Hamiltonian Extensions, Hilbert Adjoints and Singular Value Functions for Nonlinear Systems

Scherpen, Jacquelien M.A.; Fujimoto, Kenji; Gray, W. Steven

Proceedings of the 39th IEEE
Conference on Decision and Control
Sydney, Australia • December, 2000

Hamiltonian Extensions, Hilbert Adjoints and Singular Value Functions for Nonlinear Systems

Jacquelien M.A. Scherpen*, Kenji Fujimoto**, W. Steven Gray***

* Fac. ITS, Dept. of Electrical Eng., Delft University of Technology,
P.O. Box 5031, 2600 GA Delft, The Netherlands, E-mail: J.M.A.Scherpen@its.tudelft.nl
** Dept. of Systems Science, Graduate School of Informatics, Kyoto University,
Uji, Kyoto 611-0011 Japan, E-mail: fujimoto@i.kyoto-u.ac.jp
***Dept. of Electrical and Computer Engineering, Old Dominion University,
Norfolk, Virginia 23529-0246, U.S.A., E-mail: gray@ece.odu.edu

Abstract

This paper studies previously developed nonlinear Hilbert adjoint operator theory from a variational point of view and provides a formal justification for the use of Hamiltonian extensions via Gâteaux differentials. The primary motivation is its use in characterizing singular values of nonlinear operators, and in particular, the Hankel operator and its relationship to the state space notion of nonlinear balanced realizations.

1 Introduction

Adjoint operators play an important role in linear control systems theory. They provide a duality between inputs and outputs of linear systems. The properties with respect to input, e.g. controllability and stabilizability, directly reduce to the dual results with respect to output, observability and detectability. 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 an important role in the nonlinear control systems theory. So called nonlinear Hilbert adjoint operators are introduced in [6, 11] as a special class of nonlinear adjoint operators. The existence of such operators in an input-output sense was shown in [7], but their state-space realizations are only preliminary available in [5], where the main interest is the Hilbert adjoint extension with an emphasis on the use of port-controlled Hamiltonian system methods.

Here, we consider adjoint operators from a variational point of view and provide a formal justification for the use of Hamiltonian extensions via Gâteaux differentials. We investigate whether one can use their state-space realizations to characterize singular values of nonlinear operators, and in particular, the Hankel operator. We also consider the relation to 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 motivate the line of thinking for the nonlinear case. In Section 3 state-space realizations of nonlinear adjoint operators are introduced in terms of Hamiltonian extensions. In Section 4 a formal justification of the use of Hamiltonian extensions for nonlinear adjoint systems is provided. 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. [12], 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, we summarize our conclusions.

2 Linear systems as a paradigm

This section outlines the way linear adjoint operators play an important role in the linear systems theory, see e.g. [12]. The material is presented here in a way that clarifies the line of thinking in the nonlinear case. Consider a causal linear input-output system $y = \Sigma(u) : L_2^m[0,\infty) \to L_2^r[0,\infty)$ with (A,B,C) the state-space realization. The transfer function matrix is given by $G(s) := C(sI-A)^{-1}B$. Its adjoint operator is isomorphic to $y_a = \Sigma^*(u_a) : L_2^r[0,\infty) \to L_2^m[0,\infty)$ where the transfer function matrix is given by $G^*(s) := G^T(-s) = B^T(-sI-A^T)^{-1}C^T$. Here u_a and y_a have the same dimensions as y and u, respectively. Σ^* satisfies the definition for Hilbert adjoint operators, namely,

$$\langle \Sigma(u), u_a \rangle_{L_2^r} = \langle u, \Sigma^*(u_a) \rangle_{L_2^m}. \tag{1}$$

Since u_a has the same dimension as y we have that

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

by substituting $u_a = \Sigma(u)$. This relation can be utilized to derive the singular values of the corresponding 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 $[0,\infty)$. If S is BIBO stable then the system Hankel integral operator is the well defined mapping $\mathcal{H}_\Sigma : L_2^m[0,\infty) \to L_2^r[0,\infty)$

$$\mathcal{H}_\Sigma \;:\; \hat{u} \to \hat{y}(t) = \int_0^\infty H(t+\tau)\hat{u}(\tau)d\tau.$$

