On Adjoints and Singular Value Functions for Nonlinear Systems

Jacquelien M.A. Scherpen
* Fac. ITS, Dept. of Electrical Eng., Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands.
J.M.A.Scherpen@its.tudelft.nl

Kenji Fujimoto*,**
** Department of Systems Science Graduate School of Informatics, Kyoto University, Uji, Kyoto 611-0011 Japan.
fujimoto@i.kyoto-u.ac.jp, K.Fujimoto@its.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

Keywords: nonlinear adjoint systems, adjoint operators, energy functions, singular value functions.

I. Introduction

Adjoint operators play an important role in linear systems theory. They provide duality between input and output. The properties with respect to input, e.g. controllability and stabilizability issues, of linear systems directly translate to the dual results with respect to output, observability and detectability issues. Consider a linear operator (transfer function) $\Sigma(s) : E \to F$ with Hilbert spaces E and F. Then its adjoint operator $\Sigma'(s) : F' \to E'$ is isomorphic to $\Sigma^T(-s) : F \to E$. The adjoint can be easily described by a state-space realization if the operator $\Sigma(s)$ has a finite dimensional state-space realization. In this paper we study the nonlinear extension of such adjoint operators, and apply the results to Hankel theory.

Nonlinear adjoint operators can be found in the mathematics literature, e.g. [1], and they are expected to play a similar role in the nonlinear systems theory. So called nonlinear Hilbert adjoint operators are introduced in [5, 11] as a special class of nonlinear adjoint operators. The existence of such operators in an input-output sense was shown in [6], but their state-space realizations are only preliminary available in [4], where the main interest is the Hilbert adjoint extension with an emphasis on the use of port-controlled Hamiltonian system methods.

Here, we consider these adjoint operators from a variational point of view and provide a formal justification for the use of Hamiltonian extensions by using Gâteaux derivatives. We investigate whether one can use the state-space realizations given by the Hamiltonian extensions to characterize singular values of nonlinear operators, and, in particular, for the Hankel operator. We also consider the relation with the previously defined singular value functions that have been defined entirely from the controllability and observability functions corresponding to a state space representation of a nonlinear system [10].

In Section 2 we present the linear system case as a paradigm, in order to present the line of thinking for the nonlinear case. In Section 3 we present the state-space realizations of nonlinear adjoint operators, in terms of Hamiltonian extensions. In Section 4 we provide the formal justification of the use of Hamiltonian extensions for nonlinear adjoint systems. In Section 5 we concentrate on the Hankel operator, and correspondingly on the controllability and observability operators for nonlinear systems. Then, in Section 6, we extend some results of the linear case on singular values, see e.g. [13], and their relation to the Hankel operator to the nonlinear case by using the state space realizations for adjoint systems as given in Section 3. Finally, some conclusions are given.

II. Linear systems as a paradigm

This section gives some examples of linear adjoint operators which play an important role in the linear systems theory, see e.g. [13]. They are presented in a way that clarifies the line of thinking in the nonlinear case. Consider a causal linear input-output system $\Sigma : L_2^m[0,\infty) \to L_2^r[0,\infty)$ with a state-space realization

$$u \mapsto y = \Sigma(u) : \left\{ \begin{array}{rcl} \dot{x} & = & Ax + Bu \\ y & = & Cx \end{array} \right. \quad (1)$$

where $x(0) = 0$. The Laplace transformation gives its transfer function matrix

$$G(s) := C(sI - A)^{-1}B. \quad (2)$$

Its adjoint operator is isomorphic to $\Sigma^* : L_2^r[0,\infty) \to L_2^m[0,\infty)$, where the transfer matrix is given by

$$G^*(s) := G^T(-s) = B^T(-sI - A^T)^{-1}C^T \quad (3)$$

with a state-space realization

$$u_a \mapsto y_a = \Sigma^*(u_a) : \left\{ \begin{array}{rcl} \dot{x} & = & -A^T x - C^T u_a \\ y_a & = & B^T x \end{array} \right. \quad (4)$$

where $x(\infty) = 0$. Here u_a and y_a have the same dimensions as y and u respectively. Σ^* satisfies the definition for Hilbert adjoint operators, namely,

$$\langle \Sigma(u), u_a \rangle_{L_2^r} = \langle u, \Sigma^*(u_a) \rangle_{L_2^m}. \quad (5)$$

Since u_a has the same dimension as y we obtain

$$\|\Sigma(u)\|_{L_2^r}^2 = \langle \Sigma(u), \Sigma(u) \rangle_{L_2^r} = \langle u, \Sigma^* \circ \Sigma(u) \rangle_{L_2^m}$$

by substituting $u_a = \Sigma(u)$. This relation can be utilized to derive the singular values of the input-output map.

Now, consider the Hankel operator of a continuous-time causal linear time-invariant input-output system $S : u \to y$ with an impulse response H which is analytic on $[0,\infty)$. $\hat{\mathcal{H}}_\Sigma = [\mathcal{H}_{\Sigma\ i,j}]$, where $\hat{\mathcal{H}}_{\Sigma\ i,j} = H_{i+j-1}$ for $i, j \geq 1$. Its rank is finite if and only if the corresponding transfer function is composed of strictly proper rational components [12]. If S is BIBO stable (take here to mean that $H \in L_1[0,\infty)$) then the system Hankel integral operator in this context is the well defined mapping

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

Define the *time flipping* operator as the injective mapping

$$\begin{array}{rcl} \mathcal{F} & : & L_2^m[0,\infty) \to L_2^m(-\infty,\infty) \\ & : & \hat{u} \to u(t) = \left\{ \begin{array}{rcl} \hat{u}(-t) & : & t < 0 \\ 0 & : & t \geq 0, \end{array} \right. \end{array}$$

then $\mathcal{H}_\Sigma = S\mathcal{F}$, where the codomain of S is restricted to $L_2^p[0,\infty)$.

