

**UNIVERSITY OF OTTAWA**  
**Faculty of Pure and Applied Science**  
**Department of Electrical Engineering**  
**Ottawa, Canada**

**A THESIS:**

**FUNCTIONAL APPROACH TO ANALYSIS AND  
SYNTHESIS OF NONLINEAR SYSTEMS**

**by**

**N. U. Ahmed**

**Submitted in partial fulfilment of  
the requirements for the degree of  
Doctor of Philosophy in Electrical  
Engineering.**

## ABSTRACT

By the term functional (more precisely functional operator) is meant a transformation (not necessarily linear) of a linear vector space  $E_1$  into a linear vector space  $E_2$ .

The methods of analysis and synthesis of linear systems have a unifying basis. This is due to the fact that the set of all bounded linear operators acting within the same Banach space forms a ring. For such systems it is only necessary to choose a suitable element from the ring. In contrast nonlinear systems have no such unifying basis.

There are two possibilities: one possibility is to pick up a particular nonlinear device and perform its detailed analysis. The other possibility is to choose a class of nonlinear operators, as large as possible, and develop techniques for analysis and synthesis of systems belonging to the class.

In mathematical literature, a great deal of work has been done on system analysis based on the two basic classes of nonlinear operators; the class of Hammerstein operators and the class of Uryson operators.

It appears from engineering literature that the Volterra-Frechet functional series has occupied the interest of many workers in the field.

The author feels that there is a need of introducing a larger class of operators so that a larger class of systems may be considered in the problems of engineering analysis and synthesis. In this thesis two new classes of operators, which are natural extensions of the Hammerstein and the Uryson operators, have been introduced. Volterra-Frechet functional series is a subclass of these new classes of operators.

This thesis is largely devoted to the development of methods of analysis and synthesis of nonlinear systems belonging to the larger class.

Some interesting results have also been obtained for the class of Volterra-Frechet operators.

The thesis has been planned into two parts. The first part, consisting of chapter one and chapter two, considers mainly analysis and the second part, consisting of chapter three and chapter four, considers mainly synthesis.

In chapter 1, an algebra of nonlinear systems is developed after studying some of the fundamental properties of the basic nonlinear operators; which will be useful in the following chapters.

In chapter II, this algebra is used in the analysis of feedback control systems. The plant is assumed to be an element belonging to the algebra. Application of feedback gives rise to interesting nonlinear functional equations. The properties of the resolvents of these functional equations are studied in some detail by use of the principle of contraction mapping. It is also shown that solutions to nonlinear functional equations of Volterra-Frechet type (with variable upper limits of integration) can be obtained under much less restrictive conditions than usually demanded by the principle of contraction mapping. Stability of steady state solutions of these equations is briefly studied.

In chapter III, two types of problems of optimum synthesis are considered.

In the first type of problem the plant is assumed to be some undetermined element belonging to the algebra. The statistical properties of the input and the desired output are assumed to be known. An element from the algebra must be chosen such that the system performance is optimized in some sense.

In the second type of problem, the plant is assumed to be any functional operator possessing two Gateaux derivatives at each point

in some Banach space  $E$ . The problem considered is the design of an optimal control signal in  $E$ , from the knowledge of the desired output, the characteristics of the plant and a suitable performance criterion.

Analysis and synthesis of nonlinear systems become substantially simplified if a set of orthogonal functionals, complete and closed in the class of functionals under consideration, can be constructed. This problem is considered in chapter IV. A set of orthogonal functionals is constructed on the measure space  $(\Omega, B, \mu)$  [where  $\Omega$  may be either a separable Hilbert space  $H$  or the space  $C$ ,  $B$  is the  $\sigma$ -ring of measurable cylinder sets in  $\Omega$  and  $\mu$  is the Gaussian measure on the  $\sigma$ -ring]. This set of orthogonal functionals is complete in the class  $L^2(\Omega, B, \mu)$  and therefore any functional  $g(\cdot) \in L^2(\Omega, B, \mu)$  has a Fourier development. This leads to a systematic method of analysis and synthesis of a class of nonlinear systems.

The techniques developed in the first three chapters are applicable to a large class of nonlinear systems. The practical disadvantage is the complexity. In this respect the technique presented in chapter IV is superior but suffers from many limitations which are discussed in the context.

Most of the propositions appearing in the context are proved by the author. Those propositions which have been borrowed are indicated by the name of their original authors.

