_2 \\ &:= x_1^T P x_2 \qquad (4) \\ < x_1, \hat{\mathcal{O}}^*\hat{\mathcal{O}} x_2 > &= x_1^T \int_0^\infty e^{A^T t}C^T C e^{At} dt \; x_2 \\ &:= x_1^T Q x_2. \qquad (5)\end{aligned}$$

P and Q are the usual controllability and observability Gramian matrices. It is well known that the nonzero eigenvalues of PQ are equivalent to the squared Hankel singular values of S. When (A,B,C) is in input normal form, i.e., when $P = I_n$, then the Hankel singular values are determined completely by the eigenvalues of Q. Now consider the following definition for the *energy functions* first described by Scherpen for general affine nonlinear systems [13].

Definition 2.1 *The* **controllability** *and* **observability functions** *for the system* (A,B,C) *are defined, respectively, as*

$$L_c(x) = \min_{\substack{u \in L_2(-\infty,0) \\ x(-\infty)=0,\; x(0)=x}} \frac{1}{2}\int_{-\infty}^0 \|u(t)\|^2 \, dt$$

$$L_o(x) = \frac{1}{2}\int_0^\infty \|y(t)\|^2 \, dt,$$

when $x(0) = x$*, and* $u(t) = 0$ *for* $0 \leq t < \infty$.

Clearly, $L_c(x)$ has the interpretation of being the minimum amount of input energy required to drive the system from zero at $t = -\infty$ to $x(0) = x$, while $L_o(x)$ is equivalent to the energy generated by the natural response of the system to $x(0) = x$. These functions need not be well defined for all $x \in \mathbb{R}^n$. If the realization is reachable and asymptotically stable, then in light of equations (4) and (5), it can be shown directly in the linear case that we have

$$L_c(x) = \frac{1}{2}x^T P^{-1} x = \frac{1}{2}\langle x, (\hat{\mathcal{C}}\hat{\mathcal{C}}^*)^{-1} x \rangle \tag{6}$$

$$L_o(x) = \frac{1}{2}x^T Q x = \frac{1}{2}\langle x, (\hat{\mathcal{O}}^*\hat{\mathcal{O}}) x \rangle. \tag{7}$$

2.2 The nonlinear case

Now, let S be a given input-output map represented by a convergent generating series

$$S : u \to y(t) = \sum_{\eta \in I^*} c(\eta) \, E_\eta(t,t_0)[u], \;\; t \geq t_0,$$

where I^* is the set of multi-indices for the index set $I = \{0,1,\ldots,m\}$, $c(\eta) \in \mathbb{R}^p$, and

$$E_{i_k \cdots i_0}(t,t_0)[u] = \int_{t_0}^t u_{i_k}(\tau) E_{i_{k-1}\cdots i_0}(\tau,t_0)[u]\, d\tau$$

with $E_\emptyset(t,t_0)[u] := 1$ and $u_0(t) := 1$. The mapping S can also be represented by a formal power series in non commuting monomials $\mathcal{Z} = \{z_0, z_1, \ldots, z_m\}$ via $c = \sum_{\eta \in I^*} c(\eta) z_\eta$, where $z_\eta := z_{i_k} \ldots z_{i_0}$ when $\eta = (i_k \ldots i_0)$ (see, for example, [4, 7]). In the following development, we use the convention that L_2-stability of an input-output system, S, means that $u \in L_2^m(-\infty,0]$ implies that $S(u)$ restricted to $[0,+\infty)$ is in $L_2^p[0,+\infty)$. Similarly, L_2 input-to-state stability on a set W of a state space realization implies that when $u \in L_2^m(-\infty,0]$ then the corresponding state vector, $x(t)$, (assuming initial condition $x(-\infty) = 0$) is finite on $(-\infty,0]$ and always contained in W. In this context, consider the following definition.

Definition 2.2 **[6]** *For any causal* L_2*-stable input-output system* S*, the corresponding* **Hankel operator** *is*

$$\begin{aligned}\hat{\mathcal{H}} &: L_2^m[0,+\infty) \to L_2^p[0,+\infty) \\ &: \hat{u} \to \hat{y} = (S\mathcal{F})(\hat{u}),\end{aligned}$$

where $\mathcal{F}$ *is the time-flipping operator as given in (1).*

Observe that the usual interpretation from linear system theory that $\hat{\mathcal{H}}$ maps *past inputs* to *future outputs* is preserved by this definition.

