J. Nonlinear Funct. Anal. 2016 (2016), Article ID 40

MATHRES
Journal of Nonlinear Functional Analysis
Available online at http://jnfa.mathres.org

VOLTERRA SERIES AND NONLINEAR INTEGRAL EQUATIONS OF THE SECOND KIND ARISING FROM OUTPUT FEEDBACK AND THEIR OPTIMAL CONTROL

N. U. AHMED

Department of EECS, University of Ottawa, Ottawa, Canada

Abstract. In this paper we consider nonlinear systems using infinite Volterra series describing the input-output model. We present sufficient conditions under which the nonlinear operator determined by the infinite Volterra series maps L_2 space into itself. Using output feedback we arrive at a nonlinear integral equation of the second kind. We prove existence of solutions of this equation and local input-output stability. Then we consider a control problem where the input is to be chosen to optimize certain objective functional. We prove existence of suboptimal control (or input) policy from the class of regular controls. We discuss the question of nonexistence of optimal policy from the class of regular controls, and then prove the existence of an optimal relaxed control from the class of finitely additive measures defined on a related compact Hausdorff space (the space of regular controls endowed with the weak topology). Next we consider the inverse problem identifying the kernels of the Volterra series. These results are then extended to an infinite dimensional Hilbert space.

Keywords. Nonlinear integral equation; Output feedback; Suboptimal control; Relaxed control; Inverse problem.

1. Introduction

It is well known in engineering literature that many linear systems (more precisely affine systems) admit the following input-output representation

$$y(t) = \Pi(t) + \int_0^t K(t,s)x(s)ds, t \geq 0, \tag{1}$$

E-mail address: ahmed@site.uottawa.ca

Received June 20, 2016

where y is the output and x is the input and K is the so-called impulse response. Indeed, if $x(s) = a\delta(s-t_0), s \geq t_0$, the output y at time t is given by $y(t) = aK(t,t_0)$ where a denotes the size or intensity of the impulsive force applied at time t_0. The function Π may be given by

$$\Pi(t) = \int_{-\infty}^{0} K(t,s)x_0(s)ds, t \geq 0,$$

where x_0 denotes the past history of input. The function Π represents the impact of the past history on the present state. In the case of time invariant systems, the kernel K is a function of one variable, the difference $(t-s) \geq 0$, with $K(t-s)$ giving the response of the system which was subjected to an impulsive force of unit intensity $(t-s)$ units of time before.

In case of distributed systems the input-output representation is given by

$$y(t,\xi) = \Pi(t,\xi) + \int_0^t \int_D K(t,\xi;s,\eta)x(s,\eta)dsd\eta, \xi \in D, t \geq 0, \tag{2}$$

where D is the spatial domain (like a rod or a plate, spatial region like the atmosphere around the planet earth, the human body, etc). The kernel K is given by appropriate Greens function determined by the homogeneous boundary conditions (such as Dirichlet or Neumann). Let H denote the Hilbert space $L_2(D)$ of Lebesgue measurable and square integrable functions defined on the open set D. Then define $Y(t) \equiv y(t,\cdot) \in H$ as an H-valued function of t for $t \geq 0$. . Similarly define $\hat{\Pi}(t) \equiv \Pi(t,\cdot) \in H$ and $X(t) \equiv x(t,\cdot) \in H$. Then the system (2) can be written in the abstract form

$$Y(t) = \hat{\Pi}(t) + \int_0^t \mathscr{K}(t,s)X(s)ds, t \geq 0, \tag{3}$$

as an integral equation on the Hilbert space H unlike the scalar case (1) with $\mathscr{K}(t,s)$ being the operator valued kernel that maps H to H. In fact, it has the following representation

$$(\mathscr{K}(t,s)\varphi)(\xi) \equiv \int_D K(t,\xi;s,\zeta)\varphi(\zeta)d\zeta.$$

Thus it is clear that the abstract equation (3) has similar algebraic form as (1) but with different interpretation. First we consider the scalar case and then in the final section we extend the results to infinite dimensional Hilbert spaces. Equation (1) is linear in x and can indeed cover a very large class of linear systems.

In 1883, it was the Italian mathematician, Vito Volterra, who introduced the notion of functionals which he considered as functions of functions, lines, area, volume etc. Here y, given by the expression (1), can be considered as a function (functional) of x, which itself is a function defined for $t \geq 0$, and this is linear. We are predominantly interested in nonlinear functionals or operators given by multiple integrals as follows

$$y(t) = F(x)(t), t \in I \equiv [0,T], T < \infty, \tag{4}$$

with the operator F given by the sum of multiple Lebesgue integrals of the input process x,

$$F(x)(t) \equiv \sum_{n=1}^{\infty} \int_{I_t^n} K_n(t, s_1, s_2, \cdots, s_n) x(s_1) x(s_2) \cdots x(s_n) ds_1 ds_2 \cdots ds_n, \tag{5}$$

where $I_t^n \equiv \otimes^n I_t$ is the cartesian product of n-copies of the interval $I_t = [0,t]$ and K_n is the n-th order kernel which is a function of $n+1$ variables. This generalizes the linear input-output system (1). The expression (5), known as Volterra series, can be used to model a very large class of nonlinear systems. The convergence of the series on the space of continuous functions $C(I,R)$ was studied by Fréchet and can be found in [1, p21]. This was extended to L_p spaces by the author in [2, Theorem 2, p147] where it was shown that any continuous functional defined on a closed bounded subset of the L_p space, $1 \leq p < \infty$, can be approximated as closely as necessary by functional polynomials of the form (5). Sodorov, in his recent monograph [5], has considered some interesting applications of Volterra series for modeling integral dynamics of physical systems applied to power engineering and image processing. Volterra series has been extensively used in the representation of strong solutions of stochastic differential equations and stochastic functional differential equations driven by Wiener process as well as Lévy process [4] and the references therein. Also it has been used in the study of generalized functionals of Brownian motion or Wiener process [4].

The rest of the paper is organized as follows; Section 2, is devoted to the analysis of the input-output model based on infinite Volterra series on spaces like $L_2(I)$, $C(I)$ and $L_\infty(I)$. In section 3, we consider nonlinear integral equation of the second kind arising from the output feedback and prove existence of solutions and regularity properties thereof. This section is concluded with a nonlinear output-feedback operator and a result on the question of existence of solution of the resulting integral equation. In section 4, we consider optimal control of the nonlinear system and prove existence of suboptimal control policy from a closed ball in L_2 space. Following this, we consider relaxed or measure valued controls on the closed ball in L_2 space and prove existence of optimal relaxed controls. In section 5, we consider stochastic input and present existence of solutions using stopping time arguments. Then we consider the inverse problem to identify the kernels of the Volterra series. The technique for identification of kernels of Volterra series proposed by Sodorov [6, p136-] is rather complicated. Here in this paper we use stochastic technique to identify all the kernels in a much more simpler way. In section 6, we extend all the preceding results to infinite dimensional Hilbert spaces using tensor products.

2. Nonlinear systems: input-output models

In this section we consider the input-out model (4) and present sufficient conditions on the kernels $\{K_n\}$ that guarantee convergence of the series and input-output stability of the system. For simplicity, we consider only $L_2(I) \equiv L_2([0,T])$ space as the domina and range space. Let $\{L_n\}$ denote the sequence of Lebesgue measurable functions defined on the cartesian product of n copies of the interval I denoted by $\{I^n\}$ and that $L_n \in L_2(I^n)$ for each $n \in N$. Consider the vector space (of infinite sequence of Kernels)

$$X \equiv \left\{ L \equiv \{L_1, L_2, L_3, \cdots, L_n \cdots\}, L_n \in L_2(I^n), n \in N \equiv \{1,2,3\cdots\} \right\}. \tag{6}$$

Clearly, for $L \in X$ and $\alpha \in R$, $\alpha L \in X$. For $K, L \in X$, it is clear that $K + L \in X$. Thus X is a linear vector space. Let us introduce a norm on it as follows:

$$\| L \|_X \equiv \left(\sum_{n=1}^{\infty} n! \| L_n \|^2_{L_2(I^n)} \right)^{1/2},$$

where

$$\| L_n \|^2_{L_2(I^n)} = \int_{I^n} |L_n(s_1, s_2, \cdots, s_n)|^2 ds_1 ds_2 \cdots ds_n.$$