It is well known that the composition $\mathcal{H}_\Sigma^*\mathcal{H}_\Sigma$ is a compact positive semi-definite self-adjoint operator with a well defined spectral decomposition [8]:

$$\mathcal{H}_\Sigma^*\mathcal{H}_\Sigma = \sum_{j=1}^{\infty} \sigma_j^2 \langle \cdot, v_j\rangle_{L_2} v_j, \quad \sigma_j \geq 0, \quad v_j \in L_2^m[0,\infty) \tag{6}$$

$$\langle v_j, v_k\rangle_{L_2} = \delta_{jk}, \quad \langle v_j, (\mathcal{H}_\Sigma^*\mathcal{H}_\Sigma)(v_j)\rangle_{L_2} = \sigma_j^2. \tag{7}$$

The nonnegative real numbers $\sigma_1 \geq \sigma_2 \geq \ldots$ are called the *Hankel singular values* for the input-output system S.

When there exists a finite integer n such that $\sigma_n \neq 0$ and $\sigma_j = 0$ for all $j > n$, or equivalently $rank(\hat{\mathcal{H}}_\Sigma) = rank(\mathcal{H}_\Sigma) = n$, then there exists a state space realization (A, B, C) of S with dimension n. Any such realization induces a factorization of the system Hankel matrix into the form $\hat{\mathcal{H}}_\Sigma = \hat{\mathcal{O}}_\Sigma\hat{\mathcal{C}}_\Sigma$, where $\hat{\mathcal{O}}_\Sigma$ and $\hat{\mathcal{C}}_\Sigma$ are the (extended) observability and controllability matrices. If the realization is asymptotically stable (i.e., A is Hurwitz) then the Hankel operator can be written as the composition of uniquely determined observability and controllability operators; that is, $\mathcal{H}_\Sigma = \mathcal{O}_\Sigma\mathcal{C}_\Sigma$, where the observability and controllability operators, $\mathcal{O}_\Sigma : \mathbb{R}^n \to L_2^r[0,\infty)$ and $\mathcal{C}_\Sigma : L_2^m[0,\infty) \to \mathbb{R}^n$, respectively, are given by

$$x^0 \mapsto y = \mathcal{O}_\Sigma(x^0) \quad := \quad Ce^{At}x^0 \tag{8}$$

$$u \mapsto x^0 = \mathcal{C}_\Sigma(u) \quad := \quad \int_0^\infty e^{A\tau}Bu(\tau)d\tau. \tag{9}$$

Note that these operators $\mathcal{O}_\Sigma$ and $\mathcal{C}_\Sigma$ are also operators on Hilbert spaces, hence their adjoint operators are given by $\mathcal{O}_\Sigma^* : L_2^m[0,\infty) \to \mathbb{R}^n$ and $\mathcal{C}_\Sigma^* : \mathbb{R}^n \to L_2^r[0,\infty)$

$$u_a \mapsto x^0 = \mathcal{O}_\Sigma^*(u_a) \quad := \quad \int_0^\infty e^{A^T\tau}C^T u_a(\tau)d\tau \tag{10}$$

$$x^0 \mapsto y_a = \mathcal{C}_\Sigma^*(x^0) \quad := \quad B^T e^{A^T t}x^0. \tag{11}$$

It can be easily checked that they satisfy

$$\langle \mathcal{O}_\Sigma(x^0), u_a\rangle_{L_2^r} \quad = \quad \langle x^0, \mathcal{O}_\Sigma^*(u_a)\rangle_{\mathbb{R}^n} \tag{12}$$

$$\langle \mathcal{C}_\Sigma(u), x^0\rangle_{\mathbb{R}^n} \quad = \quad \langle u, \mathcal{C}_\Sigma^*(x^0)\rangle_{L_2^m}. \tag{13}$$

These adjoint operators can be used to calculate the observability and controllability Gramian, respectively:

$$\begin{aligned} \|\mathcal{O}_\Sigma(x^0)\|_{L_2^r}^2 &= \langle x^0, \mathcal{O}_\Sigma^* \circ \mathcal{O}_\Sigma(x^0)\rangle_{\mathbb{R}^n} \\ &= \langle x^0, \textstyle\int_0^\infty CA^\tau A^{T\tau}C^T d\tau\ x^0\rangle_{\mathbb{R}^n} \qquad (14) \\ &= \langle x^0, Q\ x^0\rangle_{\mathbb{R}^n} \\ \|\mathcal{C}_\Sigma^*(x^0)\|_{L_2^m}^2 &= \langle x^0, \mathcal{C}_\Sigma^{**} \circ \mathcal{C}_\Sigma^*(x^0)\rangle_{\mathbb{R}^n} \qquad (15) \\ &= \langle x^0, \textstyle\int_0^\infty B^T A^{T\tau}A^\tau B d\tau\ x^0\rangle_{\mathbb{R}^n} \qquad (16) \\ &= \langle x^0, P\ x^0\rangle_{\mathbb{R}^n} \end{aligned}$$

These imply $Q = \mathcal{O}_\Sigma^* \circ \mathcal{O}_\Sigma$ and $P = \mathcal{C}_\Sigma^{**} \circ \mathcal{C}_\Sigma^* = \mathcal{C}_\Sigma \circ \mathcal{C}_\Sigma^*$.