Let M be an n-dimensional analytic state space manifold, and let

$$\dot{x} = f(x) + g(x)u, \quad y = h(x) \tag{8}$$

be a system defined in terms of local coordinates on M. We assume that f, g, and h are analytic on M. A realization (f,g,h) defined locally about $x^o \in M$ is said to *realize* a formal power series c if for every $\eta = (i_k \dots i_0) \in I^*$

$$c(\eta) = L_{g_\eta} h(x^o) := L_{g_{i_0}} L_{g_{i_1}} \dots L_{g_{i_k}} h(x^o),$$

where $g_0 := f$ and g_i is the ith column of g when $i > 0$. Consider the corresponding controllability/observability factorization of the system Hankel operator.

Theorem 2.1 [6] *Let (f,g,h) be an analytic realization in a neighborhood W of 0 of an L_2-stable input-output mapping $S : u \to y(t) = \sum_{\eta\in I^*} c(\eta)\, E_\eta(t,-\infty)[u]$. If the realization is L_2 input-to-state stable on W then the corresponding Hankel operator $\hat{\mathcal{H}} : \hat{u} \to \hat{y}$ can be written as the composition*

$$\hat{\mathcal{H}} = \hat{\mathcal{O}}\hat{\mathcal{C}},$$

where the controllability and observability operators are defined, respectively, as

$$\begin{aligned} \hat{\mathcal{C}} &: L_2^m[0,+\infty) \to W_c \\ &: \hat{u} \to x = \sum_{\eta\in I^*} L_{g_\eta} \mathcal{I}(0)\, E_\eta(0,-\infty)[u] \end{aligned} \qquad (9)$$

$$\begin{aligned} \hat{\mathcal{O}} &: W_c \to L_2^p[0,+\infty) \\ &: x \to \hat{y}(t) = \sum_{i=0}^{\infty} L_{g_0}^i h(x)\, E_{\underbrace{0\dots0}_{i}}(t,0) \end{aligned} \qquad (10)$$

with $W_c := \hat{\mathcal{C}}(L_2^m[0,+\infty)) \subset W$, and $\mathcal{I} : W \to W$ denoting the identity map on W, $u = \mathcal{F}(\hat{u})$ and $g_0 := f$.

In Section 3, it is shown that these generalized controllability and observability operators can be related to the energy functions. Here we review how to define Hankel singular value functions from a given realization of S using the energy functions L_c and L_o under the following additional assumptions:

1. f is asymptotically stable on some neighborhood Y of 0.
2. L_c and L_o are smooth and finite functions on Y.
3. the system is zero-state observable on Y.
4. $\dfrac{\partial^2 L_c}{\partial x^2}(0) > 0$ and $\dfrac{\partial^2 L_o}{\partial x^2}(0) > 0$.

The following lemma and theorem describe the origin of the singular value functions.

Lemma 2.1 [13] *There exists a coordinate transformation $x = \phi(\bar{x})$, $\phi(0) = 0$ (defined on a neighborhood of 0), such that in the new coordinates $\bar{x} = \phi^{-1}(x)$ the function $L_c(x)$ is of the form*

$$L_c(\phi(\bar{x})) = \frac{1}{2}\bar{x}^T \bar{x}.$$

Furthermore, in the new coordinates $\bar{x} = \phi^{-1}(x)$, we can write $L_o(x)$ in the form

$$L_o(\phi(\bar{x})) = \frac{1}{2}\bar{x}^T M(\bar{x})\bar{x} \;\text{ where }\; M(0) = \frac{\partial^2 L_o}{\partial x^2}(0)$$

with $M(\bar{x})$ an $n \times n$ symmetric matrix such that its entries are smooth functions of $\bar{x}$.