From now on we assume that X is equipped with this norm topology $(X, \| \cdot \|_X)$. Completion of the vector space X with respect to this norm topology is a Hilbert space and we continue to denote this by simply the same character X. The inner product in this space is given by

$$< L, K >_X = \sum_{n=1}^{\infty} n! (L_n, K_n)_{L_2(I^n)}.$$

Clearly, it follows from this that $| < L,K >_X | \leq \| L \|_X \| K \|_X$ for all $L,K \in X$. For the input-output model given by (4)-(5) (representing time-varying systems) we need to consider time-varying kernels $\mathscr{K} : I \longrightarrow X$ satisfying

$$\| \mathscr{K} \|_{L_2(I,X)} \equiv \left(\int_I \| \mathscr{K}(t) \|_X^2 \, dt < \infty \right)^{1/2}.$$

Here $\mathscr{K}(t)$ stands for the X valued kernel given by

$$\mathscr{K}(t) = \{K_1(t;\cdot), K_2(t;\cdot,\cdot), K_3(t;\cdot,\cdot,\cdot), \cdots, K_n(t;\cdots), \cdots\cdots\}.$$

Note that here the sequence of kernels $\{K_n\}$ are functions of $(n+1)$ variables, that is,

$$\{K_n(t;s_1,s_2,s_3,\cdots s_n), s_i \in [0,t]\}, t \in I, n \in N.$$

Throughout the rest of the paper we may assume that the kernels $\{K_n\}$ are causal in the sense that $K_n(t,s_1,s_2,\cdots,s_n) = 0$ for $s_i > t$ for any $1 \leq i \leq n$. However, the presentation is independent of whether the kernels are causal or not. First we present the following result.

Theorem 2.1. *Consider the input-output model given by the equations (4)-(5) and suppose the kernel* $\mathscr{K}$, *determining the integral operator* F, *is* X *valued causal and* $\mathscr{K} \in L_2(I,X)$. *Then, for every input* $x \in L_2(I)$, *the output* $y \in L_2(I)$. *In other words* $F : L_2(I) \longrightarrow L_2(I)$.

Proof. For any $n \in N$, consider the n-th term of F given by

$$F_n(x)(t) \equiv \int_{I_t^n} K_n(t;s_1,s_2,\cdots,s_n)x(s_1)x(s_2)\cdots x(s_n)ds_1ds_2\cdots,ds_n, t \in I. \tag{7}$$

It follows from Hölder inequality that

$$|F_n(x)(t)| \leq \left(\int_{I_t^n} |K_n(t;s_1,s_2,s_3,\cdots,s_n)|^2 ds_1ds_2,\cdots,ds_n \right)^{1/2} \bullet$$
$$\bullet \left(\int_{I_t} |x(s)|^2 ds \right)^{n/2}. \tag{8}$$

Now integrating the square of both sides of the above inequality we obtain

$$\int_I |F_n(x)(t)|^2 dt \leq \int_I \left(\int_{I_t^n} |K_n(t;s_1,s_2,\cdots,s_n)|^2 ds_1ds_2\cdots ds_n \right) dt \bullet \left(\int_I |x(s)|^2 ds \right)^n. \tag{9}$$

Clearly, it follows from the above inequality and the causality assumption that

$$\| F_n(x) \|^2_{L_2(I)} \leq \| K_n \|^2_{L_2(I^{n+1})} (\| x \|_{L_2(I)})^{2n} \tag{10}$$

and hence we have the following inequality,

$$\| F_n(x) \|_{L_2(I)} \leq \| K_n \|_{L_2(I^{n+1})} (\| x \|_{L_2(I)})^{n}. \tag{11}$$

This shows that each component F_n of the operator F maps $L_2(I)$ to $L_2(I)$. Using this estimate and the triangle inequality it follows from the expression for the operator F that

$$\| F(x) \|_{L_2(I)} \leq \sum_{n=1}^{\infty} \| K_n \|_{L_2(I^{n+1})} \| x \|^n_{L_2(I)} . \tag{12}$$

Now using Hölder inequality it follows from (12) that

$$\begin{aligned} \| F(x) \|_{L_2(I)} &\leq \sum_{n=1}^{\infty} \sqrt{n}! \| K_n \|_{L_2(I^{n+1})} \frac{\| x \|^n_{L_2(I)}}{\sqrt{n}!} \\ &\leq \left(\sum_{n=1}^{\infty} n! \| K_n \|^2 \right)^{1/2} \left(\sum \frac{\| x \|^{2n}_{L_2(I)}}{n!} \right)^{1/2} . \end{aligned} \tag{13}$$

Since by our assumption $\mathscr{K} \in L_2(I,X)$, it follows from the above inequality that

$$\begin{aligned} \| F(x) \|_{L_2(I)} &\leq \| \mathscr{K} \|_{L_2(I,X)} \left(\exp \| x \|^2_{L_2(I)} - 1 \right)^{1/2} \\ &\leq \| \mathscr{K} \|_{L_2(I,X)} \| x \| \exp(\| x \|^2_{L_2(I)} /2). \end{aligned} \tag{14}$$

This shows that the nonlinear operator F maps $L_2(I)$ space into itself. And, further, the L_2 norm of the output y is bounded above as follows

$$\| y \|_{L_2(I)} \leq \| \mathscr{K} \|_{L_2(I,X)} \| x \| \exp(\| x \|^2_{L_2(I)} /2).$$

This proves that whenever the input signal $x \in L_2(I)$, the output signal $y \in L_2(I)$ also. This completes the proof. •

Another interpretation of the above result is that the input-output model given by the infinite Volterra series, $x \longrightarrow F(x) \equiv y$, is stable in the L_2 space mapping norm bounded sets into norm bounded sets. In other words, both the domain and the range of the nonlinear operator F is the whole L_2 space. In engineering literature, this is known as bounded-input bounded-output (BIBO) stability.

In case the kernels $\{K_n\}$ are continuous in $t \in I$ with values in X, the system (4)-(5) maps $L_2(I)$ space into the Banach space $C(I)$ with the usual sup-norm topology. Indeed,

$$\sup_{t\in I}|F(x)(t)| \le \sup_{t\in I} \parallel \mathscr{K}(t) \parallel_X \left(\exp(\parallel x \parallel^2_{L_2(I)}) - 1\}\right)^{1/2}. \tag{15}$$

Further we can also consider both the input and output spaces to be the space of continuous functions $C(I)$ with the standard sup-norm topology. In this case $x \in C(I)$ and the output $y \in C(I)$ also. For each $n \in N$, define the $L_1(I^n)$-valued function $\hat{K}_n(t) \equiv K_n(t, \cdots)$. That is

$$|\hat{K}_n(t)|_{L_1(I^n)} = \left(\int_{I^n} |K_n(t, s_1, s_2, s_3, \cdots, s_n)| ds_1 ds_2 \cdots ds_n\right) < \infty.$$

We prove the following result for this case.

Corollary 2.2. *Suppose that for each $n \in N$, and $t \in I$, the kernel $\hat{K}_n(t) \in L_1(I^n)$ and that*

$$\left(\sum n! \big(\sup_{t\in I} \parallel \hat{K}_n(t) \parallel_{L_1(I^n)}\big)^2\right)^{1/2} \equiv b < \infty, \tag{16}$$

and the function

$$t \longrightarrow \parallel \hat{K}_n(t) \parallel_{L_1(I^n)} \equiv \int_{I^n} |K_n(t, s_1, s_2, \cdots, s_n)| ds_1 ds_2 \cdots, ds_n, \tag{17}$$

is continuous on I. Then $F : C(I) \longrightarrow C(I)$.

Proof. First note that

$$|F_n(x)(t)| \le \left(\int_{I^n} |K_n(t; s_1, s_2, \cdots, s_n)| ds_1 ds_2 \cdots ds_n\right) \left(\sup_{s\in I} |x(s)|\right)^n,$$

$$\le \parallel \hat{K}_n(t) \parallel_{L_1(I^n)} (\parallel x \parallel_{C(I)})^n. \tag{18}$$

Hence, following similar steps as before, we obtain the following estimate

$$\sup_{t\in I}|F(x)(t)| \le \left(\sum n! \big(\sup_{t\in I} \parallel \hat{K}_n(t) \parallel_{L_1(I^n)}\big)^2\right)^{1/2} \left(\sum_{n=1}^{\infty} \frac{\parallel x \parallel^{2n}_{C(I)}}{n!}\right)^{1/2}. \tag{19}$$

Now it follows from the continuity hypothesis (17) that $t \longrightarrow F(x)(t)$ is continuous and bounded in norm as shown in (19). Further, it follows from the above estimate that

$$\sup_{t\in I}|F(x)(t)| \le b\left(\exp\{\parallel x \parallel^2_{C(I)}\} - 1\right)^{1/2}. \tag{20}$$

Hence we conclude that $F : C(I) \longrightarrow C(I)$. This completes the proof. •

In case of essentially bounded inputs we use the space of essentially bounded Lebesgue measurable functions $L_\infty(I)$ in place of the space $C(I)$.