Useful comments on the results, their limitations, and possible extensions are made at the end of each chapter.

ACKNOWLEDGEMENT

I am very grateful to my advisor, Professor G. S. Glinski, for his generous encouragement, understanding and guidance.

I would also like to express my sincere thanks to Professor V. I. Linis, Professor J. L. Howland and Professor D. A. De Salvo of the Department of Mathematics, and Professor J. A. Brzozowski of the Department of Electrical Engineering for their helpful suggestions and valuable criticism.

Thanks are also due to the Atomic Energy Commission of Pakistan and Colombo Plan of Canada who sponsored the project.

## TABLE OF CONTENTS

	PAGE
<b>CHAPTER 1</b>	<b>ALGEBRA OF OPERATORS</b>
1.1:	Introduction. 1
1.2:	Linear Operators. 1
1.3:	Nonlinear Operators. 13
<b>CHAPTER II</b>	<b>NONLINEAR FUNCTIONAL EQUATIONS IN THE ANALYSIS OF FEEDBACK CONTROL SYSTEMS.</b>
2.1:	Examples of Nonlinear Systems of Type $A_U \quad A_H \quad A_V$
2.1-1:	Introduction 30
2.1-2:	Some Examples of Nonlinear Physical Systems. 30
2.2:	Nonlinear Functional Equations and their Resolvents. 44
2.2-1:	The Principle of Contraction Mapping. 44
2.2-2:	Resolvent of Nonlinear Operators and their Properties. 47
2.2-3:	Existence of Local and Nonlocal Solutions. 54
2.2-4:	Nonlinear Causal Systems. 57
<b>CHAPTER III</b>	<b>OPTIMUM SYNTHESIS OF NONLINEAR SYSTEMS.</b>
3.1:	Introduction 74
3.2:	Integral Equations in the Optimum Synthesis of a Class of Nonlinear Time-variant Systems. 75
3.3:	Formulation of an Optimal Control Problem and its Solution. 90
3.4:	Discussions. 106

## CHAPTER IV

THE PROBLEM OF CANONICAL  
REPRESENTATION OF NONLINEAR  
SYSTEMS BY ORTHOGONAL  
FUNCTIONALS.

4.1:	Introduction	112
	4.1-1: Stochastic Processes.	114
	4.1-2: Measurable Functionals.	117
4.2:	Orthogonal Functionals and Their Properties.	
	4.2-1: Construction of Orthogonal Functionals on the Measure Space $(\Omega_I, \mathcal{B}_I, \mu_B)$	120
	4.2-2: Properties of the Orthogonal Functionals.	126
4.3:	Remarks.	134
	References.	141

**CHAPTER I**  
**ALGEBRA OF OPERATORS**

## 1.1 INTRODUCTION

In the design of a system it is sometimes necessary to combine a set of subsystems whose characteristics may be known to the designer. The properties of the overall systems depend very much on how the subsystems are combined. Therefore, in order to be able to predict the behavior of the overall system from the knowledge of the properties of the subsystems, it is very essential to study the algebraic structure of the operators.

In section 1.2, we study the algebraic properties of linear operators of Fredholm type. Linear operators of Volterra type, being a special case of those of Fredholm type, will possess similar properties even under less restrictive conditions.

In section 1.3 we study the algebraic structure of a few basic nonlinear integral operators in which we are interested in the following chapters.

## 1.2 LINEAR OPERATORS

We consider linear operators acting on  $L^2$  spaces.

We will denote by  $A$  the class of all linear operators of Fredholm type acting within  $L^2$  spaces. There are two reasons for the choice of this special space as the domain of this class of operators:

(i) This is the most commonly encountered space in the study of engineering systems.

(ii) Mathematically, the properties of this class of spaces are very wellknown and can be easily treated.

Definition 1.1

By a linear operator of Fredholm type we mean the following integral operator,

$$y = Kx \quad (1.1)$$

$$\text{ie } y(t) = \int_I K(t, \tau) x(\tau) d\tau$$

$$\forall t \in I.$$

where  $I$  is any fixed interval on the real line.

We will assume that the Kernel  $K(t, \tau) \in L^2(I^2)$ , where  $I^2 = I \times I$ , and that  $x(t) \in D(K)$  the domain of the operator  $K$  and that it is an  $L^2$  function on  $I$ . The Lebesgue measure of the set  $I$  may be finite or infinite. The set  $A(K_1, L_1, R_1, M_1, \dots)$  will be used to denote the class of all Fredholm operators, and the elements  $K_1, L_1, \dots$  etc will be used to denote the Kernels corresponding to these operators which are Lebesgue square integrable on  $I \times I$ .