Theorem 2.2 [13] *Consider a system (f,g,h) and assume that there exists a neighborhood of 0 where the number of eigenvalues $M(\bar{x})$ is constant. Then, on a neighborhood U of 0 there exists a coordinate transformation $x = \psi(z)$, $\psi(0) = 0$, such that in the new coordinates $z \in W := \psi^{-1}(U)$ the function L_c is of the form*

$$\check{L}_c(z) := L_c(\psi(z)) = \frac{1}{2}z^T z,$$

and the function L_o is of the form

$$\check{L}_o(z) := L_o(\psi(z)) = \frac{1}{2}z^T \begin{pmatrix} \tau_1(z) & & 0 \\ & \ddots & \\ 0 & & \tau_n(z) \end{pmatrix} z,$$

where $\tau_1(z) \geq \dots \geq \tau_n(z)$ are smooth functions of z, referred to as the **singular value functions***.*

To put the controllability and observability functions into *balanced form*, we need the additional coordinate transformation $\bar{z}_i = \eta_i(z_i) := \tau_i(0,..,0,z_i,0,..0)^{\frac{1}{4}} z_i$, $i = 1,..,n$ and hence $\bar{z} = \eta(z) := (\eta_1(z_1)\dots\eta_n(z_n))$ on $\bar{z} \in \bar{W} := \eta(W)$. Then in the new coordinates it follows that on the coordinates axes

$$\begin{aligned} \hat{L}_c(0,..,0,\bar{z}_i,0,..,0) &= \frac{1}{2}\bar{z}_i^2 \sigma_i(\bar{z}_i)^{-1} \\ \hat{L}_o(0,..,0,\bar{z}_i,0,..,0) &= \frac{1}{2}\bar{z}_i^2 \sigma_i(\bar{z}_i), \end{aligned}$$

where $\sigma_i(\bar{z}_i) = \tau_i(0,..,0,\eta_i^{-1}(\bar{z}_i),0,..,0)^{\frac{1}{2}}$ for $i = 1,..,n$.

3 New Developments

In this section we present new theoretical results concerning singular value functions for nonlinear systems. In the first subsection, we introduce the notion of an adjoint operator for a nonlinear operator. This concept helps establish a connection between the controllability/observability factorization of the Hankel operator presented in Theorem 2.1 and the energy functions. It is also a critical device in the final subsection where we develop a relationship between singular value functions and the system Hankel operator.

3.1 Energy functions from the Hankel operator factorization

Let F be a topological vector space over $\mathbb{R}$ with dual space F' [12]. Let E be a nonempty set, and $\mathcal{A}$ a collection of nonempty subsets of E. Let E^β be a linear space of real-valued functions x^β on E with the property that the restriction x^β_A to every $A \in \mathcal{A}$ is bounded. A mapping $T : E \to F$ is called *$\mathcal{A}$-bounded* if T maps the sets of $\mathcal{A}$ into bounded subsets of F. For any $\mathcal{A}$ -bounded mapping $T : E \to F$ define the *dual map* of T as

$$\begin{aligned} T' &: F' \to E^\beta \\ &: y' \to (T'(y'))(x) := (y' \circ T)(x), \quad x \in E \end{aligned}$$

(see, for example, [1]). Now if F is endowed with an inner product $\langle\cdot,\cdot\rangle_F$ then it follows from the Riesz representation theorem that for any $y' \in F'$ there exists a unique $y \in F$ such that $y'(\cdot) = \langle y,\cdot\rangle_F$. Hence one can write the identity

$$(T'(y'))(x) = \langle y, T(x)\rangle_F, \quad x \in E.$$

Now suppose E has an inner product $\langle\cdot,\cdot\rangle_E$, and let $y \in F$ be fixed. We are interested in the problem of determining a corresponding $\tilde{x}_y \in E$ such that

$$\langle T(x), y\rangle_F = \langle x, \tilde{x}_y\rangle_E, \quad x \in E. \qquad (11)$$

If T were a linear operator then such an $\tilde{x}_y$ is known to always exist, in fact $\tilde{x}_y = T^*(y)$, where T^* is the Hilbert adjoint of T. But in this more general context, the existence of $\tilde{x}_y$ is not automatic. In fact, it is conjectured that the identity (11) may only be meaningful if $\tilde{x}_y$ is also a function of x as well. In what follows below, we simply assume the existence of a well defined mapping $T^* : F \times E \to E$, called the *nonlinear Hilbert adjoint*, such that

$$\langle T(x), y\rangle_F = \langle x, T^*(y,x)\rangle_E, \quad x \in E, \quad y \in F. \qquad (12)$$

The fact that the domain of T^* has the form $F \times E$ agrees with the state space notion of adjoint systems based on the Hamiltonian extension given in [2, 17].