Furthermore, it is known that

Lemma II.1 *[13] The operator $\mathcal{H}_\Sigma^*\mathcal{H}_\Sigma$ and the matrix QP have the same nonzero eigenvalues.*

III. State-space realization of nonlinear Hilbert adjoint operators

This section is devoted to the state-space characterization of nonlinear Hilbert adjoint operators as an extension of the properties given in the previous section. We will show a relationship between nonlinear Hilbert adjoint operators and Hamiltonian extensions.

The precise definition of *nonlinear Hilbert adjoint* operators is given as follows [5, 6, 11].

Definition III.1 *Consider an operator $T : E \to F$ with Hilbert spaces E and F. An operator $T^* : F \times E \to E$ such that*

$$\langle T(u), y\rangle_F = \langle u, T^*(y,u)\rangle_E, \ \forall u \in E, \ \forall y \in F \tag{17}$$

holds is said to be a nonlinear Hilbert adjoint of T.

Remark III.2 In the most general setting, let F be a topological vector space over $\mathbb{R}$ with dual space F' [1]. 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_A^β 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 $T : E \to F$, the *dual map* of T is defined as

$$\begin{aligned} T' \ &: \ F' \to E^\beta \\ &: \ y' \mapsto (T'(y'))(u) = (y' \circ T)(u) \qquad (18) \\ &\quad \forall u \in E, \ \forall y \in F. \end{aligned}$$

Hence a nonlinear Hilbert adjoint operator T^* yields an adjoint operator in the usual sense by

$$(T'(y'))(u) := \langle u, T^*(y,u)\rangle_E, \quad u \in E, \ y \in F. \tag{19}$$

The converse result can be found in [6].

If T is a linear operator then T^* always exists and is equivalent to T'. Of course T^* is a function only of F, i.e.

$$\langle T(u), y\rangle_F = \langle u, T^*(y)\rangle_E, \ \forall u \in E, \ \forall y \in F \tag{20}$$

in the previous section.

Adjoint operators and Hamiltonian extensions

This subsection gives some relations between nonlinear Hilbert adjoint operators and Hamiltonian extensions. Let us consider an input-output system $\Sigma : L_2^m(\Omega) \to L_2^r(\Omega)$ defined on a (possibly infinite) time interval $\Omega = [t^0, t^1] \subseteq \mathbb{R}$ which has a state-space realization

$$u \mapsto y = \Sigma(u) : \left\{ \begin{array}{rcl} \dot{x} &=& f(x,u) \quad x(t^0) = 0 \\ y &=& h(x,u) \end{array} \right. \tag{21}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^r$. Here we assume the origin is an equilibrium, i.e. $f(0,0) = 0$, $h(0,0) = 0$ holds, and that all signals and functions are sufficiently smooth.

Before giving the Hamiltonian extension of Σ, we have to introduce the variational system of Σ. It is given by

$$(u, u_v) \mapsto y_v = \Sigma_v(u, u_v) : \left\{ \begin{array}{rcl} \dot{x} &=& f(x,u) \\ \dot{x}_v &=& \frac{\partial f}{\partial x}x_v + \frac{\partial f}{\partial u}u_v \\ y_v &=& \frac{\partial h}{\partial x}x_v + \frac{\partial h}{\partial u}u_v \end{array} \right. \tag{22}$$

with $x(t^0) = 0$ and $x_v(t^1) = 0$. The input-state-output set (u_v, x_v, y_v) are the so called variational input, state, and output, respectively, and they represent the variation along the trajectory (u, x, y) of the original system Σ.

The Hamiltonian extension Σ_a of Σ is given by a Hamiltonian control system [2] which has an adjoint form of the variational system. It is given by

$$(u, u_a) \mapsto y_a = \Sigma_a(u, u_a) :$$
$$\begin{cases} \dot{x} &= \frac{\partial H}{\partial p}^T &= f(x,u) \\ \dot{p} &= -\frac{\partial H}{\partial x}^T &= -\left(\frac{\partial f}{\partial x}^T p + \frac{\partial h}{\partial x}^T u_a\right) \\ y_a &= \frac{\partial H}{\partial u}^T &= \frac{\partial f}{\partial u}^T p + \frac{\partial h}{\partial u}^T u_a \\ y &= \frac{\partial H}{\partial u_a}^T &= h(x,u) \end{cases} \quad (23)$$

with $x(t^0) = 0$, $p(t^1) = 0$, and with the Hamiltonian

$$H(x, p, u, u_a) := p^T f(x,u) + u_a^T h(x,u). \quad (24)$$

Remark III.3 In Section 4, we show that such a Hamiltonian control system is a realization of the Gâteaux derivative of the adjoint of the operator. This interpretation results from taking the Gâteaux derivative from the squared L_2 norm of the nonlinear operator. Therefore, it is a more restricted interpretation than is given above by the Hilbert adjoint definition in terms of the inner product.