Corollary 2.3. *Suppose that for each $n \in N$, the kernel $\hat{K}_n(\cdot) \in L_\infty(I, L_1(I^n))$ and that*

$$\Big(\sum n!\big(\| \hat{K}_n(\cdot) \|_{L_\infty(I,L_1(I^n))}\big)^2\Big)^{1/2} \equiv b < \infty,$$

Then the operator $F : L_\infty(I) \longrightarrow L_\infty(I)$ and, it is continuous, and bounded on bounded sets.

3. From output feedback to nonlinear integral equations

Consider the input-output model given by (4)-(5) and let $z \in L_2(I)$ denote the exogenous input of the feedback system described by the following chain of relations (composition maps)

$$x \longrightarrow F(x) = y \longrightarrow \lambda y + z = x \tag{21}$$

completing the feedback loop with λ denoting the feedback gain. The feedback signal given by λy, where y is the output, is added to the exogenous input z to produce the signal x driving the system F. In other words the feedback system is governed by the following nonlinear integral equation of the second kind

$$x = z + \lambda F(x), \tag{22}$$

where the operator F is given by the expression (5) with values $F(x)(t) \in R$ for all $t \geq 0$. This is a nonlinear Volterra integral equation of the second kind. We show that, under certain assumptions, the feedback scheme is viable in the sense that equation (22) has a unique solution in the Hilbert space $L_2(I)$. Let $B_r(L_2)$ denote the closed ball in $L_2(I)$ of radius r centered at the origin. That is,

$$B_r(L_2) \equiv \{u \in L_2(I) : \| u \|_{L_2(I)} \equiv \big(\int_I |u(t)|^2 dt\big)^{1/2} \leq r\}.$$

Theorem 3.1. *Consider the feedback system (22) and suppose the assumptions of Theorem 2.1 hold. Then, for any (input signal) $z \in B_{\alpha r}(L_2)$ for any $r \in (0, \infty)$ and $\alpha \in (0,1)$; and λ sufficiently small (depending on r), the integral equation (22) has a unique solution $x \in B_r(L_2)$.*

Proof. Define the operator G_λ by

$$G_\lambda(x) \equiv z + \lambda F(x), \lambda > 0, x \in L_2(I). \tag{23}$$

From our preceding analysis, it is easy to verify that for any $x \in L_2$, we have

$$\| G_\lambda(x) \|_{L_2} \leq \| z \| + |\lambda| \, \| \mathscr{K} \|_{L_2(I,X)} \| x \| \exp\{(1/2) \| x \|^2\}. \tag{24}$$

For any $r > 0$, $\alpha \in (0,1)$, $z \in B_{\alpha r}(L_2)$ and $x \in B_r(L_2)$, it follows from the inequality (24) that

$$\| G_\lambda(x) \|_{L_2} \leq \alpha r + |\lambda| \, \| \mathscr{K} \|_{L_2(I,X)} \, r \exp(r^2/2). \tag{25}$$

Thus for λ sufficiently small satisfying

$$|\lambda| \leq \frac{(1-\alpha)}{\| \mathscr{K} \| \exp(r^2/2)}, \tag{26}$$

we have $\| G_\lambda(x) \|_{L_2} \leq r$. In other words G_λ maps B_r into itself. We show now that G_λ is locally Lipschitz, in particular, on the ball $B_r(L_2)$ for any finite positive number r. First note that

$$\begin{aligned} &|\Pi_{i=1}^n x_2(s_i) - \Pi_{i=1}^n x_1(s_i)| \\ &\quad = \sum_{m=0}^{n-1} \left(\{\Pi_{i=1}^m |x_2(s_i)|\}\{|x_2(s_{m+1}) - x_1(s_{m+1})|\}\{\Pi_{i=m+2}^n |x_1(s_i)|\} \right) \end{aligned} \tag{27}$$

with the convention that, for all $q < k$, $\Pi_{i=k}^q \eta_i = 1$; and, for $q \geq k$, $\Pi_{i=k}^q \eta_i = \eta_k \eta_{k+1} \cdots \eta_q$ where $\{\eta_i\}$ are real numbers. Now using the expression (27) and straightforward but laborious algebraic manipulations one can verify that

$$\int_{I_t^n} |\Pi_{i=1}^n x_2(s_i) - \Pi_{i=1}^n x_1(s_i)|^2 ds_1 ds_2 \cdots ds_n \leq n r^{2(n-1)} \| x_2 - x_1 \|_{L_2}^2 \tag{28}$$

for all $x_1, x_2 \in B_r(L_2)$. Thus it follows from Hölders inequality, applied to the expression $F_n(x_2) - F_n(x_1)$, that

$$\begin{aligned} |F_n(x_2)(t) - F_n(x_1)(t)| &\leq \| \hat{K}_n(t) \|_{L_2(I_t^n)} \left(n r^{2(n-1)} \| x_2 - x_1 \|_{L_2(I)}^2 \right)^{1/2} \\ &\leq \| \hat{K}_n(t) \|_{L_2(I^n)} \left(n r^{2(n-1)} \| x_2 - x_1 \|_{L_2(I)}^2 \right)^{1/2}. \end{aligned} \tag{29}$$

Now considering the infinite series $[F(x_2)-F(x_1)]$ and computing it's L_2 norm it follows from Hölder and Schwartz inequalities that

$$\| F(x_2)-F(x_1) \|^2_{L_2(I)} \leq \left(\sum_{n=1}^{\infty} n! \int_I \| \hat{K}_n(t) \|^2_{L_2(I^n)} \, dt\right) \bullet$$

$$\bullet\left(\sum_{n=1}^{\infty} (n/n!) r^{2(n-1)}\right) \| x_2-x_1 \|^2_{L_2(I)} . \tag{30}$$

Hence

$$\| F(x_2)-F(x_1) \|_{L_2(I)} \leq \| \mathscr{K} \|_{L_2(I,X)} \exp(r^2/2) \| x_2-x_1 \|_{L_2(I)} . \tag{31}$$

Using the above estimate and the expression for the operator G_λ, we arrive at the following inequality,

$$\| G_\lambda(x_2)-G_\lambda(x_1) \|_{L_2(I)} \leq |\lambda| \| \mathscr{K} \|_{L_2(I,X)} \exp(r^2/2) \| x_2-x_1 \|_{L_2(I)} . \tag{32}$$

Thus for λ sufficiently small satisfying the inequality (26) we have

$$\| G_\lambda(x_2)-G_\lambda(x_1) \|_{L_2(I)} \leq (1-\alpha) \| x_2-x_1 \|_{L_2(I)}, \forall\, x_1,x_2 \in B_r(L_2). \tag{33}$$

Collecting the above results we conclude that, for λ sufficiently small satisfying the bound (26), the operator $G_\lambda : B_r(L_2) \longrightarrow B_r(L_2)$ and that it is a contraction on $B_r(L_2)$. Hence it follows from Banach fixed point theorem (contraction mapping principle) that the integral equation (22) has a unique solution $x^* \in B_r(L_2)$. •

Under the assumptions of Theorem 3.1, the solution x^*, whose existence is guaranteed by the Banach fixed point theorem, can actually be computed by successive approximation. This is presented in the following Corollary.

Corollary 3.2. *Suppose the assumptions of Theorem 3.1 hold. Then the solution x^* can be constructed by successive approximation.*

Proof. For the given data $z \in B_{\alpha r}(L_2)$, define the first approximation by $x_1 \equiv z$ and then define the sequence $\{x_n\}$ as follows

$$x_2 = z+\lambda F(x_1),\ x_3 = z+\lambda F(x_2),\ x_{n+1} = z+F(x_n), n \in N. \tag{34}$$

Under the assumptions of Theorem 3.1, it is easy to verify that for any $m \in N$

$$\| x_{m+1} - x_m \| \leq (1-\alpha)^m \| z \| . \tag{35}$$

Hence for any $p(\in N) \geq 1$ and $n \in N$, we have

$$\begin{aligned} \| x_{n+p} - x_n \| &\leq \sum_{m=n}^{n+p-1} \| x_{m+1} - x_m \| \leq \sum_{m=n}^{n+p-1} (1-\alpha)^m \| z \| \\ &\leq [(1-\alpha)^n/\alpha] \| z \| . \end{aligned} \tag{36}$$

Since $\alpha \in (0,1)$, it follows from the above inequality that, for any $p \geq 1$,

$$\lim_{n\to\infty} \| x_{n+p} - x_n \| = 0.$$

Hence $\{x_n\}$ is a Cauchy sequence in L_2 (a Banach space) and therefore there exists a unique element $x^o \in L_2$ to which it converges. The reader can easily verify that x^o coincides with the fixed point x^* of the operator G_λ as seen in Theorem 3.1. This completes the proof. •

In the following corollary, we state a result on continuous dependence of solutions of the integral equation (22) on the exogenous input z. From this result follows the continuous dependence of the output of the feedback system (21) (governed by the integral equation (22)) on the input z.