Consider now a realization (f,g,h) from Theorem 2.1 with the additional assumptions that there is an equilibrium at 0, i.e., $f(0) = 0$, and this equilibrium is asymptotically stable on W. Without loss of generality we may assume that $h(0) = 0$. If the realization is asymptotically reachable on W (see [15, 16]), then it is clear that for every $x \in W$ there exists at least one $\hat{u} \in L_2^m[0,+\infty)$ such that $\hat{\mathcal{C}}(\hat{u}) = x$. The existence of a unique minimum energy control thus guarantees a well defined pseudo-inverse of $\hat{\mathcal{C}}$ on W, denoted here by $\hat{\mathcal{C}}^\dagger$. The following identities are immediate after applying (12) with $T = \hat{\mathcal{C}}^\dagger$ and $y = \hat{\mathcal{C}}^\dagger(x)$:

$$\begin{aligned} L_c(x) &= \frac{1}{2}\|\hat{\mathcal{C}}^\dagger(x)\|_{L_2}^2 = \frac{1}{2}\langle \hat{\mathcal{C}}^\dagger(x), \hat{\mathcal{C}}^\dagger(x)\rangle_{L_2} \\ &= \frac{1}{2}\langle x, \hat{\mathcal{C}}^{\dagger *}(\hat{\mathcal{C}}^\dagger(x),x)\rangle := \frac{1}{2}\langle x, p(x)\rangle. \qquad (13)\end{aligned}$$

It was shown in [13] that L_c must always have a local minimum at $x = 0$, i.e., $\frac{\partial L_c}{\partial x}(0) = 0$. Thus it is clear, via [10], that one can always write $p(x) = \hat{P}(x)x$ for some matrix-valued function $\hat{P}$.

The corresponding notion for $L_o(x)$ follows analogously:

$$\begin{aligned} L_o(x) &= \frac{1}{2}\|\hat{\mathcal{O}}(x)\|_{L_2}^2 = \frac{1}{2}\langle \hat{\mathcal{O}}(x), \hat{\mathcal{O}}(x)\rangle_{L_2} \\ &= \frac{1}{2}\langle x, \hat{\mathcal{O}}^*(\hat{\mathcal{O}}(x),x)\rangle := \frac{1}{2}\langle x, q(x)\rangle. \qquad (14)\end{aligned}$$

In this case, if the system is zero-state observable, then it is known that L_o must have a local minimum at $x = 0$, i.e., $\frac{\partial L_o}{\partial x}(0) = 0$ [13]. Thus, after differentiating the expression for L_o given in (14), it follows that $q(x) = \hat{Q}(x)x$ for some matrix-valued function $\hat{Q}$. Comparing the functions $\hat{P}$ and $\hat{Q}$ defined here to the expression given in (6) and (7), respectively, allows one to conclude that the linear case always results in the trivial situation where the functions $\hat{P}$ and $\hat{Q}$ are constant matrix functions, specifically, $\hat{P}(x) = P^{-1}$ and $\hat{Q}(x) = Q$ for all $x \in W$.

Some specific nonlinear adjoint expressions

Using equation (12) one can obtain more explicit expressions which $\mathcal{O}^*(\mathcal{O}(x),x)$ and $\mathcal{C}^*(\mathcal{C}(\hat{u}),\hat{u})$ must satisfy. Consider the identities:

$$\langle \hat{\mathcal{O}}(x), \hat{\mathcal{O}}(x)\rangle_{L_2} = \int_0^\infty \left(\sum_{i=0}^\infty L_{g_0}^i h(x)\, E_{\underbrace{0\ldots0}_{i}}(t,0)\right)^T \cdot \left(\sum_{i=0}^\infty L_{g_0}^i h(x)\, E_{\underbrace{0\ldots0}_{i}}(t,0)\right) dt$$

$$\langle x, \hat{\mathcal{O}}^*(\hat{\mathcal{O}}(x),x)\rangle = x^T \hat{\mathcal{O}}^* \left(\sum_{i=0}^\infty L_{g_0}^i h(x)\, E_{\underbrace{0\ldots0}_{i}}(t,0)\, ,\, x\right).$$

