On input/output maps for nonlinear systems via continuity in a locally convex topology

Ravi R. Mazumdar\(^a\)*, Raghavan Kannurpatti\(^b\), Arunabha Bagchi\(^c\)

\(^a\)INRS-Telecommunications, Université du Québec, Ile des Soeurs, Què. H3E 1H6, Canada
\(^b\)Department of Electrical Engineering, Indian Institute of Technology, Hauz Khas, New Delhi 110016, India
\(^c\)Department of Applied Mathematics, University of Twente, PO Box 217, 7500AE Enschede, Netherlands

(Received 24 November 1993; revised 12 February 1994)

Abstract

In this paper we show that the output of a nonlinear system with inputs in \((L_2[0, T]; \mathbb{R}^m)\) whose state satisfies a nonlinear differential equation with standard smoothness conditions can be written as the composition of a nonlinear map with a linear Hilbert–Schmidt operator acting on the input. The result also extends to abstract semi-linear infinite dimensional systems. The approach is via the study of the continuity of the solution in a locally convex topology generated by seminorms of Hilbert–Schmidt operators in Hilbert space. The result reveals an entirely new structure related to nonlinear systems which can lead to useful approximation results.

Keywords: Nonlinear systems; Smoothness; Approximation; Infinite-dimensional systems

1. Introduction

In this paper we point out the role of the locally convex topology induced on a Hilbert space of \(L_2\) functions by Hilbert–Schmidt operators, showing thereby that the input/output map for a general class of nonlinear differential systems is given by the composition of a nonlinear map with a linear Hilbert–Schmidt map on the input which is taken to be in \(L_2[0, T]\). Not only does this result generalize the well known results of Freedman and Willems [6] on the representation of I/O maps of nonlinear systems involving derivatives of inputs as smooth mappings on the undifferentiated inputs (see also [12]) but also it has important implications in the I/O representations in terms of Volterra series (more correctly polynomials).

Consider the following nonlinear system (for simplicity we consider the scalar case):

\[
\frac{d}{dt} X_t = f(X_t) + g(X_t)u_t, \quad t \in [0, T].
\] (1.1)

Suppose the observation is a smooth measurable mapping of the state, i.e. \(Y_t = h(X_t)\) where \(h(\cdot)\) is a smooth function.
Then the result of Freedman and Willems [6] states that if \( f(\cdot), g(\cdot) \) and \( h(\cdot) \) are \( C^\infty \) with bounded first derivatives then the solution can be written as:

\[
Y_t = h \circ r \left( \int_0^t u_s \, ds \right),
\]  

(1.2)

where \( r(\cdot) \) is a continuous map not depending on \( \{u_s\} \).

Their approach was based on differential geometric arguments. The extension to the vector case was obtained by an additional commutativity assumption on the control term \( g(\cdot) \).

In this paper, by studying the continuity of the solution to (1.1) (but without the \( C^\infty \) requirement) in a locally convex topology generated by semi-norms of Hilbert–Schmidt operators, we obtain a generalization of the results using a purely functional analytic framework. The locally convex topology is termed the S-topology and plays an important role in the study of probability measures on infinite dimensional spaces. See [1, 2, 7, 9, 10] and the references therein.

The basic aim of this paper is to show that the S-topology has important implications in the study of deterministic nonlinear systems as well. In particular by studying a special class of S-continuous maps, i.e., those which are uniformly S-continuous around the origin (uscaco), we can prove a generalization of the results of [6]. The representation result also leads to a very simple proof of the representation of I/O maps as finite Volterra series via the Weierstrass theorem on Hilbert spaces [11] without unduly restrictive conditions. A good discussion on Volterra series can be found in the book of Isidori [8] and the paper of Brockett [3].

The organization of the paper is as follows: in Section 2 we recall the basic facts of S-topology. In Section 3 we show that the representation result holds for nonlinear systems under usual hypotheses for which solutions over finite intervals can be shown to exist. In Section 4 we show how the representation result leads to very simple proofs of Volterra polynomial approximations.

2. S-topology and the representation theorem

In this section we recall the basic facts on the S-topology which will be the framework of study in this paper. We then state and prove a representation result for a particular class of S-continuous maps which occur in the study of nonlinear systems. For details we refer the reader to [7, 9, 10].