Corollary 3.3. *Under the assumptions of Theorem 3.1, the solution x of the integral equation (22) (feedback system) and the corresponding output y are not only continuously dependent on the exogenous input; they are also locally Lipschitz.*

Proof. Consider the integral equation (22) and let $z_1, z_2 \in B_{\alpha r}(L_2)$ be any two exogenous inputs. By virtue of Theorem 3.1, there exist $x_1, x_2 \in B_r(L_2)$ which are the unique solutions of equation (22) corresponding to the inputs z_1, z_2 respectively. Then it follows from (22) and (31) that

$$\begin{aligned} \| x_1 - x_2 \| &\leq \| z_1 - z_2 \| + |\lambda| \, \| F(x_1) - F(x_2) \| \\ &\leq \| z_1 - z_2 \| + |\lambda| \, \| \mathscr{K} \|_{L_2(I,X)} \exp(r^2/2) \| x_1 - x_2 \| . \end{aligned} \tag{37}$$

Hence for λ satisfying (26) it follows from this inequality that

$$\| x_1 - x_2 \| \leq (1/\alpha) \| z_1 - z_2 \| . \tag{38}$$

Since $\alpha \in (0,1)$ this proves that the input to solution map $z \longrightarrow x$ is Lipschitz continuous. We may denote this map by Φ and write $x = \Phi(z)$. So the map Φ is locally Lipschitz on $B_{\alpha r}(L_2)$. Note that for the feedback system (21) the output is denoted by y and it is given by $y = F(x)$. Thus it follows from the preceding analysis that the outputs $\{y_1, y_2\}$ corresponding to inputs $\{z_1, z_2\}$ satisfy the following inequalities:

$$\begin{aligned} \parallel y_1 - y_2 \parallel = \parallel F(x_1) - F(x_2) \parallel &\leq \parallel \mathscr{K} \parallel_{L_2(I,X)} \exp(r^2/2) \parallel x_1 - x_2 \parallel \\ &\leq L(r) \parallel z_1 - z_2 \parallel, \end{aligned} \tag{39}$$

where $L(r) \equiv \{(1/\alpha) \parallel \mathscr{K} \parallel_{L_2(I,X)} \exp(r^2/2)\}$. This shows that the input to output map $z \longrightarrow y$ is also locally Lipschitz. We denote this map by Ψ and write $y = \Psi(z)$. This completes the proof. •

If the input to the feedback system (21) denoted by z is identically zero, we obtain the homogeneous integral equation $x = \lambda F(x)$. This equation has the trivial solution $(x \equiv 0)$ as proved in the following corollary.

Corollary 3.4. *Under the assumptions of Theorem 3.1, the homogeneous integral equation,* $x = \lambda F(x) \equiv G^o_\lambda(x)$, *has only the trivial solution* $x \equiv 0$.

Proof. By virtue of the inequality (26), G^o_λ maps $B_r(L_2)$ into itself. Further, for $x_1, x_2 \in B_r$, we have $\parallel G^o_\lambda(x_1) - G^o_\lambda(x_2) \parallel \leq (1-\alpha) \parallel x_1 - x_2 \parallel$. Since $\alpha \in (0,1)$, G^o_λ is a contraction on B_r and hence, again by Banach fixed point theorem, it has a unique fixed point $x^o \in B_r$. The fact that $x^o \equiv 0$ follows from uniqueness of solution and the fact that $x \equiv 0$ is a solution of the homogeneous equation. This completes the proof. •

The feedback system represented by the chain of relations (21) can be generalized to allow nonlinear operator in the feedback link. Let $h : I \times R \longrightarrow R$ be a Carathéodory function. That is, h is measurable in the first argument and continuous in the second. The nonlinear operator H (known as Nemytski operator) given by

$$H(y)(t) \equiv h(t, y(t)) \tag{40}$$

represents the feedback operator giving the following complete feedback system:

$$x \longrightarrow F(x) = y \longrightarrow H(y) + z = x, \tag{41}$$

where z is the exogenous input and y is the output. Thus the feedback system can be abstractly represented by the relation $x = z + HoF(x) = z + H(F(x))$. The equivalent integral equation arising from the above model can be written as

$$x(t) = z(t) + h(t, F(x)(t)), t \in I. \tag{42}$$

Following the same procedure as in Theorem 3.1 one can prove the following result.

Theorem 3.5. *Consider the integral equation (42) and suppose the nonlinear operator F satisfies the assumptions of Theorem 3.1 and the Carathèodory function h satisfy the following conditions: there exists an $a \in L_2^+(I)$ and $b, \eta \geq 0$ such that*

$$(H1) : |h(t,\xi)| \leq a(t) + b|\xi| \; \forall \; \xi \in R; \tag{43}$$

$$(H2) : |h(t,\xi) - h(t,\zeta)| \leq \eta|\xi - \zeta| \; \forall \; t \in I and \; \xi, \zeta \in R. \tag{44}$$

Further, suppose that for any finite $r > 0$, and a $\theta \in (0,1)$,

$$0 < b \leq \left(\frac{(1-\theta)r}{\| \mathscr{K} \|_{L_2(I,X)} (\exp(r^2/2) - 1)} \right)$$
$$0 < \eta < \frac{\theta}{\| \mathscr{K} \|_{L_2(I,X)} \exp(r^2/2)}.$$

Then, for any exogenous input $z \in L_2(I)$ satisfying $\| z \| + \| a \| \leq \theta r$, the integral equation (42) has a unique solution $x \in B_r(L_2)$.

Proof. The proof is similar to that of Theorem 3.1.

4. Optimal input (or control) for the feedback system

Consider the feedback system (21) governed by the nonlinear integral equation (22) with output y given by $y = F(x)$. Let $y_d \in L_2(I)$ denote the desired output trajectory. The objective is to find an exogenous input called control $z \in B_{\alpha r}(L_2(I))$ to minimize the following cost functional,

$$J(z)=\int_0^T |y(t)-y_d(t)|^2dt+\int_0^T |z(t)|^2dt. \tag{45}$$

By virtue of Corollary 3.3, the output y is given by $y=\Psi(z)$ and hence the cost functional $J(z)$ is given by

$$J(z)=\parallel \Psi(z)-y_d \parallel^2+\parallel z \parallel^2, \tag{46}$$

where Ψ is the input-output map. Since Ψ is Lipschitz continuous it is easy to verify that $z \longrightarrow J(z)$ is continuous on $B_{\alpha r}(L_2)$. Indeed, let $z_n \stackrel{s}{\longrightarrow} z_0$ in L_2 and recall that for $u,v \in L_2$, we have

$$|\parallel u \parallel - \parallel v \parallel| \leq \parallel u-v \parallel .$$

Using this inequality and the expression (46) one can easily prove that $\lim_{n\to\infty} J(z_n)=J(z_0)$. Thus J is (strongly) continuous on $B_{\alpha r}$. Since the ball $B_{\alpha r} \equiv B_{\alpha r}(L_2) \subset L_2(I)$ is not compact (only weakly compact), we can not assert from the continuity of J alone that it attains its minimum on $B_{\alpha r}$. However, we can consider the problem of ε-optimality. Consider the optimization problem (45) subject to the feedback system (21) governed by the integral equation (22). Let m^* denote the infimum of J defined by

$$m^* \equiv \inf\{J(z); z \in B_{\alpha r}\}.$$

The problem is that there may not exist any element $z \in B_{\alpha r}$ at which the infimum is attained. So we may consider the possibility of a near optimal solution.

Definition 4.1. If for every $\varepsilon>0$, there exists an element $z_\varepsilon \in B_{\alpha r}$ such that $J(z_\varepsilon) \leq m^*+\varepsilon$, then the element z_ε is said to be ε-optimal.

We prove the following result on ε-optimality.