By careful consideration of the Hamiltonian, we can relate the Hamiltonian extension idea to the Hilbert adjoint as follows (for more details, see [4]):

Theorem III.4 *[4] Consider the system Σ as in (21) and let $\Sigma : L_2^m(\Omega) \to L_2^r(\Omega)$ where $\Omega = [t^0, t^1] \subseteq \mathbb{R}$ denotes the mapping $u \mapsto y$. Suppose f and h are input-affine, i.e. $f(x,u) \equiv g_0(x) + g(x)u$ and $h(x,u) \equiv k_0(x) + k(x)u$ for some smooth functions g_0, g, k_0 and k. Suppose moreover that*

$$\begin{aligned} &u \in L_2^m(\Omega),\ u_b \in L_2^r(\Omega) \quad \Rightarrow \\ &|x(t^1)| < \infty,\ |p_1(t^0)| < \infty,\ |p_2(t^0)| < \infty \end{aligned} \quad (25)$$

for the state-space system

$$(u_b, u) \mapsto y_b = \Sigma^*(u_b, u) :$$
$$\begin{cases} \dot{x} &= g_0(x) + g(x)u \\ \dot{p}_1 &= -\frac{\partial g_0}{\partial x}^T p_1 - \frac{\partial k_0}{\partial x}^T p_2 \\ \dot{p}_2 &= u_b \\ y_b &= \left(\frac{\partial (g^T p_1)}{\partial x} + \frac{\partial (k^T p_2)}{\partial x}\right) g_0(x) \\ & \quad -g^T(x)\left(\frac{\partial g_0}{\partial x}^T p_1 + \frac{\partial k_0}{\partial x}^T p_2\right) + k^T(x) u_b \\ y &= k_0(x) + k(x)u \end{cases} \quad (26)$$

with $x(t^0) = 0$, $p_1(t^1) = 0$ and $p_2(t^1) = 0$. Then a state-space realization of the nonlinear Hilbert adjoint $\Sigma^* : L_2^{m+r}(\Omega) \to L_2^m(\Omega)$ of Σ is given by (26).

There also exists a relation between adjoint operators and port-controlled Hamiltonian systems, as has been established in [4]. Then, instead of the interpretation in terms of the Gâteaux derivative of the norm, the interpretation is more general, and can be given in terms of the Hilbert adjoint and the inner product. Despite this more general interpretation for the port-controlled case, we only consider here the Hamiltonian extensions as defined in [2], since we then have explicit solutions for the "dual" coordinates p of the system. Much more can be said about port-controlled Hamiltonian systems, however, that falls beyond the scope of this paper, and we refer to [4] for more details.

IV. Gâteaux differentiation of dynamical systems

This section develops the concept of Gâteaux differentiation for dynamical systems from an input-output point of view. In Remark III.3 we mention that it is of importance for understanding the meaning of the Hamiltonian extensions and adjoint systems as presented in the previous section. Also, Gâteaux differentiation of Hankel operators plays an important role in the analysis of the properties of Hankel operators, which is the topic of Section 5 and 6. To this end, we state the definition of Gâteaux differentiation.

Definition IV.1 *(Gâteaux differentiation) Suppose X and Y are Banach spaces, $U \subseteq X$ is open, and $T : U \to Y$. Then T has a Gâteaux derivative at $x \in X$ if, for all $\xi \in U$ the following limit exits:*

$$dT(x)(\xi) = \lim_{\varepsilon \to 0} \frac{T(x + \varepsilon\xi) - T(x)}{\varepsilon} = \frac{d}{d\varepsilon} T(x + \varepsilon\xi)|_{\varepsilon=0}. \quad (27)$$

We write $dT(x)(\xi)$ for the Gâteaux derivative of T at x in the "direction" ξ.

Next, we state the chain rule of Gâteaux derivative for convenience.

Lemma IV.2 *The derivative of a composition is given by the following equation:*

$$d(T \circ S)(x)(\xi) \quad = \quad dT(S(x), dS(x)(\xi)). \quad (28)$$

Perhaps more well-known than the Gâteaux derivative is the Fréchet derivative, which is especially useful for the analysis of nonlinear static functions. Fréchet differentiation is a special case of Gâteaux differentiation. In the sequel, we concentrate on Gâteaux differentiation, since that is the most suitable in our framework.

Theorem IV.3 *Suppose that $\Sigma : u \mapsto y$ as in (21) is input-affine and has no direct feed-through, i.e. $f(x,u) \equiv g_0(x) + g(x)u$ and $h(x,u) \equiv h(x)$ for some analytic functions g_0, g and h. Furthermore, suppose that Σ is Gâteaux differentiable, namely that there exists a neighborhood $U_v \subseteq L_2^m(\Omega)$ of 0 such that*

$$u \in L_2^m(\Omega),\ u_v \in U_v \Rightarrow y_v \in L_2^r(\Omega). \quad (29)$$

Then it follows that

$$\Sigma_v(u, v) = d\Sigma(u)(v) \quad (30)$$

with the variational system Σ_v given in (22).

In order to prove Theorem IV.3, we need the following property of variational systems.

Lemma IV.4 *[2] Let $(x(t,\varepsilon), u(t,\varepsilon), y(t,\varepsilon))$, $t \in [a,b]$ be a family of state-input-output trajectories of Σ, parameterized by ε, such that $x(t,0) = x(t)$, $u(t,0) = u(t)$ and $y(t,0) = y(t)$, $t \in [a,b]$. Then the quantities*

$$x_v(t) = \frac{\partial x(t,0)}{\partial \varepsilon} \quad (31)$$

$$u_v(t) = \frac{\partial u(t,0)}{\partial \varepsilon} \quad (32)$$

$$y_v(t) = \frac{\partial y(t,0)}{\partial \varepsilon} \quad (33)$$

satisfy $y_v = \Sigma_v(u, u_v)$.

Note that in case of a fixed initial state $x(0) = x^0$ the variational state $x_v(0)$ at time 0 is necessarily 0. Now, we can give the proof of Theorem IV.3.