We now introduce some notation which will be used throughout.

Let \( H = L_2([0, T]; H_1) \) denote the separable Hilbert space of square integrable functions in \([0, T]\) which take values in another separable Hilbert space \( H_1 \).

Let \([\cdot, \cdot]_r\) denote the usual inner-product in \(H_1\) i.e.

\[
[f, g]_r = \int_0^T \langle f_s, g_s \rangle \, ds, \quad f, g \in H,
\]

where \( \langle \cdot, \cdot \rangle \) denotes the inner-product in \( H_1 \) which is the usual euclidean scalar product when \( H_1 = \mathbb{R}^n \) and \( |\cdot| \) denotes the corresponding norm.

We denote the inner-product in \( L_2([0, t]; t \in [0, T]) \) by \([\cdot, \cdot]_t\), which can be represented in terms of the inner-product on \( H \) by restricting the functions to the interval \([0, t]\) by means of the projection operator \( Q_t \) defined by:

\[
Q_t f = \begin{cases} 
 f(s), & 0 \leq s \leq t, \\
 0, & s \geq t,
\end{cases}
\]

and hence

\[
[Q_t f, Q_t g]_t = [f, g]_t.
\]
Definition 2.1. A map $F : H \to B$, $B$ a Banach space, is said to be continuous at $x \in H$ in the $S$-topology if for any $\varepsilon > 0$ there exists a Hilbert–Schmidt (HS) operator $L_\varepsilon(x) : H \to H$ such that

$$
\|L_\varepsilon(x)(x - x')\| < 1 \Rightarrow \|F(x) - F(x')\| < \varepsilon.
$$

We say that the mapping is uniformly $S$-continuous if $L_\varepsilon(x)$ does not depend on $x$.

Definition 2.2. A map $F : H \to B$ is said to be uniformly $S$-continuous around the origin (uscao) if $F$ is uniformly $S$-continuous on sets:

$$
U_n = \{x \in H : \|A_n x\| \leq 1\}
$$

where $\{A_n\}_{n=1}^\infty$ is a sequence of Hilbert–Schmidt operators such that $\|A_n\|_{\text{HS}} \to 0$ and $\bigcup_{n=1}^\infty U_n = H$.

Remark. A mapping which is uniformly $S$-continuous is obviously $S$-continuous around the origin. In fact taking $A_n = (1/n)L$ with $L$ defined above satisfies Definition 2.2. A necessary and sufficient condition for a map to be uscao was given by Gandolfi and Germani [7].

Theorem 2.3. A map $f(\cdot) : H \to B$, where $H$ is a real, separable Hilbert space and $B$ is a Banach space, is uscao if and only if it is continuous with respect to the topology $T_{\text{f}}$ generated by a sequence of semi-norms $\{\|L_i x\|\}$ where $\{L_i\}$ is a sequence of Hilbert–Schmidt operators.

Remark. If $f : H \to B$ is uscao then $g \circ f : H \to B$ is uscao if $g : B \to B$ is continuous.

Recently, a representation result for a map to be uscao was given by Desantis et al. [5] which we state below.

Theorem 2.4. A necessary and sufficient condition for a map $F : H \to B$ to be uscao is that there exists a Hilbert–Schmidt operator $L : H \to H$ and a continuous map $g : H \to B$ such that

$$
F = g \circ L. \tag{2.1}
$$

Moreover $g(\cdot)$ is continuous on every bounded sphere in $H$.

In the next section via the use of Theorem 2.3 we show that the I/O maps of nonlinear systems whose state is governed by a nonlinear differential equation of type (1.1) with standard continuity and growth conditions on the coefficients $f(\cdot)$ and $g(\cdot)$ have the uscao property w.r.t. the inputs in $L_2[0, T]$.