Theorem 4.2. *Consider the optimization problem (45/46) subject to the feedback system (21) governed by the integral equation (22) and the set of admissible inputs $B_{\alpha r}(L_2)$. Suppose the assumptions of Theorem 3.1 hold. Then the optimization problem has an ε-optimal solution.*

Proof. Let $\{e_i, i \in N\}$ be a complete ortho-normal basis for the Hilbert space $L_2(I)$. For each $n \in N$, define the set

$$M_n \equiv \overline{\text{lin-span}}\{e_i, i = 1, 2, \cdots, n\}.$$

This is a closed linear subspace of $L_2(I)$ having (finite) dimension n. Introduce the set

$$\Gamma_n \equiv M_n \cap B_{\alpha r}.$$

Since $B_{\alpha r}$ is a closed bounded subset of L_2 space, the set Γ_n is a closed bounded subset of a finite dimensional subspace of the L_2 space. Thus it is compact with respect to the (strong) norm topology. As seen above, the cost functional J is continuous on $B_{\alpha r}$ in the norm topology and therefore it attains its minimum on Γ_n for each $n \in N$. Define

$$m_n \equiv \min\{J(z), z \in \Gamma_n\}$$

and let $z_n \in \Gamma_n$ be a corresponding minimizer giving $J(z_n) = m_n$. Note that $\{\Gamma_n\}$ is an increasing sequence of compact subsets of $B_{\alpha r}$ and since $\{e_i\}$ is a complete ortho-normal basis of the Hilbert space $L_2(I)$, it is clear that

$$\Gamma_n \longrightarrow B_{\alpha r}, \text{ as } n \to \infty.$$

Since Γ_n is an increasing sequence of compact sets, $\{m_n\}$ is a monotone decreasing sequence of nonnegative real numbers. Therefore, there exists a number $m^* \geq 0$ such that $m_n \to m^*$ as $n \to \infty$. On the other hand the sequence $\{z_n\} \in \Gamma_n \subset B_{\alpha r}$ may not converge strongly. However, since $B_{\alpha r}$ is weakly compact there exists a subsequence $\{z_{n_k}\} \subset \{z_n\}$ and an element $z_o \in B_{\alpha r}$ such that $z_{n_k} \stackrel{w}{\longrightarrow} z_o$. It is clear that

$$J(z_{n_k}) = m_{n_k} \longrightarrow m^*, \text{ as } k \to \infty.$$

Since J is only continuous in the norm topology and z_{n_k} converges to z_o weakly (as $k \to \infty$) the sequence $J(z_{n_k}) \not\to J(z_o)$ (may not converge). However, given $\varepsilon > 0$, there exists an integer $n_\varepsilon \in N$ such that

$$J(z_{n_k}) \leq m^* + \varepsilon, \text{ for all } n_k \geq n_\varepsilon.$$

Thus the element $z_{n_\varepsilon} \in B_{\alpha r}$ is an ε-optimal input. This completes the proof. •

Remark 4.3. Continuity in the strong topology is incompatible with the topology of weak convergence. This is due to the fact that, in general, weak convergence is incompatible with nonlinear composition. This is why we can not claim that $J(z_o) = m^*$. In other words, there may not exist any regular control (input), that is, controls from $B_{\alpha r}(L_2(I))$.

Existence of optimal relaxed control (or input)

Consider the L_2 space endowed with the weak topology and suppose the closed ball $B_{\alpha r}(L_2) \equiv \mathbf{S} \subset L_2(I)$ is given the relative weak topology. With respect to the weak topology it is a compact Hausdorff space. Let $\mathscr{S}$ denote the algebra of subsets of the set $\mathbf{S}$. Consider the vector space $B(\mathbf{S})$ of bounded $\mathscr{S}$ measurable real valued functions defined on $\mathbf{S}$. Equipped with the supnorm topology, $B(\mathbf{S})$ is a Banach space. It is known [Dunford 5, Theorem 1, Corollary 3, p258] that the topological dual $B^*(\mathbf{S})$ of $B(\mathbf{S})$ is given by the space of finitely additive measures on $\mathscr{S}$ having bounded total variation. We denote this space by $M_{ba}(\mathbf{S})$. This is a Banach space with respect to the norm topology determined by the total variation norm. Any continuous linear functional ℓ on $B(\mathbf{S})$ has the integral representation given by

$$\ell(\varphi) = \int_{\mathbf{S}} \varphi(s)\mu(ds), \varphi \in B(\mathbf{S})$$

for a certain $\mu \in M_{ba}(\mathbf{S})$. In fact, every $\mu \in M_{ba}(\mathbf{S})$ determines a continuous linear functional on $B(\mathbf{S})$ through the above integral expression, and that, for every continuous linear functional on $B(\mathbf{S})$, there exists a unique $\mu \in M_{ba}(\mathbf{S})$ determining the functional through the same integral expression. In other words we have an isometric isomorphism: $B^*(\mathbf{S}) \cong \mathbf{M_{ba}}(\mathbf{S})$. Let $M_1(\mathbf{S}) \subset \mathbf{M_{ba}}(\mathbf{S})$ denote the space of finitely additive probability measures on $\mathbf{S}$ (more precisely on $\mathscr{S}$). Since $y_d \in L_2(I)$, the functional J given by (45/46) is bounded on $B_{\alpha r}$ and so bounded on $\mathbf{S}$ irrespective of the topology. Since J is not continuous in the weak topology, it is not continuous on $\mathbf{S}$ but $J \in B(\mathbf{S})$. We have seen in Theorem 4.2 that $z_{n_k} \xrightarrow{w} z_o$ but $J(z_{n_k})$ may not converge to $J(z_o)$ as J is not weakly continuous. However, if we dismiss the requirement of (weak) continuity of J and relax the space of admissible inputs (controls) from $\mathbf{S}$ to the class of finitely additive measures on it, we can obtain convergence in a wider sense.

Consider the Dirac measures $\{\nu_k \equiv \delta_{z_{n_k}}\} \subset M_1(\mathbf{S})$ supported on the one point sets $\{z_{n_k}\} \subset \mathbf{S}$. With this notation, the cost functional evaluated at z_{n_k} is given by the integral

$$J(z_{n_k}) = \int_{\mathbf{S}} J(\eta)\nu_k(d\eta).$$

Let $\mathscr{E}(K)$ denote the set of extreme points of any compact convex set K of any locally convex Hausdorff topological vector space. By Alaoglu's theorem, $M_1(\mathbf{S})$ is a weak star compact (convex) subset of $M_{ba}(\mathbf{S})$. Thus it follows from Krein-Milman theorem [Dunford-Schwartz,[4],Theorem V.8.4, p440] that $M_1(\mathbf{S}) = clco^{w^*}(\mathscr{E}(M_1(\mathbf{S})))$. Note that the Dirac measures $\{\nu_k \equiv \delta_{z_{n_k}}\}$ are the extreme points of $M_1(\mathbf{S})$, that is, $\{\nu_k\} \subset \mathscr{E}(M_1(\mathbf{S}))$. Since $M_1(\mathbf{S})$ is weak star compact, there exists a generalized subsequence of the sequence $\{\nu_k\}$, relabeled as the original sequence, and a $\nu_o \in M_1(\mathbf{S})$ such that

$$\lim_{k\to\infty} \int_{\mathbf{S}} \varphi(\eta)\nu_k(d\eta) = \int_{\mathbf{S}} \varphi(\eta)\nu_o(d\eta), \text{ for any } \varphi \in B(\mathbf{S}).$$

Since the cost functional $J \in B(\mathbf{S})$ the above identity also holds for $\varphi = J$. Thus

$$\lim_{k\to\infty} \int_{\mathbf{S}} J(\eta)\nu_k(d\eta) = \int_{\mathbf{S}} J(\eta)\nu_o(d\eta).$$

On the other hand, by virtue of Theorem 4.2, we also have $\lim_{k\to\infty} \int_{\mathbf{S}} J(\eta)\nu_k(d\eta) = \lim_{k\to\infty} m_{n_k} = m^*$ and hence

$$\int_{\mathbf{S}} J(\eta)\nu_o(d\eta) = m^*.$$

Thus we have proved that an optimal control (or input) exists in the class of relaxed controls. This is formally stated as follows.

Theorem 4.4. *Consider the feedback system (21) governed by the integral equation (22) and suppose the assumptions of Theorem 3.1 hold and let J, given by (46), denote the cost functional with the relaxed inputs $M_1(\mathbf{S})$ replacing the regular or classical inputs $\mathbf{S}$. Then, there exists an input (control) $\nu_o \in M_1(\mathbf{S})$ that minimizes the cost functional*

$$\hat{J}(\nu) \equiv \int_{\mathbf{S}} J(z)\nu(dz)$$

attaining the infimum m^*.