Define the *time flipping* operator as the injective mapping $\mathcal{F}: L_2^m[0,\infty) \to L_2^m(-\infty,\infty)$ with $\mathcal{F}(\hat{u}(t)) = \hat{u}(-t)$ for $t<0$ and $\mathcal{F}(\hat{u}(t)) = 0$ for $t \geq 0$. Then clearly $\mathcal{H}_\Sigma = S\mathcal{F}$, where the codomain of S is restricted to $L_2^r[0,\infty)$. 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 [9]:

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

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

$$x^0 \mapsto y = \mathcal{O}_\Sigma(x^0) \;:=\; Ce^{At}x^0 \tag{2}$$
$$u \mapsto x^0 = \mathcal{C}_\Sigma(u) \;:=\; \textstyle\int_0^\infty e^{A\tau}Bu(\tau)d\tau. \tag{3}$$

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 $\mathcal{C}_\Sigma^* : \mathbb{R}^n \to L_2^r[0,\infty)$

$$u_a \mapsto x^0 = \mathcal{O}_\Sigma^*(u_a) \;:=\; \textstyle\int_0^\infty e^{A^T\tau}C^T u_a(\tau)d\tau \tag{4}$$
$$x^0 \mapsto y_a = \mathcal{C}_\Sigma^*(x^0) \;:=\; B^T e^{A^T t}x^0. \tag{5}$$

It can be easily checked that they satisfy $\langle \mathcal{O}_\Sigma(x^0), u_a\rangle_{L_2^r} = \langle x^0, \mathcal{O}_\Sigma^*(u_a)\rangle_{\mathbb{R}^n}$ and $\langle \mathcal{C}_\Sigma(u), x^0\rangle_{\mathbb{R}^n} = \langle u, \mathcal{C}_\Sigma^*(x^0)\rangle_{L_2^m}$. These adjoint operators can be used to calculate the observability and controllability Gramians, respectively:

$$\|\mathcal{O}_\Sigma(x^0)\|_{L_2^r}^2 = \langle x^0, \textstyle\int_0^\infty Ce^{A\tau}e^{A^T\tau}C^T d\tau\, x^0\rangle_{\mathbb{R}^n} = \langle x^0, Q\, x^0\rangle_{\mathbb{R}^n}$$
$$\|\mathcal{C}_\Sigma^*(x^0)\|_{L_2^m}^2 = \langle x^0, \textstyle\int_0^\infty B^T e^{A^T\tau}e^{A\tau}Bd\tau\, x^0\rangle_{\mathbb{R}^n} = \langle x^0, P\, x^0\rangle_{\mathbb{R}^n}$$

These imply that $Q = \mathcal{O}_\Sigma^* \circ \mathcal{O}_\Sigma$ and $P = \mathcal{C}_\Sigma^{**} \circ \mathcal{C}_\Sigma^* = \mathcal{C}_\Sigma \circ \mathcal{C}_\Sigma^*$. Furthermore, it is well-known that

Lemma 2.1 *[12] The operator $\mathcal{H}_\Sigma^*\mathcal{H}_\Sigma$ and the matrix QP have the same nonzero eigenvalues.*

3 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. The precise definition of *nonlinear Hilbert adjoint* operators is given as follows [6, 7, 11].

Definition 3.1 *Consider an operator $T: E \to F$ with Hilbert spaces E and F. An operator $T^*: F \times E \to E$ such that*

$$\langle T(u), y\rangle_F = \langle u, T^*(y,u)\rangle_E, \;\forall u \in E, \;\forall y \in F \tag{6}$$

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

Remark 3.2 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

$$T' \;:\; F' \to E^\beta \tag{7}$$
$$\;:\; y' \mapsto (T'(y'))(u) = (y' \circ T)(u), \;\forall u \in E, \;\forall y \in F \tag{8}$$

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

$$(T'(y'))(u) := \langle u, T^*(y,u)\rangle_E, \;\; u \in E, \; y \in F. \tag{9}$$

The converse result can be found in [7].