Then it follows that $\hat{\mathcal{O}}^*$ must fulfill

$$\int_0^\infty \left(\sum_{i,j=0}^\infty \left(L_{g_0}^i h(x)\right)^T \left(L_{g_0}^j h(x)\right) E_{\underbrace{0\ldots0}_{i}}(t,0) E_{\underbrace{0\ldots0}_{j}}(t,0)\right) dt$$
$$= x^T \hat{\mathcal{O}}^* \left(\textstyle\sum_{i=0}^\infty L_{g_0}^i h(x)\, E_{\underbrace{0\ldots0}_{i}}(t,0)\, ,\, x\right).$$

By factoring out an x on the left hand side (this is possible since $h(0) = 0$, e.g., [10]), we can obtain an explicit expression for $\hat{\mathcal{O}}^*(\hat{\mathcal{O}}(x),x)$. Unfortunately, such an expression will not be unique.

Similarly, we can obtain for $\hat{\mathcal{C}}^*(\hat{\mathcal{C}}(\hat{u}),\hat{u})$ the following expression:

$$\left(\sum_{\eta\in I^*} L_{g_\eta}\mathcal{I}(0)\, E_\eta(0,-\infty)[\mathcal{F}(\hat{u})]\right)^T \cdot \left(\sum_{\eta\in I^*} L_{g_\eta}\mathcal{I}(0)\, E_\eta(0,-\infty)[\mathcal{F}(\hat{u})]\right)$$
$$= \int_0^\infty \hat{u}^T \hat{\mathcal{C}}^* \left(\sum_{\eta\in I^*} L_{g_\eta}\mathcal{I}(0)\, E_\eta(0,-\infty)[\mathcal{F}(\hat{u})]\, ,\, \hat{u}\right) dt.$$

Here, one can not easily see a connection with $L_c(x)$ since that quantity is in terms of $\hat{\mathcal{C}}^\dagger(x)$. If the system is in input-normal form, i.e., $L_c(x) = \frac{1}{2}x^T x$, then obviously from (13) it follows that $x^T \hat{\mathcal{C}}^{\dagger *}(\hat{\mathcal{C}}^\dagger(x),x) = x^T x$.

Energy functions and operators of pseudo-dual systems

For a linear system, the controllability and observability Gramians are the observability and controllability Gramians, respectively, of the dual state space system (e.g., [18]). In the nonlinear case, we are able to find similar relations using the so called *pseudo-dual* nonlinear state space system. While equivalent to the usual dual system in the linear case, the pseudo-dual system is distinct from other notions of dual systems found in the literature (e.g., [2]).

Consider equation (8), and assume that $f(0) = 0$, then we can write $f(x) = A(x)x$, where $A(0) = \frac{\partial f}{\partial x}(0)$, and $A(x)$ is a $n \times n$ matrix with entries depending smoothly on x. If the controllability function L_c exists and is finite on some neighborhood W of 0, then it satisfies the following Hamilton-Jacobi type of equation [13]

$$\frac{\partial L_c}{\partial x}(x)A(x)x + \frac{1}{2}\frac{\partial L_c}{\partial x}(x)g(x)g^T(x)\frac{\partial^T L_c}{\partial x}(x) = 0, \qquad (15)$$

with $L_c(0) = 0$, and $-(f(x) + g(x)g^T(x)\frac{\partial^T L_c}{\partial x}(x))$ asymptotically stable on W. This last property is equivalent to positive definiteness of $L_c(x)$, implying that $\frac{\partial L_c}{\partial x}(0) = 0$, (e.g., [15, 16]). Thus, we can write $\frac{\partial L_c}{\partial x}(x) = x^T N(x)$ where $N(0) = \frac{\partial^2 L_c}{\partial x^2}(0)$, and $N(x)$ is a $n \times n$ matrix with entries depending smoothly on x. The coordinate transformation $z := N(x)x$, ($N(x) > 0$ on a neighborhood $U \subset W$, hence the inverse exists on U), and its inverse $x = \phi(z)$ applied to equation (15) then results in the equation

$$z^T N(\phi(z))^{-1} A(\phi(z))^T z + \frac{1}{2} z^T g(\phi(z)) g(\phi(z))^T z = 0,$$

where now $N(\phi(z))^{-1}z$ is the gradient of the observability function of the pseudo-dual system

$$\dot{z} = A(\phi(z))^T z, \quad y = g(\phi(z))^T z. \tag{16}$$

Specifically, if $\bar{\mathcal{O}}$ is the observability operator as defined in (10) for system (16), and $\hat{\mathcal{C}}^\dagger$ is, as before, the pseudo-inverse of the controllability operator as defined in (9) for the original system (8), then we have, with $x = \phi(z)$ that

$$\bar{L}_o(z) = \frac{1}{2}\langle \bar{\mathcal{O}}(z), \bar{\mathcal{O}}(z)\rangle_{L_2} = \frac{1}{2}\langle \hat{\mathcal{C}}^\dagger(x), \hat{\mathcal{C}}^\dagger(x)\rangle_{L_2} = L_c(x).$$