Remark 4.5. An interesting question is: does the relaxed control ν_o belong to the set $\mathscr{E}(M_1(\mathbf{S}))$ or, equivalently, does it have a point support. Since the set of extreme points, $\mathscr{E}(M_1(\mathbf{S}))$, is not generally closed, it is possible that $\nu_o \notin \mathscr{E}(M_1(\mathbf{S}))$. It may very well belong to the convex hull of $\mathscr{E}(M_1(\mathbf{S}))$. So we do not have an optimal policy in the set $\mathbf{S}$. This is the significance of the above theorem.

5. Stochastic input

Let $(\Omega, \mathscr{F}, \mathscr{F}_{t\geq 0}, P)$ be a complete filtered probability space where $\mathscr{F}$ is the sigma algebra of Borel subsets of the set Ω and $\mathscr{F}_{t\geq 0}$ is an increasing family of right continuous complete subsigma algebras of the sigma algebra $\mathscr{F}$ and P is the probability measure. Let $\Xi \equiv \{\xi(t), t \geq 0\}$ denote the class of real valued square integrable $\mathscr{F}_t$-adapted stochastic process defined on the filtered probability space such that for any finite positive number $b > 0$,

$$T_b = \inf\{t \geq 0 : \int_0^t |\xi(s)|^2 ds \geq b\}$$

is an $\mathscr{F}_t$-stopping time and $P\{T_b > 0\} = 1$.

Theorem 5.1. *Consider the input-output model given by the equations (4)-(5) and suppose the kernel $\mathscr{K}$, determining the integral operator F, is X valued and causal satisfying $\mathscr{K} \in L_2(I,X)$. Then, for every input process $x \in \Xi$ and $r > 0$, for which $T_r(x) > 0$ with probability one, the output process $y \in \Xi$ and satisfies*

$$P\left\{\int_0^{T_r \wedge T} |y(t)|^2 dt \leq C^2(r) < \infty\right\} = 1,$$

where $C(r) \equiv \| \mathscr{K} \|_{L_2(I,X)} \, r \exp(r^2/2)$.

Proof. In view of Theorem 2.1, the proof is intuitively obvious.

Let us now consider the integral equation (22) representing the feedback system (21). In view of the above result, the proof of the following theorem is also obvious. Define the stopping time T_α for the input process z as follows

$$T_\alpha(z) \equiv \inf\left\{\tau \geq 0, \int_0^\tau |z(t)|^2 dt \geq (\alpha r)^2\right\}$$

where $\alpha \in (0,1)$ and $r > 0$, finite.

Theorem 5.2. *Consider the integral equation (22) representing the feedback system (21) and suppose the kernel $\mathscr{K} \in L_2(I,X)$. Then, for λ sufficiently small, satisfying the bound (26) and* *any measurable input process* $\{z(t), t \geq 0\}$ *for which the stopping time* $T_\alpha(z) > 0$, *with probability one, the integral equation has a unique solution x satisfying*

$$P\left\{\int_0^{T_\alpha \wedge T} |x(t)|^2 dt \leq r^2 < \infty\right\} = 1.$$

Hence the output process y has the property,

$$P\{\int_0^{T_\alpha \wedge T} |y(t)|^2 dt \leq C^2(r) < \infty\} = 1.$$

Inverse (or identification) problems

For application of Volterra series in physical sciences, it is necessary to identify the kernel $\mathscr{K} = \{K_n, n \in N\}$ determining the nonlinear operator F given by the expression (5). To identify the kernel of a linear system, it suffices to determine the so called impulse response. In principle, the system is subjected to an impulsive force at time τ for $0 \leq \tau < t$ and the response at a later time t is recorded giving the kernel $K(t, \tau)$, a function of two variables $\{(t, \tau) : 0 \leq \tau \leq t < \infty\}$. In the case of nonlinear systems of the form (4)-(5) this is not possible. Instead, the impulsive force is replaced by white noise, generalized derivative of the Wiener process $\{w\}$. In this case the operator F, subjected to Wiener process w, is given by

$$F(\dot{w})(t) \equiv \sum_{n=1}^{\infty} H_n(w)(t), \tag{47}$$

where H_n represents the $n-th$ order homogeneous Wiener chaos given by the following multiple Wiener-Ito integral [4]

$$H_n(w)(t) = \int_{I^n} K_n(t, s_1, s_2, \cdots, s_n) dw(s_1) dw(s_2) \cdots dw(s_n). \tag{48}$$

In reference [4, S-4.4, p95] we have considered the problem of identification of the sequence of kernels $\{K_n\}$ determining the operator F. For simplicity one may assume that the kernels $\{K_n\}$ are symmetric in the last n variables $\{s_1, s_2, \cdots, s_n\} \in I^n$. If not one can always symmetrize it. For each $n \in N$, let $\hat{L}_2(I^n) \equiv \mathscr{H}_n$ denote the Hilbert space of real valued symmetric kernels defined on I^n (Cartesian product of n copies of the interval I) and let $\hat{K}_n : I \longrightarrow \mathscr{H}_n$ be an $\mathscr{H}_n$-valued function $\hat{K}_n(t) \equiv K_n(t; \cdot, \cdot, \cdots, \cdot)$ such that, for almost all $t \in I$, its $\mathscr{H}_n$ norm is given by

$$\| \hat{K}_n(t) \|_{\mathscr{H}_n} \equiv \left(\int_{I^n} |K_n(t; s_1, s_2, \cdots, s_n)|^2 ds_1 ds_2 \cdots ds_n \right)^{1/2} \tag{49}$$

and that this $\mathscr{H}_n$ norm is square integrable on I in the sense of Lebesgue. That is, $\hat{K}_n \in L_2(I, \mathscr{H}_n)$ with $\| \hat{K}_n \|^2_{L_2(I,\mathscr{H}_n)} = \int_I \| \hat{K}_n(t) \|^2_{\mathscr{H}_n} dt$.

Let $\Omega \equiv C_0(I)$ denote the space of continuous real valued functions defined on I starting from zero and μ denote the standard Wiener measure on it, more precisely, on the sigma algebra generated by cylinder sets. Let $L_2(\Omega, \mu) \equiv \mathscr{W}$ denote the Hilbert space of all random variables which are square integrable with respect to the Wiener measure. For any $\eta \in \mathscr{W}$, we write

$$\| \eta \|_{\mathscr{W}} = \left(\mathbf{E}^{\mu} |\eta|^2 \right)^{1/2} \equiv \left(\int_{\Omega} |\eta(w)|^2 \mu(dw) \right)^{1/2}.$$

A simple example is: $\eta(w) = H_n(w)$ given by the expression (48). Given the Wiener measure space, or equivalently the Brownian motion space, it is the kernel $\hat{L} \in L_2(I, \mathscr{H}_n)$ that determines the homogeneous chaos H_n. Therefore, it is more logical to denote this by $H_n(\hat{L})$ for $\hat{L} \in L_2(I, \mathscr{H}_n))$. This is true for each $n \in N$. Thus, using this notation and suppressing the wiener process w, the input-output model (47) can be rewritten as

$$F(\dot{w})(t) \equiv \Phi(\mathscr{K}(t)) \equiv \sum_{n=1}^{\infty} \Phi_n(\hat{K}_n(t)), t \in I, \tag{50}$$

where $\mathscr{K} = \{K_1, K_2, K_3, \cdots, K_n, \cdots\}$, and $\Phi_n(\hat{K}_n(t)) \equiv H_n(w)(t), t \in I$, given by the expression (48) with the kernel $\hat{K}_n$. It is now interesting to note that, for each $n \in N$, the map

$$\Phi_n : L_2(I, \mathscr{H}_n) \longrightarrow L_2(I, \mathscr{W})$$

is a bounded linear operator satisfying the following properties (for details see [4, p37] and [3]): For each $\hat{L} \in L_2(I, \mathscr{H}_n)$ and $\hat{K} \in L_2(I, \mathscr{H}_m)$, we have

$$< \Phi_n, \Phi_m >_{L_2(I,\mathscr{W})} \equiv \mathbf{E}^{\mu} \{\Phi_n(\hat{L}) \Phi_m(\hat{K})\} = \begin{cases} 0 & \text{if } n \neq m, \\ n! \int_I < \hat{L}(t), \hat{K}(t) >_{\mathscr{H}_n} dt & \text{if } n = m. \end{cases} \tag{51}$$