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

$$\langle T(u), y\rangle_F = \langle u, T^*(y)\rangle_E, \;\forall u \in E, \;\forall y \in F \tag{10}$$

in the previous section.

Now, we 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

$$u \mapsto y = \Sigma(u): \begin{cases} \dot{x} &=& f(x,u) \quad x(t^0) = 0 \\ y &=& h(x,u) \end{cases} \tag{11}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^r$. Here we assume that 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. In order to obtain a state-space characterization of the Hilbert adjoint of a system in terms of an Hamiltonian extension we have to introduce the variational system of Σ. It is given by $(u, u_v) \mapsto y_v = \Sigma_v(u, u_v)$:

$$\begin{cases} \dot{x} &=& f(x,u) & x(t^0) = 0 \\ \dot{x}_v &=& \frac{\partial f}{\partial x}x_v + \frac{\partial f}{\partial u}u_v & x_v(t^1) = 0 \\ y_v &=& \frac{\partial h}{\partial x}x_v + \frac{\partial h}{\partial u}u_v. \end{cases} \tag{12}$$

The input, state and output (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 in the form of a variational system. It is given by $(u,u_a)\mapsto y_a=\Sigma_a(u,u_a)$:

$$\begin{cases} \dot{x} &= \frac{\partial H}{\partial p}^T &= f(x,u) & x(t^0)=0 \\ \dot{p} &= -\frac{\partial H}{\partial x}^T &= -\left(\frac{\partial f}{\partial x}^T p+\frac{\partial h}{\partial x}^T u_a\right) & p(t^1)=0 \\ y_a &= \frac{\partial H}{\partial u}^T &= \frac{\partial f}{\partial u}^T p+\frac{\partial h}{\partial u}^T u_a & \\ y &= \frac{\partial H}{\partial u_a}^T &= h(x,u) & \end{cases} \tag{13}$$

with the Hamiltonian

$$H(x,p,u,u_a) := p^T f(x,u)+u_a^T h(x,u). \tag{14}$$

Remark 3.3 In Section 4, we show that such a Hamiltonian control system is a nonlinear Hilbert adjoint of the Gâteaux differential of the operator. This interpretation results from taking the Gâteaux differential of the squared L_2 norm of the nonlinear operator. Therefore, it is a more restricted interpretation than the Hilbert adjoint definition in terms of the inner product.

By careful considerations of the Hamiltonian, one can relate the Hamiltonian extension idea to the Hilbert adjoint of the original system Σ (for more details, see [5]). It boils down to extending the linear system to a $(2n+m)$-dimensional system corresponding to

$$\left(s\,\Sigma(s)\,\frac{1}{s}\right)^* = s\,\Sigma^T(-s)\,\frac{1}{s},$$

because the Hamiltonian extension was originally defined as the adjoint of the variational system. Of course this mapping coincides with $\Sigma^*(s)=\Sigma^T(-s)$ in the linear case, however, for nonlinear systems such a relation does not follow.

There also exists a relation between adjoint operators and port-controlled Hamiltonian systems, as has been established in [5]. Instead of the interpretation in terms of the Gâteaux differential of the norm (see the next section), 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 [5] for more details.

4 Gâteaux differentiation of dynamical systems

This section develops the concept of Gâteaux differentials for dynamical systems from an input-output point of view. It is not only important for understanding the meaning of the Hamiltonian extensions and adjoint systems but Gâteaux differentials of Hankel operators also play 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 differentials.

Definition 4.1 *(Gâteaux differential) Suppose X and Y are Banach spaces, $U\subseteq X$ is open, and $T:U\to Y$. Then T has a Gâteaux differential at $x\in X$ if, for all $\xi\in U$ the following limit exits:*

$$dT(x)(\xi)=\lim_{\varepsilon\to 0}\frac{T(x+\varepsilon\xi)-T(x)}{\varepsilon}=\frac{d}{d\varepsilon}T(x+\varepsilon\xi)|_{\varepsilon=0}. \tag{15}$$

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

There is also a chain rule for the Gâteaux differential, i.e., the differential of a composition is given by the following equation:

$$d(T\circ S)(x)(\xi) \;=\; dT(S(x))(dS(x)(\xi)) \tag{16}$$

Perhaps more well-known than the Gâteaux differential is the Fréchet derivative. Fréchet differentiation is a special case of Gâteaux differentiation, although in the cases where we use it, they are in fact equal. Since the directional notation of Gâteaux differentiation is more suitable for our framework, we use the Gâteaux differential.