Theorem 1.1

Let  $K_1 \in A$  and let  $x \in L^2(I)$ , then the function  $z$  defined by,

$$z = K_1 x \text{ (a.e on } I) \text{ also belongs to } L^2(I).$$

Proof: Since  $K_1 \in L^2(I^2)$ , it is a measurable function of  $(t, \tau) \in (I \times I)$ . Hence by Fubini's theorem [1, 2] for almost all  $t \in I$ ,  $K_1(t, \tau)$  is a measurable function of  $\tau$  on  $I$ . Moreover for almost all  $t \in I$ ,  $K_1(t, \tau) \in L^2(I)$  as a function of  $\tau$ . Since  $x \in L^2(I)$  as a function of  $\tau$ , it follows from Scharz inequality that for almost all  $t \in I$ , the integral  $z(t) = \int_I K_1(t, \tau) x(\tau) d\tau$  exists and that whenever

$z$  does exist,

$$|z(t)| \leq \left( \int_I |K_1(t, \tau)|^2 d\tau \int_I |x(\tau)|^2 d\tau \right)^{1/2} \quad 1.2$$

By use of Fubini's theorem on measureability of plane sets it can be proved that  $z(t)$  is a measurable function of  $t$  on  $I$ . Since  $z$  is measurable on  $I$ , it follows on squaring both sides of the inequality 1.2 and integrating over  $I$ , that  $z \in L^2(I)$  and

$$\|z\| \leq \|K_1\| \|x\|.$$

This theorem shows that  $K_1$  carries  $L^2$  functions into  $L^2$  function.

It is clear that the properties of the integral operator (1.1) depend on those of the  $L^2$  Kernel  $K$ . In the following we study the algebraic structure of the class  $A$  of Fredholm operators.

We now define the following operations on the set  $A$  :

(i) " $\Psi_1$ ":  $A \times A \rightarrow A$  is a binary operation on  $A$ , which assigns to any two elements  $K_1, L_1 \in A$  an element  $M_1$  which is also in  $A$  and is defined by  $\Psi_1(K_1, L_1) = K_1 + L_1$

(ii) " $\simeq$ " this is an equivalence relation on  $A$ .

If for each  $x \in L^2(I)$  and for any two elements  $K_1$  and  $L_1 \in A$ ,  $K_1 x = L_1 x$  holds a.e on  $I$ , then  $K_1$  will be said to be equivalent to  $L_1$  and this will be written as  $K_1 \simeq L_1$ .

(iii) " $\Psi_2$ ":  $F \times A \rightarrow A$  is defined by  $\Psi_2(a, K_1) = a \cdot K_1$  for every  $a \in F$  (the field of real or complex numbers) and for every  $K_1 \in A$ .

(iv) " $\Psi_3$ ":  $A \times A \rightarrow A$  is also a binary operation on  $A$ , which assigns to any two elements of  $A$  an element which also belongs to  $A$ ,

and is defined by,  $\Psi_3(K_1, L_1) = K_1 \odot L_1$ . That is, for every  $x \in L^2(I)$ ,  $(K_1 \odot L_1)(x) = K_1(L_1 x)$ .

(i) The set A (that is the class of integral operators as defined, in this section) is closed under the operation  $\Psi_1$ .

The proof of this fact follows directly from Minkowski's inequality [1].

Let  $x \in L^2(I)$  and let  $K_1, L_1 \in A$  then by Minkowski's inequality we have,

$$\begin{aligned} \|M_1 x\| &= \|(K_1 + L_1)x\| \leq \|K_1 + L_1\| \|x\| \\ &\leq (\|K_1\| + \|L_1\|) \|x\| \end{aligned}$$

Hence  $M_1 \triangleq K_1 + L_1$  also carries every  $L^2$  function into an  $L^2$  function and  $\|M_1\| \leq \|K_1\| + \|L_1\|$ . 1.3

Thus  $M_1$  also belongs to A, that is the set A is closed with respect to addition and  $M_1 x = K_1 x + L_1 x$  a.e on I.