In case $n = m$ and $\hat{L} = \hat{K}$, it follows from the above expression that

$$< \Phi_n(\hat{L}), \Phi_n(\hat{L}) >_{L_2(I,\mathscr{W})} \equiv n! \int_I \| \hat{L}(t) \|^2_{\mathscr{H}_n} dt \text{ for } \hat{L} \in L_2(I, \mathscr{H}_n). \tag{52}$$

Thus, for any $\hat{L} \in L_2(I, \mathscr{H}_n)$, we have

$$\| \Phi_n(\hat{L}) \|_{L_2(I,\mathscr{W})} = \sqrt{n!} \, \| \hat{L} \|_{L_2(I,\mathscr{H}_n)}$$

and therefore $\| \Phi_n \|_{\mathscr{L}(L_2(I,\mathscr{H}_n),L_2(I,\mathscr{W}))} = \sqrt{n!}$. Consider the Hilbert space given by the direct sum, $\mathscr{H} \equiv \bigoplus_{n=1}^{\infty} \sqrt{n!} \; \mathscr{H}_n$. This is endowed with the norm topology $\| \cdot \|_{\mathscr{H}}$ given by the following expression,

$$\| L \|_{\mathscr{H}} \equiv \left(\sum_{n=1}^{\infty} n! \, \| \hat{L}_n \|_{\mathscr{H}_n}^2 \right)^{1/2}.$$

Here L is independent of t, and it is given by $L \equiv \{\hat{L}_1(\cdot), \hat{L}_2(\cdot,\cdot), \cdots \hat{L}_n(\cdot,\cdot,\cdots s_n), \cdots\}$. It follows from the well known Wiener-Ito decomposition [4, Corollary 2.2.5, p44] that the Hilbert space $\mathscr{W}$ is isometrically isomorphic to the Hilbert space $\mathscr{H}$ briefly denoted by $\mathscr{W} \cong \mathscr{H}$. Hence we conclude that

$$L_2(I, \mathscr{W}) \cong L_2(I, \mathscr{H}).$$

This shows that the operator Φ (equivalently F), as defined by the expression (50), is an isometric isomorphism (topological homeomorphism). For details see [4] where an extension of Riesz-Fischer theorem on Wiener measure space [4, Theorem 2.2.3,p41] was proved to arrive at such conclusions. Thus, for every $\zeta \in L_2(I, \mathscr{W})$, there exists a unique $\mathscr{K} \in L_2(I, \mathscr{H})$ such that

$$\int_I \| \zeta(t) \|_{\mathscr{W}}^2 \, dt \; = \int_I \| \mathscr{K}(t) \|_{\mathscr{H}}^2 \, dt, \tag{53}$$

and conversely for every $\mathscr{K} \in L_2(I, \mathscr{H})$, there exists a unique (random process) $\zeta \in L_2(I, \mathscr{W})$ satisfying the above identity. Clearly, given the Wiener measure space, we can view the input-output model given by (50) as a linear operator from the Hilbert space $L_2(I, \mathscr{H})$ to the Hilbert space $L_2(I, \mathscr{W})$. Further, it is bijective and isometric. In view of the above observations, in order to identify $\mathscr{K}$ or equivalently the set of kernels $\{K_n\}$, we apply the Brownian motion to the input of the system (50) obtaining the output process $\{y(t), t \in I\} \in L_2(I, \mathscr{W})$. Let the process $\{y_n(t), t \in I\} \in L_2(I, \mathscr{W})$ denote the corresponding output of the generator of the $n-th$ order homogeneous chaos giving $y_n = H_n = \Phi_n(\hat{L})$ for any $\hat{L} \in L_2(I, \mathscr{H}_n)$. Taking the scalar product, it follows from the orthogonality property (51) that

$$\int_I \mathbf{E}^\mu\{y(t)y_n(t)\}dt = \int_I < y(t), y_n(t) >_{\mathscr{W}} dt$$

$$= n! \int_I < \hat{K}_n(t), \hat{L}(t) >_{\mathscr{H}_n} dt. \tag{54}$$

Now define the linear functional on $L_2(I, \mathscr{H}_n)$ by

$$\ell(\hat{L}) \equiv (1/n!) \int_I < y(t), y_n(t) >_{\mathscr{W}} dt = \int_I < \hat{K}_n(t), \hat{L}(t) >_{\mathscr{H}_n} dt. \tag{55}$$

Since the closed unit ball $B_1(L_2(I, \mathscr{H}_n)) \equiv \{\hat{L} \in L_2(I, \mathscr{H}_n) : \| \hat{L} \|_{L_2(I,\mathscr{H}_n)} \leq 1\}$ is weakly compact and the linear functional ℓ is weakly continuous, it attains its maximum on the boundary. Hilbert spaces are strictly convex (Banach) spaces because its closed unit ball is strictly convex. Thus ℓ attains its maximum at a unique point on the boundary. Thus maximizing this linear functional ℓ on the closed unit ball $B_1(L_2(I, \mathscr{H}_n))$ we obtain a unique $\hat{L}^o \in B_1(L_2(I, \mathscr{H}_n)$ given by

$$\hat{L}^o(t) = \frac{\hat{K}_n(t)}{\| \hat{K}_n \|_{L_2(I,\mathscr{H}_n)}}, t \in I, \tag{56}$$

belonging to the unit ball of $L_2(I, \mathscr{H}_n)$. From this we identify the n-th order kernel of the system as

$$\hat{K}_n(t) = \alpha_n \hat{L}^o(t), t \in I, \tag{57}$$

where $\alpha_n = \| \hat{K}_n \|_{L_2(I,\mathscr{H}_n)} = \ell(\hat{L}^o)$. This identifies the n-th order kernel for any $n \in N$. So we may denote $\hat{L}^o$ by $\hat{L}^o_n$. Thus we have proved the following result on the inverse problem.

Theorem 5.3. *Consider the input-output model (4) determined by an unknown kernel $\mathscr{K} \in L_2(I, \mathscr{H})$. By subjecting the system to standard Brownian motion at the input and scalar multiplying the output in the Hilbert space $L_2(I, \mathscr{W})$ with that of the generator of a homogeneous chaos, $H_n \equiv \Phi_n(\hat{L})$, of order n corresponding to any $\hat{L} \in L_2(I, \mathscr{H}_n)$ and maximizing the corresponding linear functional on the unit ball of $L_2(I, \mathscr{H}_n)$, one can uniquely identify $\mathscr{K}$ by $\mathscr{K} \equiv \{\hat{K}_n, n = 1, 2, 3 \cdots\}$ where*

$$\hat{K}_n = \ell(\hat{L}^o_n)\hat{L}^o_n \in L_2(I, \mathscr{H}_n), n \in N.$$

Remark 5.4. In order to carry out the above computation one is required to use the Monte Carlo simulation for the term on the left of the expression (54).

6. Extensions to infinite dimensional Hilbert spaces

The results presented in sections 2-5 can be extended to infinite dimensional Hilbert spaces using tensor products. We present only a brief outline. Let H be a real Hilbert space with norm denoted by $|\cdot|_H$ and let $H^{\otimes n}$ denote the (projective) tensor product of n-copies of H and $\mathscr{L}(H^{\otimes n}, H)$ the space of bounded linear operators from $H^{\otimes n}$ to H. For an operator $L \in \mathscr{L}(H^{\otimes n}, H)$, the operator norm is given by

$$\| L \|_{\mathscr{L}(H^{\otimes n}, H)} \equiv \sup\Big\{ |L(h_1 \otimes h_2 \otimes h_3 \cdots \otimes h_n)|_H, |h_i|_H \leq 1, i = 1, 2, \cdots, n \Big\}.$$

Let $\mathscr{L}_2(H^{\otimes n}, H) \subset \mathscr{L}(H^{\otimes n}, H)$ denote the space of Hilbert-Schmidt operators with the topology induced by the trace norm, $\sqrt{Tr(L^*L)} = \| L \|_{\mathscr{L}_2(H^{\otimes n}, H)}$. In order to avoid introducing new notations, we let $\mathscr{H}_n \equiv L_2(I^n, \mathscr{L}_2(H^{\otimes n}, H))$ denote the space of uniformly measurable (uniform operator topology) L_2 kernels defined on I^n and taking values in $\mathscr{L}_2(H^{\otimes n}, H)$. Using this notation the operator determined by the infinite Volterra series is given

$$F(x)(t) \equiv \sum_{n=1}^{\infty} F_n(x)(t), \tag{58}$$

where, for each $t \in I$ and $n \in N$,

$$F_n(x)(t) \equiv \int_{I_t^n} K_n(t; s_1, s_2, \cdots, s_n)\big(x(s_1) \otimes x(s_2) \otimes \cdots \otimes x(s_n)\big) ds_1 ds_2 \cdots ds_n.$$