Proof of Theorem IV.3 Let $u(t,\varepsilon) = u(t) + \varepsilon v(t)$ in Lemma IV.4. Then we have

$$\begin{aligned} \Sigma(u+\varepsilon v)(t) &= y(t,\varepsilon) \\ &= y(t,0) + \frac{\partial y(t,0)}{\partial \varepsilon}\varepsilon + \sum_{i=2}^{\infty} \frac{1}{i!}\frac{\partial^i y(t,0)}{\partial \varepsilon^i}\varepsilon^i \\ &= \Sigma(u)(t) + \Sigma_v(u,v)(t)\ \varepsilon + \sum_{i=2}^{\infty} R_i(u,v)(t)\ \varepsilon^i \end{aligned}$$

where $R_i(u,v)(t) := \frac{1}{i!}\frac{\partial^i y(t,0)}{\partial \varepsilon^i}$. This implies

$$\begin{aligned} (d\Sigma(u)(v))(t) &= \lim_{\varepsilon \to 0} \frac{\Sigma(u+\varepsilon v)(t) - \Sigma(u)(t)}{\varepsilon} \\ &= \lim_{\varepsilon \to 0} \left(\Sigma_v(u,v)(t) + \sum_{i=2}^{\infty} R_i(u,v)(t)\ \varepsilon^{i-1} \right) \\ &= \Sigma_v(u,v)(t). \end{aligned}$$

This proves the theorem. □

The Hamiltonian extension Σ_a also has a relation with Gâteaux differentiation and provides a justification for being called the adjoint form of the variational system in [2].

Theorem IV.5 *Suppose that the assumptions in Theorem IV.3 hold, and that $u \in L_2^m(\Omega)$, $u_a \in L_2^r(\Omega) \Rightarrow \|x(t^1)\| < \infty$, $\|p(t^0)\| < \infty$. Then it follows that*

$$\Sigma_a(u,v) = (d\Sigma(u))^*(v) \quad (34)$$

with the Hamiltonian extension Σ_a given in (23).

The fact that the Hamiltonian extension $\Sigma_a(u,v)$ is linearly dependent on v is crucial to prove Theorem IV.5. A more general version, related to the Hilbert adjoint definition, can be derived from the differential version of Proposition 2 in [4], but falls beyond the scope of this paper.

V. The Hankel operator and its derivative

This section gives a state-space realization for the nonlinear Hilbert adjoint of some particular energy functions and operators, namely the observability and controllability functions and operators and the Hankel operator. Furhtermore, a relation with singular value analysis of the Hankel operator is given. We only consider time invariant input-affine nonlinear systems without direct feed-through in the form of

$$\Sigma : \left\{ \begin{array}{rcl} \dot{x} &=& f(x) + g(x)u \\ y &=& h(x) \end{array} \right. \quad (35)$$

defined on the time interval $\Omega := (-\infty, \infty)$. Here Σ is L_2-stable in the sense that $u \in L_2^m(-\infty, 0]$ implies that $\Sigma(u)$ restricted to $[0,\infty)$ is in $L_2^r[0,\infty)$. Suppose that the input-output mapping $u \mapsto y$ of this system can be described by a Chen-Fliess functional expansion [3, 7], i.e. the mapping $u \mapsto y$ is represented by the following convergent generating series

$$u \mapsto y(t) = \sum_{\eta \in I^*} c(\eta) E_\eta(t,t^0)(u), \quad t \geq t^0 \quad (36)$$

where I^* is the set of multi-indices for the index set $I = \{0, 1, \ldots, m\}$ and

$$E_{i_k \ldots i_0}(t,t^0)(u) = \int_{t^0}^{t} u_{i_k}(\tau) E_{i_{k-1}\ldots i_0}(\tau, t^0)(u) d\tau \quad (37)$$

with $E_\emptyset(u) := 1$ and $u_0(t) := 1$. Here $c(\eta) \in \mathbb{R}^r$ is described by

$$c(\eta) = L_{g_\eta} h(0) := L_{g_{i_0}} L_{g_{i_1}} \ldots L_{g_{i_k}} h(0) \quad (38)$$

with $g_0 := f$. Let us consider the observability and controllability operators $\mathcal{O}_\Sigma : \mathbb{R}^n \to L_2^r(\Omega_+)$ and $\mathcal{C}_\Sigma : L_2^m(\Omega_+) \to \mathbb{R}^n$ with $\Omega_+ := [0,\infty)$ of Σ given in [5, 6, 11] which are defined by

$$x^0 \mapsto y(t) = \mathcal{O}_\Sigma(x^0) := \sum_{i=0}^{\infty} L_{g_0}^i h(x^0) E_{\underbrace{0\ldots0}_{i}}(t,0) \quad (39)$$

$$u \mapsto x^1 = \mathcal{C}_\Sigma(u) := \sum_{\eta \in I^*} (L_{g_\eta} x)(0) E_\eta(0,-\infty)\mathcal{F}_-(u). \quad (40)$$

Here $\mathcal{F}_- : L_2^m(\Omega_+) \to L_2^m(\Omega_-)$ with $\Omega_- := (-\infty, 0]$ denotes the so called *flipping operator* defined by

$$\mathcal{F}_-(u)(t) := \left\{ \begin{array}{cl} u(-t) & t \in \Omega_- \\ 0 & t \in \Omega_+ \end{array} \right. . \quad (41)$$

These are a natural generalization of the linear case (8) and (9).