Theorem 4.2 *Suppose that $\Sigma: u\mapsto y$ as in (11) 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*

$$u\in L_2^m(\Omega), u_v\in U_v \Rightarrow y_v\in L_2^r(\Omega), \tag{17}$$

where y_v is the output of system (12). Then it follows that

$$\Sigma_v(u,u_v)=d\Sigma(u)(u_v) \tag{18}$$

with the variational system Σ_v given in (12).

In order to prove Theorem 4.2 the following property of variational systems is needed.

Lemma 4.3 *[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*

$$x_v(t)=\frac{\partial x(t,0)}{\partial\varepsilon},\quad u_v(t)=\frac{\partial u(t,0)}{\partial\varepsilon},\quad y_v(t)=\frac{\partial y(t,0)}{\partial\varepsilon}$$

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. In [4] Theorem 4.2 and Lemma 4.3 have been extended to the more general non-input-affine case. Now, we can give the

Proof of Theorem 4.2 Let $u(t,\varepsilon)=u(t)+\varepsilon u_v(t)$ in Lemma 4.3. Then we have

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

where $R_i(u,u_v)(t) := \frac{1}{i!}\frac{\partial^i y(t,0)}{\partial \varepsilon^i}$. This implies (18). □

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

Theorem 4.4 *Suppose that the assumptions in Theorem 4.2 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*

$$\Sigma_a(u,u_v) = (d\Sigma(u))^*(u_v) \tag{19}$$

with the Hamiltonian extension Σ_a given in (13).

The fact that the Hamiltonian extension $\Sigma_a(u,u_v)$ is linearly dependent on u_v is crucial in the proof of Theorem 4.4. A more general version, related to the Hilbert adjoint definition, can be derived from the differential version of Proposition 2 in [5], but falls beyond the scope of this paper.

5 The Hankel operator and its derivative

This section gives state-space realizations for nonlinear Hilbert adjoints of various operators, and relates it to singular value analysis of energy functions and operators, i.e., the Hankel operator. We only consider time-invariant input-affine nonlinear systems without direct feed-through in the form of

$$\Sigma : \begin{cases} \dot{x} &= f(x) + g(x)u \\ y &= h(x) \end{cases} \tag{20}$$

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, 8], i.e. the mapping $u \mapsto y$ is represented by the following convergent generating series

$$u \mapsto y(t) = \sum_{\eta \in I^*} c(\eta) E_\eta(t,t^0)(u), \;\; t \geq t^0, \tag{21}$$

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

$$E_{i_k \ldots i_0}(t,t^0)(u) = \int_{t_0}^{t} u_{i_k}(\tau) E_{i_{k-1}\ldots i_0}(\tau,t_0)(u)\,d\tau \tag{22}$$

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

$$c(\eta) = L_{g_\eta} h(0) := L_{g_{i_0}} L_{g_{i_1}} \ldots L_{g_{i_k}} h(0) \tag{23}$$

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 [6, 7, 11] which are defined by

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

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

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

$$\mathcal{F}_-(u)(t) := \begin{cases} u(-t) & t \in \Omega_- \\ 0 & t \in \Omega_+ \end{cases}. \tag{24}$$

These are natural generalizations of (2) and (3).