In the special case where the original system (8) already has a controllability function in the input-normal form of Lemma 2.1 it is easily obtained that $z = x$ and $\bar{L}_o(x) = \frac{1}{2}x^T x$ is the observability function of the system

$$\dot{x} = f(x), \quad y = g(x)^T x. \tag{17}$$

This system is not the pseudo-dual system, since it has the original drift vector field. However, the above, more general analysis for pseudo-dual systems, of course includes the input-normal form case. Thus, in the input-normal form case, the controllability function of system (8) equals the observability function of both systems (16) and (17).

3.2 Singular value functions from the Hankel operator

We are now in a position to more directly connect singular value functions of a given realization of an input-output mapping S to the corresponding Hankel operator. It turns out, however, that we can not yet completely eliminate the need for a state space model. They are still very useful in *parameterizing* these functions. Let (f,g,h) be an n dimensional realization of S on a neighborhood W of 0 in input-normal/output-diagonal form (as in Theorem 2.2). Define on a neighborhood W of 0, the collection of component vectors $\tilde{x}_i = (0,\ldots,0,x_i,0,\ldots,0)$ for $i = 1,2,\ldots n$, and the functions $\hat{\sigma}_i^2(x_i) := \tau_i(\tilde{x}_i)$. Let u_i be a minimum energy input corresponding to driving the state from $x(-\infty) = 0$ to $x(0) = \tilde{x}_i$. Define $\hat{u}_i = \mathcal{F}(u_i)$.

Theorem 3.1 *The functions $\{\hat{\sigma}_i(x_i)\}_{i=1}^n$ are singular values functions of the Hankel operator $\hat{\mathcal{H}}$ in the following sense:*

$$\langle \hat{u}_j, (\hat{\mathcal{H}}^*\hat{\mathcal{H}})(\hat{u}_j)\rangle_{L_2} = \hat{\sigma}_j^2(x_j)\langle \hat{u}_j, \hat{u}_j\rangle_{L_2}, \quad i = 1,2,\ldots n, \tag{18}$$

where $(\hat{\mathcal{H}}^*\hat{\mathcal{H}}) : L_2^m \to L_2^m : \hat{u} \to \hat{\mathcal{H}}^*(\hat{\mathcal{H}}(\hat{u}),\hat{u})$.

Proof: The following equalities follow from the various assumptions above:

$$\begin{aligned}&\left\langle \hat{u}_j, (\hat{\mathcal{H}}^*\hat{\mathcal{H}})(\hat{u}_j)\right\rangle_{L_2} = \langle \hat{u}_j, \hat{\mathcal{H}}^*(\hat{\mathcal{H}}(\hat{u}_j),\hat{u}_j)\rangle_{L_2} = \\ &\langle \hat{\mathcal{H}}(\hat{u}_j), \hat{\mathcal{H}}(\hat{u}_j)\rangle_{L_2} = \langle \hat{\mathcal{O}}\hat{\mathcal{C}}(\hat{u}_j), \hat{\mathcal{O}}\hat{\mathcal{C}}(\hat{u}_j)\rangle = \\ &\langle \hat{\mathcal{O}}(\tilde{x}_j), \hat{\mathcal{O}}(\tilde{x}_j)\rangle_{L_2} = 2L_o(\tilde{x}_j) = \tau_j(\tilde{x}_j)x_j^2 = \\ &\hat{\sigma}_j^2(x_j)\cdot 2L_c(\tilde{x}_j) = \hat{\sigma}_j^2(x_j)\langle \hat{u}_j, \hat{u}_j\rangle_{L_2}.\end{aligned}$$

■

A comparison of equations (3) and (18) reveals the sense in which the $\hat{\sigma}_i$'s are singular value functions of $\hat{\mathcal{H}}$. This is really a more limited sense than in the linear case because it does not necessarily yield a spectral decomposition analogous to (2). But as was shown in [13], this concept is still useful in model reduction problems because it measures how important each coordinate direction in the state space is from the point of view of the input-output map.