3. Representation of I/O maps

Consider the nonlinear system given below:

$$
\begin{align*}
\frac{d}{dt} X_t &= f(X_t) + g(X_t)U_t, \quad 0 \leq t \leq T, \\
Y_t &= h(X_t), \quad 0 \leq t \leq T,
\end{align*}
$$

where $X_t \in \mathbb{R}^n, f(\cdot) : \mathbb{R}^n \to \mathbb{R}^n, g(\cdot) : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m, h(\cdot) : \mathbb{R}^n \to \mathbb{R}^p$ is a continuous map and $U \in L_2([0, T], \mathbb{R}^m)$. 
Furthermore it is assumed that the coefficients $f(\cdot), g(\cdot)$ satisfy the following assumptions:

(A1) $f(\cdot)$ is uniformly Lipschitz with Lipschitz constant $K < \infty$ i.e.
\[ |f(x) - f(y)| \leq K|x - y|, \quad x, y \in \mathbb{R}^n. \]

(A2) $g_{i,j}(\cdot) \in C^1$ and $|\partial g_{i,j}(x)/\partial x_i| < K \forall x \in \mathbb{R}^n$ where $g_{i,j}$ denotes the $(i,j)$ element of $g(\cdot)$ and $1 \leq i, l \leq n, 1 \leq j \leq m$.

(A3) $f(\cdot)$ satisfies the following growth condition:
\[ |f(x)| \leq C_1(1 + |x|). \]

(A4) $\sup_x |g(x)|_0 < \infty$ where $|g(x)|_0$ denotes the matrix norm defined by $|A|_0 = \sup_{|h|_1 \leq 1} |Ah|$ for any matrix $A$.

Assumptions (A2) and (A4) imply that $g(\cdot)$ satisfies uniform Lipschitz and growth conditions like $f(\cdot)$ w.r.t. the matrix norm. We state this below without proof since it is just a simple consequence of the mean value theorem and the definition of the matrix norm.

Lemma 3.1. Under the assumptions that $g(\cdot)$ satisfies (A2) and (A4) the following is true. There exists a constant $C_2 < \infty$ such that:
\[ |g(x) - g(y)|_0 \leq C_2 |x - y|, \quad \forall x, y \in \mathbb{R}^n \quad \text{(Lipschitz condition)} \]
\[ |g(x)|_0 \leq C_2(1 + |x|), \quad \forall x \in \mathbb{R}^n \quad \text{(Growth condition)} \]

Remark. With the above lemma the nonlinear differential equation describing the state satisfies the “usual” conditions for the existence and uniqueness of solutions for measurable inputs. The existence and uniqueness can be established by the usual Picard iteration technique. For details see any standard text on differential equations, e.g. Coddington and Levinson [4].

We now state the main representation result.

Theorem 3.2. Consider the nonlinear system given by (3.1) and (3.2) and assume that the coefficients satisfy (A1)-(A4).

Then for any input $Q, U \in (L_2[0, T]; \mathbb{R}^m)$, there exists a Hilbert–Schmidt operator $K: (L_2[0, T]; \mathbb{R}^m) \rightarrow (L_2[0, T]; \mathbb{R}^m)$ and a continuous mapping $\Psi(\cdot): (L_2[0, T]; \mathbb{R}^m) \rightarrow \mathbb{R}^n$ such that the output $\{Y_t\}$ satisfies
\[ Y_t = h(\Psi \circ KU(t)), \]
where
\[ v = KU, \quad v_t = \int_0^t K(t, s)U_s ds, \]
and $\Psi(\cdot)$ does not depend on $U$.

In order to prove the theorem we need the following estimate on the solution of the differential equation.

Lemma 3.3. Let $\{X_t\}_{t \in [0, T]}$ denote the solution of (3.1) under the assumptions (A1)–(A3). Then there exist finite constants $K_1, K_2$ and $K_3$ such that $X_t$ for $t \in [0, T]$ satisfies:
\[ \sup_{0 \leq t \leq T} |X_t| \leq (K_1 + K_2 \|U\|) e^{K_3\|U\|}. \]

(3.4)
Proof. Integrating (3.1) gives:

\[ X_t = X_0 + \int_0^t f(X_s) \, ds + \int_0^t g(X_s)U_s \, ds. \]

Therefore:

\[ |X_t| \leq |X_0| + \int_0^t |f(X_s)| \, ds + \int_0^t |g(X_s)U_s| \, ds. \]