One can employ state-space systems to describe the observability and controllability operators which are operators of $\mathbb{R}^n \to L_2^r$ and $L_2^m \to \mathbb{R}^n$, specifically:

$$x^0 \mapsto y = \mathcal{O}_\Sigma(x^0) : \begin{cases} \dot{x} &= f(x), \quad x(0) = x^0 \\ y &= h(x) \end{cases} \tag{25}$$

$$u \mapsto \tilde{x}^1 = \mathcal{C}_\Sigma(u) : \begin{cases} \dot{\tilde{x}} &= f(\tilde{x}) + g(\tilde{x})\mathcal{F}_-(u) \\ \tilde{x}^1 &= \tilde{x}(0) \end{cases} \tag{26}$$

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

$$\mathcal{H}_\Sigma := \Sigma \circ \mathcal{F}_-, \tag{27}$$

and $\mathcal{H}_\Sigma = \mathcal{O}_\Sigma \circ \mathcal{C}_\Sigma$ holds. This has been proven in [6, 7], along with a deeper and more detailed analysis of the Hankel operator. We can state the differential version of this fact by using (16) as

$$d\mathcal{H}_\Sigma(u)(u_v) = d\mathcal{O}_\Sigma(\mathcal{C}_\Sigma(u))(d\mathcal{C}_\Sigma(u)(u_v)). \tag{28}$$

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 5.1 *Consider the system Σ, and suppose the assumptions of Theorem 4.2 hold. Then*

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

This theorem directly follows from the definition of $\mathcal{O}_\Sigma$, $\mathcal{C}_\Sigma$, $\mathcal{H}_\Sigma$, Σ and the Gâteaux differential $d(\cdot)$. Furthermore their adjoints can be obtained by using Theorem 4.4.

Theorem 5.2 *Consider the operator Σ as in (20). Suppose that the assumptions of Theorem 4.2 and Theorem 4.4 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* $(x^0,u_a) \mapsto p^0 = (d\mathcal{O}_\Sigma(x^0))^*(u_a)$:

$$\begin{cases} \dot{x} &= f(x), & x(0) = x^0 \\ \dot{p} &= -\frac{\partial f}{\partial x}^T(x)\,p - \frac{\partial h}{\partial x}^T(x)\,u_a, & p(\infty) = 0 \\ p^0 &= p(0) \end{cases} \tag{29}$$

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

$$\begin{cases} \dot{x} &= f(x) + g(x)\mathcal{F}_-(u), & x(-\infty) = 0 \\ \dot{p} &= -\frac{\partial(f(x)+g(x)\mathcal{F}_-(u))}{\partial x}^T p, & p(0) = p^1 \\ y_a &= \mathcal{F}_+(g^T(x)\,p) \end{cases} \tag{30}$$

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

$$\left\{ \begin{array}{rcll} \dot{x} &=& f(x) + g(x)\,\mathcal{F}_-(u), & x(-\infty) = 0 \\ \dot{p} &=& -\frac{\partial f}{\partial x}^T(x)\,p - \frac{\partial h}{\partial x}^T(x)\,u_a, & p(\infty) = 0 \\ y_a &=& \mathcal{F}_+(g^T(x)\,p), & \end{array} \right. \quad (31)$$

respectively. Here $\mathcal{F}_+ : L_2^m(\Omega_-) \to L_2^m(\Omega_+)$ denotes another flipping operator defined by

$$\mathcal{F}_+(u)(t) := \left\{ \begin{array}{ll} 0 & t \in \Omega_- \\ u(-t) & t \in \Omega_+ \end{array} \right. . \quad (32)$$

The proof of this theorem is obtained by applying the adjoint Hamiltonian extensions of Section 3, and using techniques from [5].

6 Energy functions and singular values

In order to proceed, first consider the following energy functions:

Definition 6.1 *The observability function $L_o(x)$ and the controllability function $L_c(x)$ of Σ as in (20) are defined by*

$$L_o(x^0) := \frac{1}{2}\int_0^\infty \|y(t)\|^2 dt,\ x(0) = x^0,\ u(t) \equiv 0 \quad (33)$$

$$L_c(x^1) := \min_{\substack{u \in L_2^m(\Omega_-) \\ x(-\infty) = 0 \\ x(0) = x^1}} \frac{1}{2}\int_{-\infty}^0 \|u(t)\|^2 dt \quad (34)$$

respectively.