One can employ state-space systems to describe the observability and controllability operators, which are operators of the form $\mathbb{R}^n \to L_2^r$ and $L_2^m \to \mathbb{R}^n$. Specifically, their state-space realizations are given by

$$x^0 \mapsto y = \mathcal{O}_\Sigma(x^0) \quad : \quad \left\{ \begin{array}{rcl} \dot{x} &=& f(x) \\ y &=& h(x) \end{array} \right. \quad (42)$$

$$u \mapsto \tilde{x}^1 = \mathcal{C}_\Sigma(u) \quad : \quad \left\{ \begin{array}{rcl} \dot{\tilde{x}} &=& f(\tilde{x}) + g(\tilde{x})\mathcal{F}_-(u) \\ \tilde{x}^1 &=& \tilde{x}(0) \end{array} \right. \quad (43)$$

where $x(0) = x^0$ and $\tilde{x}(-\infty) = 0$. Furthermore, the Hankel operator $\mathcal{H}_\Sigma : L_2^m(\Omega_+) \to L_2^r(\Omega_+)$ of Σ is given by

$$\mathcal{H}_\Sigma := \Sigma \circ \mathcal{F}_-. \quad (44)$$

and $\mathcal{H}_\Sigma = \mathcal{O}_\Sigma \circ \mathcal{C}_\Sigma$ holds. This has been proven in [5, 6], along with a deeper and more detailed analysis of the Hankel operator. We can state the differential version of this fact using Lemma IV.2 as

$$d\mathcal{H}_\Sigma(u)(u_v) = d\mathcal{O}_\Sigma(\mathcal{C}_\Sigma(u))(d\mathcal{C}_\Sigma(u)(u_v)). \quad (45)$$

The state-space realizations of the Gâteaux differentiations $d\mathcal{O}_\Sigma$, $d\mathcal{C}_\Sigma$ and $d\mathcal{H}_\Sigma$ are then characterized by the following theorem.

Theorem V.1 *Consider the system Σ, and suppose the assumptions of Theorem IV.3 hold. Then*

$$\begin{aligned} d\mathcal{O}_\Sigma &= \mathcal{O}_{d\Sigma} \\ d\mathcal{C}_\Sigma &= \mathcal{C}_{d\Sigma} \\ d\mathcal{H}_\Sigma &= \mathcal{H}_{d\Sigma}. \end{aligned}$$

This theorem directly follows from the definition of $\mathcal{O}_\Sigma$, $\mathcal{C}_\Sigma$, $\mathcal{H}_\Sigma$ and the Gâteaux derivative $d(\cdot)$. Furthermore their adjoints can be obtained by using Theorem IV.5.

Theorem V.2 *Consider the operator Σ as in (35). Suppose that the assumptions of Theorem IV.3 and Theorem IV.5 hold. Then state-space realizations of $(d\mathcal{O}_\Sigma(x^0))^* : L_2^r(\Omega_+)(\times\mathbb{R}^n) \to \mathbb{R}^n$, $(d\mathcal{C}_\Sigma(u))^* : \mathbb{R}^n(\times L_2^m(\Omega_+)) \to L_2^m(\Omega_+)$ and $(d\mathcal{H}_\Sigma(u))^* : L_2^r(\Omega_+)(\times L_2^m(\Omega_+)) \to L_2^m(\Omega_+)$ are given by*

$$(x^0, u_a) \mapsto p^0 = (d\mathcal{O}_\Sigma(x^0))^*(u_a):$$

$$\left\{ \begin{array}{rcl} \dot{x} &=& f(x) \\ \dot{p} &=& -\frac{\partial f}{\partial x}^T(x)\ p - \frac{\partial h}{\partial x}^T(x)\ u_a \\ p^0 &=& p(0) \end{array} \right. \tag{46}$$

with $x(0) = x^0$ and $p(\infty) = 0$,

$$(p^1, u) \mapsto y_a = (d\mathcal{C}_\Sigma(u))^*(p^1):$$

$$\left\{ \begin{array}{rcl} \dot{x} &=& f(x) + g(x)\mathcal{F}_-(u) \\ \dot{p} &=& -\frac{\partial f}{\partial x}^T(x)\ p \\ y_a &=& \mathcal{F}_+(g^T(x)\ p) \end{array} \right. \tag{47}$$

with $x(-\infty) = 0$ and $p(0) = p^1$,

$$(u_a, u) \mapsto y_a = (d\mathcal{H}_\Sigma(u))^*(u_a):$$

$$\left\{ \begin{array}{rcl} \dot{x} &=& f(x) + g(x)\ \mathcal{F}_-(u) \\ \dot{p} &=& -\frac{\partial f}{\partial x}^T(x)\ p - \frac{\partial h}{\partial x}^T(x)\ u_a \\ y_a &=& \mathcal{F}_+(g^T(x)\ p) \end{array} \right. \tag{48}$$

with $x(-\infty) = 0$ and $p(\infty) = 0$, respectively. Here $\mathcal{F}_+ : L_2^m(\Omega_-) \to L_2^m(\Omega_+)$ denotes another time flipping operator defined by

$$\mathcal{F}_+(u)(t) := \left\{ \begin{array}{cc} 0 & t \in \Omega_- \\ u(-t) & t \in \Omega_+ \end{array} \right. . \tag{49}$$

The proof of this theorem is easily obtained by applying the adjoint Hamiltonian extensions of Section 3 and using techniques from [4].

VI. Energy functions and singular values

Define the following energy functions of a system.

Definition VI.1 *The observability function $L_o(x)$ and the controllability function $L_c(x)$ of Σ as in (35) are defined by*

$$L_o(x^0) := \frac{1}{2}\int_0^\infty \|y(t)\|^2 dt,\ x(0) = x^0,\ u(t) \equiv 0 \tag{50}$$

$$L_c(x^1) := \min_{\substack{u \in L_2^m(\Omega_-) \\ x(-\infty) = 0 \\ x(0) = x^1}} \frac{1}{2}\int_{-\infty}^0 \|u(t)\|^2 dt \tag{51}$$

respectively.