Using the growth conditions on \( f(\cdot) \) and \( g(\cdot) \) we obtain

\[ |X_t| \leq |X_0| + C_1 \int_0^t (1 + |X_s|) \, ds + \int_0^t |g(X_s)|_0 U_s \, ds \]

\[ \leq |X_0| + C_2 \sqrt{t} \|U\|_t + C_1 t + \int_0^t (C_1 + C_2 |U_s|) |X_s| \, ds. \]

Using the Gronewall–Bellman inequality applied to \( |X_t| \) we obtain

\[ |X_t| \leq (|X_0| + C_2 \sqrt{T} \|U\| + C_1 T) e^{C_1 T + C_2 \sqrt{T} \|U\|}, \]

and since the bound on the r.h.s. is true for all \( t \in [0, T] \) the result follows by identifying \( K_1 = (|X_0| + TC_1) e^{C_1 T}, K_2 = \sqrt{TC_2} e^{C_1 T} \) and \( K_3 = C_2 \sqrt{T}. \)

Remark. In the following we denote the bound on the r.h.s. above by \( K_1(\|U\|) \) for convenience. Note \( K_1(\|U\|) \) is increasing in \( \|U\| \). With the aid of the estimates we now prove the theorem.

Proof of Theorem 3.2. We prove the theorem by showing that the solution of (3.1) as a mapping on the input in \( (L_2(0, T), \mathbb{R}^m) \to (L_2(0, T), \mathbb{R}^n) \) is uscao by establishing continuity with respect to a topology of a sequence of Hilbert–Schmidt operators \( \{nL\} \) for an appropriate Hilbert–Schmidt operator \( L \). Then the required conclusions follow from Theorems 2.3 and 2.4.

Let \( \{X_t(U)\} \) and \( \{X_t(U')\} \) be two solutions to (3.1) corresponding to two inputs \( U \) and \( U' \), respectively. Then,

\[ \|X(U) - X(U')\|^2 = \int_0^T \left| \int_0^T (f(X_s(U)) - f(X_s(U')) \, ds + \int_0^T (g(X_s(U))U_s - g(X_s(U'))U'_s) \, ds \right|^2 dt. \]  

Using the Lipschitz conditions on \( f(\cdot) \) and \( g(\cdot) \) we obtain

\[ \|X(U) - X(U')\|^2 \leq 3K^2 \int_0^T \left| X_s(U) - X_s(U') \right|^2 \, ds \, dt \]

\[ + 3C_2^2 \int_0^T \left| X_s(U) - X_s(U') \right|^2 \, ds \, dt \]

\[ + 3 \int_0^T \left| \int_0^T g(X_s(U'))(U_s - U'_s) \, ds \right|^2 \, dt \]

\[ \leq 3K^2 T \int_0^T \|X(U) - X(U')\|^2 \, dt + 3C_2^2 \|U\|^2 \int_0^T \|X(U) - X(U')\|^2 \, dt \]

\[ + 3 \int_0^T \left| \int_0^T g(X_s(U'))(U_s - U'_s) \, ds \right|^2 \, dt. \]
Define the Hilbert–Schmidt (and Volterra) operator \( L \) by
\[
Lf = g, \quad g_i = \int_0^t f_s \, ds, \quad f \in (L_2(0, T], \mathbb{R}^m).
\]
and by abuse of notation we define \( L_t f = \int_0^t f_s \, ds \).

Let us now consider the integrand of the third integral above. Then via integration by parts since \( g(\cdot) \) is assumed to be in \( C^1 \) we obtain:
\[
\int_0^t g(\mathbf{X}_s(U'))(\mathbf{U}_s - \mathbf{U}_t') \, ds = g(\mathbf{X}_t(U'))L_t(U' - U') - \int_0^t \frac{dg(\mathbf{X}_s(U'))}{ds} L_s(U - U') \, ds.
\]

Then by definition
\[
\frac{dg(\mathbf{X}_s(U'))}{ds} = \sum_{i=1}^n \frac{\partial g(\mathbf{X}_s(U'))}{\partial X_s(l, U')} f_l(X_s(U')) + \sum_{j=1}^m g_i(X_s(U'))U_s'(j),
\]
where the index \( l (j) \) denotes the \( l \)th \( (j) \)th component of the corresponding vector. Therefore, using the Cauchy–Schwarz inequality we obtain:
\[
\int_0^t \left| \frac{dg(\mathbf{X}_s(U'))}{ds} \right|^2 \leq t4nK^2C(t)(K_1(\| U' \|)^2 + 1) + 4nmK^2C(t)(1 + K_1(\| U' \|)^2)\| U' \|^2.
\]