For each $n \in N$, let $\hat{K}_n(t) \equiv K_n(t; \cdots) \in \mathscr{H}_n$, such that $\hat{K}_n \in L_2(I, \mathscr{H}_n)$. Consider the vector space (of infinite set of Kernels), again denoted by X,

$$X \equiv \Big\{ L \equiv \{L_1, L_2, L_3, \cdots, L_n \cdots\}, L_n \in \mathscr{H}_n \equiv \hat{L}_2(I^n, \mathscr{L}_2(H^{\otimes n}, H)), n \in N \Big\}. \tag{59}$$

Again, we furnish this with the norm topology

$$\| L \|_X \equiv \left(\sum_{n=1}^{\infty} n! \| L_n \|_{\mathscr{H}_n}^2 \right)^{1/2}.$$

In other words, X is given by the direct sum of Hilbert spaces, $X \equiv \bigoplus_{n=1}^{\infty} \sqrt{n}! \mathscr{H}_n$. We are concerned with the Hilbert space $L_2(I, X)$. Following similar steps as in section 2, it is easy to verify that, for each $\mathscr{K} \in L_2(I,X)$, $F : L_2(I,H) \longrightarrow L_2(I,H)$ satisfying the following inequality,

$$\| F(x) \|_{L_2(I,H)} \leq \| \mathscr{K} \|_{L_2(I,X)} \| x \|_{L_2(I,H)} \exp(\| x \|^2 /2).$$

Thus the results of section 2 also hold in the infinite dimensional Hilbert space H. Considering now the integral equation (22),

$$x = z + \lambda F(x),$$

on the Hilbert space $L_2(I,H)$, and defining the operator $x \longrightarrow G_\lambda(x) \equiv z + \lambda F(x)$ as in section 3, one can again verify that for each $z \in B_{\alpha r}(L_2(I,H))$, $\alpha \in (0,1)$, $0 < r < \infty$, and $|\lambda| \leq \frac{(1-\alpha)}{\|\mathscr{K}\|_{L_2(I,X)} \exp(r^2/2)}$, the operator

$$G_\lambda : B_r(L_2(I,H)) \longrightarrow B_r(L_2(I,H))$$

is a contraction. Thus all the results of section 3-5 hold also in the infinite dimensional Hilbert space H. We note that for the subsection 4.1, considering the question of existence of relaxed control, the set $\mathbf{S}$ is now given by $\mathbf{S} \equiv B_{\alpha r}(L_2(I,H))$. Endowed with the weak topology this is a compact Hausdorff space. Thus Theorem 4.4 also remains valid in the infinite dimensional Hilbert space.

Now let us consider the inverse problem. We need some minor modifications. Let U,H be two separable (real) Hilbert spaces with norms denoted by $|\cdot|_U, |\cdot|_H$ respectively and let $W \equiv \{W(t), t \geq 0\}$ be an U valued cylindrical Brownian motion with covariance being the identity operator in U. Let $U^{\otimes n}$ denote the tensor product of n-copies of U and $\mathscr{L}(U^{\otimes n},H)$ the space of bounded linear operators from $U^{\otimes n}$ to H. For an operator $L \in \mathscr{L}(U^{\otimes n},H)$, the operator norm is given by

$$\| L \|_{\mathscr{L}(U^{\otimes n},H)} \equiv \sup\left\{ |L(u_1 \otimes u_2 \otimes u_3 \cdots \otimes u_n)|_H, |u_i|_U \leq 1, i = 1,2,\cdots,n \right\}.$$

Let $\mathscr{L}_2(U^{\otimes n},H) \subset \mathscr{L}(U^{\otimes n},H)$ denote the space of Hilbert-Schmidt operators with the topology induced by the trace $Tr(L^*L) = \| L \|^2_{\mathscr{L}_2(U^{\otimes n},H)}$. Let $\hat{L}_2(I^n, \mathscr{L}_2(U^{\otimes n},H)) \equiv \mathscr{H}_n$ denote the Hilbert space of all symmetric kernels which are Lebesgue measurable in the uniform operator topology and square integrable on I^n in the norm topology of the space $\mathscr{L}_2(U^{\otimes n},H)$. For convenience of notation we have used the same notation as in section 5. In this infinite dimensional case also, the n-th order homogeneous Wiener chaos has the representation,

$$H_n(L) \equiv \int_{I_t^n} L(s_1,s_2,\cdots,s_n)dW(s_1)\otimes dW(s_2)\otimes\cdots,\otimes dW(s_n), \tag{60}$$

for $L\in\hat{L}_2(I^n,\mathscr{L}_2(U^{\otimes n},H))$. For each such L, the H-valued random variable $H_n(L)\in L_2(\Omega,H)\equiv \mathscr{W}$ with the norm given by $|H_n(L)|_{\mathscr{W}}$ where

$$\begin{aligned}|H_n(L)|^2_{\mathscr{W}} &= n!\int_{I^n} Tr(L^*L)(s_1,s_2,\cdots,s_n)ds_1ds_2\cdots ds_n\\ &= n!\int_{I^n} \parallel L(s_1,s_2,\cdots,s_n)\parallel^2_{\mathscr{L}_2(U^{\otimes n},H)} ds_1ds_2\cdots ds_n.\end{aligned} \tag{61}$$

Again, the sequence of homogeneous chaos $\{H_n, n\in N\}$ satisfy the orthogonality property as in (51). For $L\in\hat{L}_2(I^n,\mathscr{L}_2(U^{\otimes n},H))$ and $K\in\hat{L}_2(I^m,\mathscr{L}_2(U^{\otimes m},H))$, we have

$$\begin{aligned}&< H_n(L),H_m(K) >_{\mathscr{W}}\\ &\equiv \mathbf{E}^{\mu}\{\big(H_n(L),H_m(K)\big)_H\} = \begin{cases}0 & \text{if } n\neq m,\\ n!\int_{I^n} Tr(L^*K)ds_1ds_2\cdots ds_n & \text{if } n=m.\end{cases}\end{aligned} \tag{62}$$

Using the same character $\mathscr{H}$ for the space

$$\mathscr{H}\equiv\Big\{L\equiv\{L_1,L_2,L_3,\cdots,L_n\cdots\},L_n\in\hat{L}_2(I^n,\mathscr{L}_2(U^{\otimes n},H)),n\in N\Big\} \tag{63}$$

and introducing the norm topology as in section 2, we have

$$\parallel L\parallel_{\mathscr{H}}\equiv\left(\sum_{n=1}^{\infty} n!\parallel L_n\parallel^2_{\hat{L}_2(I^n,\mathscr{L}_2(U^{\otimes n},H))}\right)^{1/2}.$$

The space $\mathscr{H}$ equipped with the above norm topology is a Hilbert space and again $L_2(\Omega,\mu)\equiv \mathscr{W}\cong\mathscr{H}$ and hence $L_2(I,\mathscr{W})\cong L_2(I,\mathscr{H})$. Thus the results of section 5.1, in particular Theorem 5.3, also remain valid for the infinite dimensional case. For details see [4, S-2.6, p56] and [3].

Acknowledgements

This work was partially supported by the National Science and Engineering Research Council of Canada under grant No. A7101.

REFERENCES

[1] V. Volterra, Theory of Functionals and of Integral and Integro-differential Equations, Dover Publications, Inc. New York, (1959).

[2] N.U. Ahmed, An approximation theorem for continuous functions on $L_p(1 \leq p < \infty)$ spaces including representation, Information and Control, 30 (1976), 143-150.

[3] N.U. Ahmed, Functional analysis on Wiener measure space and evolution on fock spaces, Proc. Dyn. Sys. Appl. 1 (1994), 1-10.

[4] N.U. Ahmed, Generalized Functionals of Brownian Motion and Their Applications, (Nonlinear Functionals of Fundamental Stochastic Processes), World Scientific, New Jersey, London, Singapore, Beijing, Shanghai, Hong Kong, Taipei, Chennai, (2012).

[5] N. Dunford and J.T. Schwartz, Linear Operators, Part 1: General Theory, Interscience publishers Inc., New York, second printing, (1964).

[6] D. Sidorov, Integral Dynamical Models: Singularities, Signals and Control,World Scientific, New Jersey, London, Singapore, Beijing, Shanghai, Hong Kong, Taipei, Chennai, (2015).