These functions are closely related to the observability and controllability operators and Gramians in the linear case. In [10] these functions have been used for defining balanced realizations and singular value functions of nonlinear systems. They also fulfill certain Hamilton-Jacobi equations, in a similar way as the observability Gramian and the inverse of the controllability Gramian are solutions of a Lyapunov/Riccati equation. In order to proceed, we first review what is meant by input-normal/output-diagonal form, see [10]:

Theorem VI.2 *[10] Consider a system (f, g, h) that fulfills certain technical conditions. Then there exists on a neighborhood $U \subset V$ of 0, a coordinate transformation $x = \psi(z)$, $\psi(0) = 0$, which converts the system into an input-normal/output-diagonal form, where*

$$\begin{aligned} \tilde{L}_c(z) &:= L_c(\psi(z)) = \frac{1}{2}z^T z, \\ \tilde{L}_o(z) &:= L_o(\psi(z)) = \frac{1}{2}z^T diag(\tau_1(z), \ldots, \tau_n(z))z \end{aligned}$$

with $\tau_1(z) \geq \ldots \geq \tau_n(z)$ being the so called smooth singular value functions on $W := \psi^{-1}(U)$.

The relation between the observability function, operator and Gramian is

$$\begin{aligned} L_o(x^0) &= \frac{1}{2}\|\mathcal{O}_\Sigma(x^0)\|^2_{L_2^r} \\ &= \frac{1}{2}\langle \mathcal{O}_\Sigma(x^0), \mathcal{O}_\Sigma(x^0)\rangle_{L_2^r} \\ &= \frac{1}{2}\langle x^0, \mathcal{O}_\Sigma^*(\mathcal{O}_\Sigma(x^0), x^0)\rangle_{\mathbb{R}^n} \\ &=: \langle x^0, \phi(x^0)\rangle_{\mathbb{R}^n}. \end{aligned} \tag{52}$$

The function $\phi(x^0)$ can always be rewritten as $\phi(x^0) = Q(x^0)\, x^0$ using a square symmetric matrix $Q(x^0)$. This matrix coincides with the observability Gramian in the linear case.

In the controllability case, there does not hold such a simple relation. Instead, it follows that

$$\begin{aligned} L_c(x^1) &= \frac{1}{2}\|\mathcal{C}_\Sigma^\dagger(x^1)\|^2_{L_2^m} \\ &= \frac{1}{2}\langle x^1, \mathcal{C}_\Sigma^{\dagger\,*}(\mathcal{C}_\Sigma^\dagger(x^1), x^1)\rangle_{\mathbb{R}^n} \\ &=: \frac{1}{2}\langle x^1, \varphi(x^1)\rangle_{\mathbb{R}^n} \end{aligned} \tag{53}$$

with $\mathcal{C}_\Sigma^\dagger : \mathbb{R}^n \to L_2^m(\Omega_+)$, which is the pseudo-inverse of $\mathcal{C}_\Sigma$ defined by

$$\mathcal{C}_\Sigma^\dagger(x^1) := \arg\min_{\mathcal{C}_\Sigma(u) = x^1} \|u\|_{L_2^m}. \tag{54}$$

Now, we can state the result from [5, 6] that relates the singular value functions to the Hankel operator:

Theorem VI.3 *[5] Let (f, g, h) be an analytic n dimensional input-normal/output-diagonal realization of a causal L_2-stable input-output mapping S on a neighborhood W of 0. Define on W the collection of component vectors $\tilde{z}_j = (0, \ldots, 0, z_j, 0, \ldots, 0)$ for $j = 1, 2, \ldots, n$, and the functions $\hat{\sigma}^2(z_j) = \tau(\tilde{z}_j)$. Let v_j be the minimum energy input which drives the state from $z(-\infty) = 0$ to $z(0) = \tilde{z}_j$ and define $\hat{v}_j = \mathcal{F}(v_j)$. Then the functions $\{\hat{\sigma}_j\}_{j=1}^n$ are singular value functions of the Hankel operator $\mathcal{H}_\Sigma$ in the following sense:*

$$\langle \hat{v}_j, (\mathcal{H}_\Sigma^*\mathcal{H}_\Sigma)(\hat{v}_j)\rangle_{L_2} = \hat{\sigma}_j^2(z_j)\langle \hat{v}_j, \hat{v}_j\rangle_{L_2}, \quad j = 1, 2, \ldots n. \tag{55}$$

The above result is quite limited in the sense that it is dependent on the input-normal/output-diagonal coordinate frame To give a more general relation, the idea is to extend Lemma II.1, by extending the proof of that lemma as given in [13]. To this effect, we consider the Gâteaux derivative of the Hankel operator in the following way

$$\begin{aligned} d\|\mathcal{H}_\Sigma(u)\|_2^2(v) &= 2\ \langle d\mathcal{H}_\Sigma(u,v), \mathcal{H}_\Sigma(u)\rangle & (56)\\ &= 2\ \langle v, (d\mathcal{H}_\Sigma(u))^* \circ \mathcal{H}_\Sigma(u)\rangle & (57)\end{aligned}$$

and consider the eigenstructure of the operator $u \mapsto (d\mathcal{H}_\Sigma(u))^* \circ \mathcal{H}_\Sigma(u)$ as

$$(d\mathcal{H}_\Sigma(u))^* \circ \mathcal{H}_\Sigma(u) = \lambda(u)u \tag{58}$$

where $\lambda(u)$ is an eigenvalue, and u the corresponding eigenvector. However, since we want to relate it to the notion of singular value functions, which depend on x^0, an additional step is needed. Therefore, we propose to consider the eigenvalues $\tilde\sigma(x^0)$ and corresponding eigenvectors x^0 of the following:

$$\begin{aligned} &\mathcal{C}_\Sigma \circ d\mathcal{H}_\Sigma^* \circ \mathcal{H}_\Sigma(u) = \mathcal{C}_\Sigma \circ d\mathcal{H}_\Sigma^* \circ \mathcal{O}_\Sigma(x^0) = \tilde\sigma(x^0)x^0, \\ &\mathcal{C}_\Sigma(u) = x^0 \end{aligned} \tag{59}$$