Using the growth conditions on \( f(\cdot) \), \( g(\cdot) \) and the result of Lemma 3.3 we obtain
\[
\left\| \frac{dg(\mathbf{X}_t(U'))}{ds} \right\|_t^2 \leq 4nK^2C(t)(K_1(\| U' \|)^2 + 1) + 4nmK^2C(t)(1 + K_1(\| U' \|)^2)\| U' \|^2.
\]

Therefore combining all the estimates we obtain:
\[
\| \mathbf{X}(U) - \mathbf{X}(U') \|^2_t \leq 3K^2T \int_0^T \| \mathbf{X}(U) - \mathbf{X}(U') \|^2_t \, dt
\]
\[
+ 3C^2\| U \|^2 \int_0^T \| \mathbf{X}(U) - \mathbf{X}(U') \|^2_t \, dt + 3(K_3(\| U' \|)\| L(U - U') \|^2_t),
\]
where \( K_3(\| U' \|) = nK^2C(t)(1 + K_1(\| U' \|)) + 4nmK^2C(t)(1 + K_1(\| U' \|))\| U' \|^2. \)

Now applying the Gronwall–Bellman inequality to \( \| \mathbf{X}(U) - \mathbf{X}(U') \|^2_t \) we obtain:
\[
\| \mathbf{X}(U) - \mathbf{X}(U') \|^2 \leq K(\| U \|, \| U' \|)\| L(U - U') \|^2_t,
\]
where \( K(\| U \|, \| U' \|) \) is an increasing constant (depending on \( \| U \|, \| U' \| \)) given by
\[
K(\| U \|, \| U' \|) = K_3(\| U' \|) e^{(3K^2T + 4nmK^2C(t))T}(1 + K_1(\| U' \|)^2).
\]

Hence for all integers \( n \geq K(\| U \|, \| U' \|) \) we have that \( \| \mathbf{X}(U) - \mathbf{X}(U') \| \) is continuous w.r.t. the sequence of semi-norms \( \{ nL \} \) where \( L \) is the Hilbert–Schmidt integral operator defined by
\[
Lf = g, \quad g_i = \int_0^t f_s \, ds.
\]

Therefore by Theorem 2.3 the solution \( \{ \mathbf{X}_t(U) \} \) is uscao and by Theorem 2.4 the representation follows. \( \square \)

Remarks. Theorem 3.4 can be seen as a generalization of the result of Freedman and Willems [6] since the representation as the integral of the input is a particular case of the general result which states that the representation can be written as a Hilbert–Schmidt mapping on the input. Our conditions are however much weaker. On the other hand the result of Theorem 3.4 does not identify the Hilbert–Schmidt explicitly as is the
The above result can also be shown to hold for a class of infinite-dimensional systems which we show below. These are the class of semi-linear abstract differential systems of the type given below.

Consider the abstract semi-linear differential equation:

$$\frac{dX(t,h)}{dt} = AX(t,h) + f(X(t,h)) + Bh(t), \quad 0 \leq t \leq T \tag{3.6}$$

where $X(\cdot, h) \in (L^2[0, T]; H_1)$, $A$ is the infinitesimal generator of a strongly continuous $C_0$ semigroup, $f(\cdot): H_1 \to H_1$ is homogeneous and satisfies a uniform Lipschitz condition with Lipschitz constant $K$, $B: H_2 \to H_1$ is a bounded, Hilbert–Schmidt operator and the input $h(\cdot) \in (L^2[0, T]; H_2)$ where $H_1$ and $H_2$ are separable Hilbert spaces.

Then we can show the following result for the process $X(\cdot, h)$ which is the solution of (3.6). We omit the proof since it is similar to the one above except that under the above assumptions the solution can be shown to be uniformly $S$-continuous (and hence uscao).