4 Conclusions and Future Research

In this paper, the problem of explicitly connecting the notions of singular value functions and Hankel operators was addressed for nonlinear systems. While this could not be done in a state space (coordinate) free setting, the result was very reminiscent of the way singular values are usually defined for compact linear operators. Along the way, we also considered some Gramian generalizations and identities since in the linear case Hankel singular values are often computed from Gramians. Future work on this problem will be in the direction of apply existing theory for compact nonlinear operators to further explore when the nonlinear Hilbert adjoint operator is well defined, and perhaps this will further clarify some of the connections between the Hankel operator factorization and the Gramian generalizations $\hat{P}$ and $\hat{Q}$ present here. In addition, more research will be done with regard to the influence of the factorization on the singular value functions.

Acknowledgment

This research was supported in part by the North Atlantic Treaty Organization through the NATO Collaborative Research Grant CRG-971135.

References

[1] J. Batt, "Nonlinear Compact Mappings and Their Adjoints," *Math. Ann.*, vol. 189, pp. 5-25, 1970.

[2] P. E. Crouch and A. J. van der Schaft, *Variational and Hamiltonian Control Systems*, Lect. Notes Contr. Inf. Sci. No. 101, Springer Verlag, Berlin, 1987.

[3] M. Fliess, "Matrices de Hankel," *J. Math. Pures. Appl.*, vol. 53, pp. 197-222, 1974.

[4] M. Fliess, M. Lamnabhi and F. Lamnabhi-Lagarrigue, "An Algebraic Approach to Nonlinear Functional Expansions," *IEEE Trans. Circuits Syst.*, vol. CAS-30, no. 8, pp. 554-570, 1983.

[5] B. A. Francis, *A Course in H_∞ Control*, Number 88 in the Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin, 1987.

[6] W. S. Gray and J. M. A. Scherpen, "Hankel Operators and Gramians for Nonlinear Systems," *Proc. 37th Conf. Dec. Contr. 1998*, Tampa, Florida, pp. 1416-1421, 1998.

[7] A. Isidori, *Nonlinear Control Systems*, 3rd Edition, Springer-Verlag, London, 1995.

[8] E. A. Jonckheere and L. M. Silverman, "Singular Value Analysis of Deformable Systems," *Circuits, Systems and Signal Processing*, vol. 1, no. 3-4, pp. 447-470, 1982.

[9] T. Kato, *Perturbation Theory for Linear Operators*, 2nd Edition, Springer-Verlag, Heidelberg, 1976.

[10] J. Milnor, *Morse Theory*, Annals of Math. Stud. 51, Princeton University Press, 1963.

[11] B. C. Moore, "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction," *IEEE Trans. on Autom. Contr.*, vol. AC-26, no. 1, pp. 17-32, 1981.

[12] A. P. Robertson and W. J. Robertson, *Topological Vector Spaces*, 2nd Edition, University Press, Cambridge, 1973.

[13] J. M. A. Scherpen, "Balancing for Nonlinear Systems," *Systems & Control Letters*, vol. 21, pp. 143-153, 1993.

[14] — "On the Similarity Invariance of Balancing for Nonlinear Systems," *Preprints IFAC Nonl. Contr. Syst. Des. (NOLCOS)*, Tahoe City, CA, pp. 783–788, 1995.

[15] J. M. A. Scherpen and W. S. Gray, “Sufficient Conditions for Minimality of a Nonlinear Realization via Controllability and Observability Functions,” *Proc. 1998 American Contr. Conf.*, Philadelphia, PA, pp. 3349-3353, 1998.

[16] —, “Minimality and Similarity Invariants of a Nonlinear State Space Realization,” *IEEE Trans. on Autom. Contr.*, to appear.

[17] J. M. A. Scherpen and A. J. van der Schaft, “Normalized Coprime Factorizations and Balancing for Unstable Nonlinear Systems,” *Int. J. of Contr.*, vol. 60, no. 6, pp. 1193-1222, 1994.

[18] K. Zhou, J. C. Doyle and K. Glover, *Robust and Optimal Control*, Prentice Hall, Upper Saddle River, NJ, 1996.