We obtain the following result:

Theorem VI.4 *Assume all technical conditions for Theorem VI.2 are fulfilled. Let* $\phi(\tilde x) := \frac{\partial^T L_c}{\partial \tilde x}(\tilde x) = M_c(\tilde x)\tilde x$, *for* $\tilde x \in W$ *such that* M_c *is invertible on* W, *then*

$$\begin{aligned} \mathcal{C}_\Sigma \circ d\mathcal{H}_\Sigma^* \circ \mathcal{H}_\Sigma(u) &= \mathcal{C}_\Sigma \circ d\mathcal{C}_\Sigma^* \circ d\mathcal{O}_\Sigma^* \circ \mathcal{O}_\Sigma(x^0) \\ &= \mathcal{C}_\Sigma(\lambda(u)u) & (60) \\ &= M_c(\psi(x_0))^{-1}\frac{\partial L_o}{\partial x}(x^0) \end{aligned}$$

for $x^0 = \mathcal{C}_\Sigma(u)$, *and* $\psi(x^0) = \phi^{-1}\left(\frac{\partial^T L_o}{\partial x}(x^0)\right)$.

Proof: First, observe that the solution of system (46) is given by $p = \frac{\partial^T L_o}{\partial x}(x)$, where x is the solution of system (42), and $u_a = y = h(x)$. Thus,

$$p^0 = d\mathcal{O}_\Sigma \circ \mathcal{O}_\Sigma(x^0) = \frac{\partial^T L_o}{\partial x}(x^0).$$

Furthermore, observe that $\tilde p = \frac{\partial^T L_c}{\partial \tilde x}(\tilde x)$ is the solution of system (47), where $\tilde x$ is the solution of system (43) and where $u = y_a = \mathcal{F}_+(g^T(\tilde x)p)$. Thus,

$$\tilde x^1 = \mathcal{C}_\Sigma \circ d\mathcal{C}_\Sigma^*(p^0) = \left(M_c \underbrace{\left(\phi^{-1}\left(\frac{\partial^T L_o}{\partial x}(x^0)\right)\right)}_{(\psi(x^0))}\right)^{-1} p^0.$$

This proves the theorem. □

Remark VI.5 It is straightforward to obtain that the above theorem applied to a linear system yields $M_c(\psi(x^0))^{-1} = P$, where P is the controllability Gramian, and $\frac{\partial L_o}{\partial x}(x^0) = Qx^0$, where Q is the observability Gramian. Hence, the above theorem can be seen as a nonlinear extension of the proof of Lemma II.1, which has been given in [13]

By taking x^0 to be an eigenvector of the above operator, we obtain the relation (59). Observe that the $\tilde\sigma(x^0)$'s do not equal the singular value functions as defined in Theorem VI.2, due to the fact that here we deal with the gradients of the controllability and observability functions, instead of the functions themselves.

VII. Conclusions

We studied the use of Hamiltonian extensions for the nonlinear adjoint systems. We formalized the basic concepts, and then applied them to study the singular value functions of the nonlinear Hankel operator. In future research, we use these results to establish more direct relations between state space notions stemming from energy functions and input-output notions like the Hankel operator.

References

[1] Batt, J. (1970). Nonlinear compact mappings and their adjoints. *Math. Ann.*, 189,pp. 5–25.

[2] Crouch, P. E. and A. J. van der Schaft (1987). *Variational and Hamiltonian Control Systems.* Vol. 101 of *Lecture Notes on Control and Information Science*. Springer-Verlag. Berlin.

[3] Fliess, M. (1974). Matrices de Hankel. *J. Math. Pures. Appl.* pp. 197–222.

[4] Fujimoto, K., J.M.A. Scherpen, and W.S. Gray (2000). Hamiltonian realizations of nonlinear adjoint operators. *to appear at the IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control*, Princeton, March 2000.

[5] Gray, W. S. and J. M. A. Scherpen (1998). Hankel operators and Gramians for nonlinear systems. *Proc. 37th IEEE Conf. on Decision and Control* pp. 3349–3353.

[6] Gray, W. S. and J. M. A. Scherpen (1999). Hankel operators, sigular value functions and Gramian generalizations for nonlinear systems. Submitted.

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

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

[9] Maschke, B. M. J. and A. J. van der Schaft (1992). Port-controlled hamiltonian systems: modelling origins and system-theoretic properties. *IFAC Symp. NOLCOS* pp. 282–288.

[10] Scherpen, J. M. A. (1993). Balancing for nonlinear systems. *Systems & Control Letters*, 21, pp. 143–153.

[11] Scherpen, J. M. A. and W. S. Gray (1999). On singular value functions and Hankel operators for nonlinear systems. *Proc. ACC'99*, pp. 2360–2364.

[12] E. D. Sontag, "Linear systems over commutative rings: a survey," *Ricerche di Automatica*, vol. 7, no. 1, pp. 1-34, 1976.

[13] Zhou, K., J. C. Doyle and K. Glover (1996). *Robust and Optimal Control.* Prentice-Hall, Inc.. Upper Saddle River, N.J.