(ii) " $\simeq$ ", the equivalence relation on A as defined must be consistent, in the sense that if any two elements  $K_1, L_1 \in A$  satisfy the relation

$$K_1 x = L_1 x \quad \forall x \in L^2(I), \text{ then } K_1 \simeq L_1.$$

Thus we have the following theorem:

Theorem 1.2

Any two elements,  $K_1, L_1$  belonging to the class A are equivalent if, and only if for all  $x \in L^2(I)$   $K_1 x = L_1 x$  a.e on I.

Proof: The "only if" part follows from the definition. We prove the "if" part. Let us define  $R_1 = K_1 - L_1$ . Then from the hypothesis

$$y(t) = \int_I R_1(t, \tau) x(\tau) d\tau = 0 \text{ a.e on I.} \quad 1.4$$

$\forall x \in L^2(I)$ .

since  $x \in L^2(I)$ , and otherwise arbitrary we can define  $x$  on  $(I \setminus I_0)$ , where the Lebesgue measure of the set  $I_0 = 0$ , by the function  $\overline{R_1(t, \tau)}$  with  $t \in (I \setminus I_0)$ . It is clear that  $\overline{R_1(t, \tau)} \in L^2(I \times I)$ . Hence

$$\begin{aligned} y(t) &= \int_I R_1(t, \tau) \overline{R_1(t, \tau)} d\tau \\ &= \int_I |R_1(t, \tau)|^2 d\tau = 0 \text{ a.e on } I. \end{aligned} \quad 1.5$$

Since  $R_1 \in L^2(I \times I)$ , on integrating both sides of 1.5 with respect to  $t$  on  $(I \setminus I_0)$ , we will have,

$$\int_{I \setminus I_0} y(t) dt = \int_{I \setminus I_0} \int_I |R_1(t, \tau)|^2 d\tau dt = 0$$

Thus,  $R_1(t, \tau) = 0$  a.e on  $(I \times I)$  which implies that  $K_1 \simeq L_1$ . Q.E.D. It follows from this theorem that if  $K_1$  and  $L_1$  are any two operators belonging to the class  $A$  then for any  $x \in L^2(I)$

$$(K_1 - L_1)x = K_1x - L_1x. \quad 1.6$$

Further it is easily seen that the relation " $\simeq$ " is an equivalence relation on  $A$ , satisfying the relations (a)-(c).

- (a),  $K_1 \simeq K_1$  (reflexivity)
- (b) If  $K_1 \simeq L_1$  then  $L_1 \simeq K_1$  (symmetric)
- (c) If  $K_1 \simeq L_1$ ,  $L_1 \simeq R_1$  then  $K_1 \simeq R_1$  (transitivity)

(iii) It is easily seen that the set  $A$  is closed with respect to ordinary multiplication by any real or complex number.

Let  $a$  be any real or complex number and let  $K_1 \in A$ , then  $R_1 \triangleq a.K_1$  also belongs to  $A$  for  $\|R_1\| = |a| \|K_1\|$ .

(iv) The set  $A$  is closed under the operation  $\Upsilon_3$ .

Let  $K_1$  and  $L_1 \in A$ , then  $\Upsilon_3: A \times A \rightarrow A$ .

For any  $x \in L^2(I)$

$$M_1 x = (K_1 \odot L_1)x \triangleq \int_I \left( \int_I K_1(t, s) L_1(s, \tau) ds \right) x(\tau) d\tau$$

where

$$M_1(t, \tau) \triangleq \int_I K_1(t, s) L_1(s, \tau) ds \quad 1.7$$

Theorem 1.3

The operation " $\odot$ " as defined above assigns to each pair of elements  $K_1$  and  $L_1$  belonging to  $A$  an element, say,  $M_1$  which also belongs to  $A$ .

Proof: Let  $x$  be any element in  $L^2(I)$ , then

$$M_1 x(t) = \int_I M_1(t, \tau) x(\tau) d\tau \quad 1.8$$

$$= \iint_{I \times I} K_1(t, s) L_1(s, \tau) x(\tau) ds d\tau \quad 1.9$$

By Fubini's theorem [2], we have

$$M_1 x(t) = \int_I K_1(t, s) L_1 x(s) ds$$

$$(K_1 \odot L_1)x = K_1(L_1 x) \quad 1.10$$

So that

$$M_1 x = K_1(L_1 x) \quad \forall x \in L^2(I). \quad 1.11$$

By the previous theorem, any other operator belonging to the class  $A$  satisfying this property is equivalent to  $M_1$ , so that

$$M_1 = K_1 \odot L_1 \quad \text{and} \quad \|M_1\| \leq \|K_1\| \|L_1\|. \quad \text{Q.E.D.}$$

Remark:

In general (time varying system), the product by composition of the elements belonging to the class  $A$  does not satisfy the commutative law, since the operator  $(L_1 \odot K_1) = \int_I L_1(t, s) K_1(s, \tau) ds$  is generally different from that defined by equality 1.7.