These functions are closely related to the observability and controllability operators and Gramians in the linear case, see Section 2. In [10] these functions have been used for the definition of balanced realizations and singular value functions of nonlinear systems. Also they fulfill corresponding 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 we mean by input-normal/output-diagonal form, see [10]:

Theorem 6.2 *[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*

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

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

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

Now, we present the relation between the observability function, operator and Gramian.

$$\begin{array}{rcl} L_o(x^0) &=& \frac{1}{2}\|\mathcal{O}_\Sigma(x^0)\|_{L_2^r}^2 = \frac{1}{2}\langle x^0, \mathcal{O}_\Sigma^*(\mathcal{O}_\Sigma(x^0), x^0)\rangle_{\mathbb{R}^n} \\ &=:& \frac{1}{2}\langle x^0, \phi(x^0)\rangle_{\mathbb{R}^n} \end{array} \quad (35)$$

The function $\phi(x^0)$ can always be rewritten by $\phi(x^0) = Q(x^0)\,x^0$ 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 as in the observability case. From equation (35) it does follow that

$$\begin{array}{rcl} L_c(x^1) &=& \frac{1}{2}\|\mathcal{C}_\Sigma^\dagger(x^1)\|_{L_2^m}^2 = \frac{1}{2}\langle x^1, \mathcal{C}_\Sigma^{\dagger *}(\mathcal{C}_\Sigma^\dagger(x^1), x^1)\rangle_{\mathbb{R}^n} \\ &=:& \frac{1}{2}\langle x^1, \varphi(x^1)\rangle_{\mathbb{R}^n} \end{array} \quad (36)$$

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

$$\mathcal{C}_\Sigma^\dagger(x^1) := \arg \min_{\mathcal{C}_\Sigma(u) = x^1} \|u\|_{L_2^m}. \quad (37)$$

Now, we can state the result from [6, 7] that relates the singular value functions to the Hankel operator:

Theorem 6.3 *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 $\bar{z}_j = (0, \ldots, 0, z_j, 0, \ldots, 0)$ for $j = 1, 2, \ldots, n$, and the functions $\hat{\sigma}^2(z_j) = \tau(\bar{z}_j)$. Let v_j be the minimum energy input which drives the state from $z(-\infty) = 0$ to $z(0) = \bar{z}_j$ 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:*

$$\langle \hat{v}_j, (\mathcal{H}_\Sigma^* \mathcal{H}_\Sigma)(\hat{v}_j)\rangle_{L_2} = \hat{\sigma}_j^2(z_j)\langle \hat{v}_j, \hat{v}_j\rangle_{L_2},\quad j = 1, 2, \ldots n. \quad (38)$$

The above result is quite limited in the sense that it is dependent on the coordinate frame in which the system is in input-normal/output-diagonal form. We now give a more general relationship between the singular value functions and the Hankel operator. The idea is to give an extension of the linear result of Lemma 2.1 inspired by the proof of the latter lemma as given in [12]. To this effect, we consider the Gâteaux differential of the Hankel operator output in the following way

$$\begin{array}{rcll} d\|\mathcal{H}_\Sigma(u)\|_2^2(v) &=& 2\,\langle d\mathcal{H}_\Sigma(u, v), \mathcal{H}_\Sigma(u)\rangle & (39) \\ &=& 2\,\langle v, (d\mathcal{H}_\Sigma(u))^* \circ \mathcal{H}_\Sigma(u)\rangle & (40) \end{array}$$

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

$$(d\mathcal{H}_\Sigma(u))^* \circ \mathcal{H}_\Sigma(u) = \lambda(u)u, \quad (41)$$

where $\lambda(u)$ is an eigenvalue depending on eigenvector u. However, since we want to relate it to the notion of singular value functions, and thus would like to have the eigenvalue be dependent on x^0, we need an additional step. We propose to consider eigenvalues $\tilde{\sigma}(x^0)$ and corresponding eigenvectors x^0 of the following:

$$\begin{array}{l} \mathcal{C}_\Sigma \circ d\mathcal{H}_\Sigma^* \circ \mathcal{H}_\Sigma(u) = \mathcal{C}_\Sigma \circ d\mathcal{H}_\Sigma^* \circ \mathcal{O}_\Sigma(x^0) = \tilde{\sigma}(x^0)x^0 \\ \mathcal{C}_\Sigma(u) = x^0 \end{array} \quad (42)$$