**Theorem 3.4.** Consider the abstract semi-linear system given by (3.6). Then for each $h \in (L^2[0, T]; H_2)$, (3.6) has a unique solution $X(\cdot, h)$ such that $X(t,h)$ is weakly continuous in $0 \leq t \leq T$ and moreover the solution can be written as:

$$X(\cdot, h) = g \circ (Lh),$$

or equivalently

$$X(t, h) = g(Lh)(t),$$

where $L$ is a linear Hilbert–Schmidt operator mapping $(L^2[0, T]; H_2) \to (L^2[0, T]; H_2)$ and $g(\cdot)$ is a continuous map mapping $(L^2[0, T]; H_2) \to (L^2[0, T]; H_1)$.

4. Volterra approximations

The uscao property of functions on $H$ leads to very simple proofs of approximation of I/O maps of nonlinear systems in terms of Volterra polynomials [2].

Recall a mapping $P^N(\cdot): H \to H$ is said to be a finite Volterra polynomial if $P^N(\cdot)$ has the form:

$$P^N(h) = \sum_{n=0}^{N} p_n(Lh),$$

where $p_n(h)$ is a symmetric $n$-linear form on $H$ i.e. $p_n(h)$ is a finite homogeneous polynomial (of order $n$) and $L$ is a Hilbert–Schmidt, Volterra operator.

Then the uscao property of I/O maps of nonlinear systems given by (2.1) or (1.4) above leads to the following result.

**Theorem 4.1.** Let $Y_t(h) = h(X_t(U))$ be the output of a nonlinear system which is uscao. Then $Y_t(U)$ can be approximated by a finite Volterra polynomial for every input $U$ which lies in a bounded sphere in $H$.

**Proof.** The proof essentially follows from the Weierstrass theorem for separable Hilbert spaces (see Prenter [11]).
Indeed, by the Weierstrass theorem for all continuous maps \( g(\cdot) \) on \( H \), for every compact set \( C \) in \( H \) and \( \varepsilon > 0 \) there exists a polynomial of finite rank \( N_\varepsilon \) (which depends on \( C \)) such that:

\[
\sup_{h \in C} \| P^{N_\varepsilon}(h) - g(h) \| \leq \varepsilon.
\]

Now since \( Y_\varepsilon(U) \) is usc as by Theorem 3.4 \( Y_\varepsilon(U) = \phi(LU)(t) = h \circ \Psi(LU)(t) \) where \( L \) is Hilbert–Schmidt. By the decomposition of every Hilbert–Schmidt operator \( L = KA \) where \( K \) is compact and \( A \) is Hilbert–Schmidt it follows that the bounded sets \( M = \{ h \in H : \|h\| \leq m \} \) are mapped into pre-compact sets i.e. \( L(M) \) is compact. Hence by the Weierstrass theorem

\[
\sup_{LU \in L(M)} \| P^{N_\varepsilon}(LU) - \phi(LU) \| \leq \varepsilon,
\]

which is equivalent to

\[
\sup_{U \in M} \| P^{N_\varepsilon}(LU) - \phi(LU) \| \leq \varepsilon,
\]

and \( P^{N_\varepsilon}(LU) \) defines a finite rank Volterra polynomial.

**Remark.** By exploiting the fact that the representation map is continuous on every bounded sphere in \( H \) a very simple proof of BIBO stability can be given (via the Heine–Borel theorem).

5. Conclusions

In this paper we have pointed out the use of the locally convex topology associated with seminorms of Hilbert–Schmidt operators in revealing an entirely new structure for I/O maps of nonlinear systems from which the existence of Volterra expansions for such maps is very easily established. The representation result however does not give an explicit characterization of the Hilbert–Schmidt operator acting on the input. Since the inputs are assumed to be in \( L_2[0, T] \) the result does not indicate any smoothness associated with the kernel of the Hilbert–Schmidt operator as in the case of the result of Freedman and Willems [6]. However, it is well known that one can approximate kernels on \( L_2([0, T] \times [0, T]) \) by differentiable kernels and thus one can approximate the I/O maps by smooth filters acting on the inputs (in particular they can be taken to be differentiable).

Acknowledgements

This work was supported in part by a grant from the Natural Sciences and Engineering Council of Canada No. OGP 0042024. The authors wish to thank Prof. A. Germani for useful discussions and making his preprints available to us.

References