This leads to the following result:

Theorem 6.4 *Assume all technical conditions for Theorem 6.2 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*

$$\begin{aligned} &\mathcal{C}_\Sigma \circ d\mathcal{H}_\Sigma^* \circ \mathcal{H}_\Sigma(u) = \mathcal{C}_\Sigma \circ d\mathcal{C}_\Sigma^* \circ d\mathcal{O}_\Sigma^* \circ \mathcal{O}_\Sigma(x^0) = \\ &\mathcal{C}_\Sigma(\lambda(u)u) = M_c(\psi(x^0))^{-1}\tfrac{\partial L_o}{\partial x}(x^0) \end{aligned} \quad (43)$$

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 (29) is given by $p = \frac{\partial^T L_o}{\partial x}(x)$, where x is the solution of system (25), 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 (30), where $\tilde{x}$ is the solution of system (26) and where $u = y_a = \mathcal{F}_+(g^T(\tilde{x})p)$. Thus,

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

Remark 6.5 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 2.1 of [12]

By taking x^0 to be an eigenvector of the above operator, we obtain the relation (42). Observe that the $\tilde{\sigma}(x^0)$'s do not equal the singular value functions as defined in Theorem 6.2. However, we are able to relate the eigenvalues of the above theorem to the singular value functions in the following way.

Corollary 6.6 *Suppose that the system is in the form of Theorem 6.2, and write*

$$M_c(\psi(z))^{-1}\frac{\partial L_o}{\partial x}(z) = \frac{\partial L_o}{\partial x}(z) = T(z)z$$

where $T(z)$ *follows from the form of* L_o *in Theorem 6.2. Then for the collection of component vectors* $\bar{z}_j$, $j = 1,\ldots,n$*, as defined in Theorem 6.3 the eigenvalues* $\rho_i(\bar{z}_j)$ *of* $T(\bar{z}_j)$ *are given by*

$$\begin{aligned} &\rho_i(\bar{z}_j) = \tau_i(\bar{z}_j), i = 1,\ldots,j-1,j+1,\ldots,n \\ &\rho_j(\bar{z}_j) = \tau_j(\bar{z}_j) + \tfrac{1}{2}\tfrac{\partial\tau_j}{\partial z_j}(\bar{z}_j)z_j \end{aligned}$$

for $j = 1,\ldots,n$.

Proof: Note that

$$T(z) = \begin{pmatrix} \tau_1(z) & & 0 \\ & \ddots & \\ 0 & & \tau_n(z) \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \frac{\partial\tau_1}{\partial z_1}(z)z_1 & \cdots & \frac{\partial\tau_n}{\partial z_1}(z)z_n \\ \vdots & \ddots & \vdots \\ \frac{\partial\tau_1}{\partial z_n}(z)z_1 & \cdots & \frac{\partial\tau_n}{\partial z_n}(z)z_n \end{pmatrix}$$

and then plug in $z_1 = \cdots = z_{j-1} = z_{j+1} = \cdots = z_n = 0$. The result follows straightforwardly. $\square$

7 Conclusions

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

Acknowledgments

This research was supported in part by a travel grant for the first author from the Netherlands Organization for Scientific Research (NWO), file number R62-554, and by a research fellowship from Delft University of Technology for the second author.

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., M. Lamnabhi and F. Lamnabhi-Lagarrigue (1983). An algebraic approach to nonlinear functional expansions. *IEEE Trans. Circ. Syst.*, vol. 30, pp. 554-570.

[4] Fujimoto, K., J.M.A. Scherpen (2000). Nonlinear balancing based on differential eigenstructure of Hankel operators. *submitted.*

[5] Fujimoto, K., J.M.A. Scherpen, and W.S. Gray (2000). Hamiltonian realizations of nonlinear adjoint operators. *Proc. IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control*, Princeton, March 2000, pp. 39–44.

[6] 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.

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

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

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

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