This is illustrated by the figure (1.1) below.

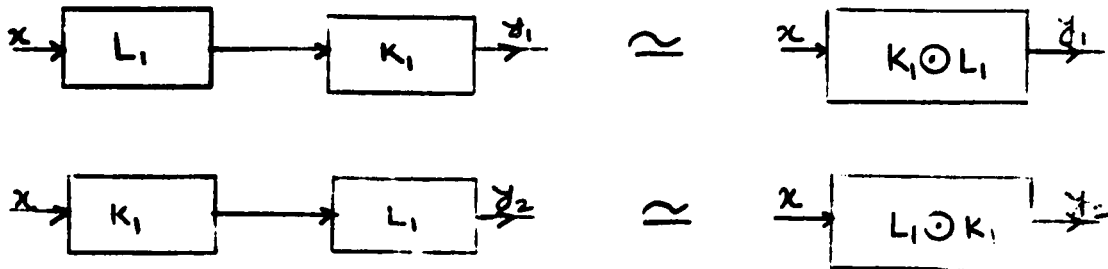


Fig. 1.1: Composition of two linear operators.

The product by composition does satisfy the associative law since for all  $K_1, L_1, R_1 \in A$  we have,

$$\begin{aligned} (K_1 \odot (L_1 \odot R_1)) x &= K_1 ((L_1 \odot R_1) x) && \text{by 1.10.} \\ &= K_1 (L_1 (R_1 x)) && \text{by 1.10} \end{aligned}$$

Similarly

$$\begin{aligned} ((K_1 \odot L_1) \odot R_1) x &= (K_1 \odot L_1) (R_1 x) && \text{by 1.10} \\ &= K_1 (L_1 (R_1 x)) && \text{by 1.10} \end{aligned}$$

This proves that

$$K_1 \odot (L_1 \odot R_1) = (K_1 \odot L_1) \odot R_1 \quad 1.11$$

and that both are equal to  $K_1 \odot L_1 \odot R_1$ .

This is illustrated in the figure below.

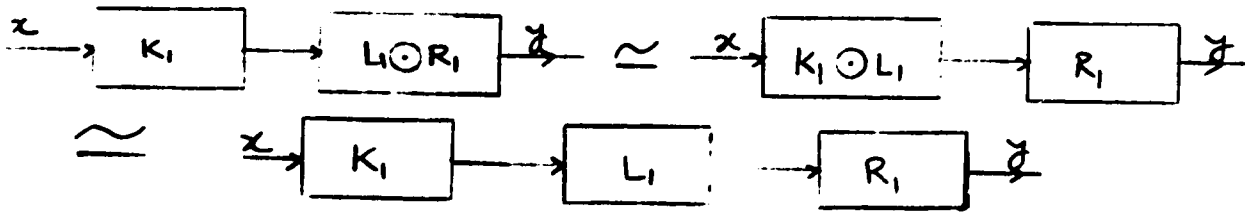


Fig.: 1.2: Associative Law.

The product by composition satisfies both the distributive laws (left and right), since for left distribution we have,

$$\begin{aligned}
 (K_1 \odot (L_1 + R_1)) x &= K_1 ((L_1 + R_1) x) \\
 &= K_1 (L_1 x + R_1 x) \\
 &= (K_1 \odot L_1) x + (K_1 \odot R_1) x \\
 &= (K_1 \odot L_1 + K_1 \odot R_1) x. \quad 1.12
 \end{aligned}$$

The first equality follows from 1.10, the second follows from the equation 1.6, the third equality follows from the linearity of the class A and 1.10, and again the fourth equality follows from 1.6.

By theorem 1.2 and equation 1.12, we have,

$$K_1 \odot (L_1 + R_1) = K_1 \odot L_1 + K_1 \odot R_1 \quad 1.13$$

Similarly it can be shown that the product by composition " $\odot$ " satisfies the right distribution, ie.

$$(L_1 + R_1) \odot K_1 = L_1 \odot K_1 + R_1 \odot K_1 \quad 1.14$$

The equalities 1.13 and 1.14 are illustrated by the following diagram. Fig. 1.3.

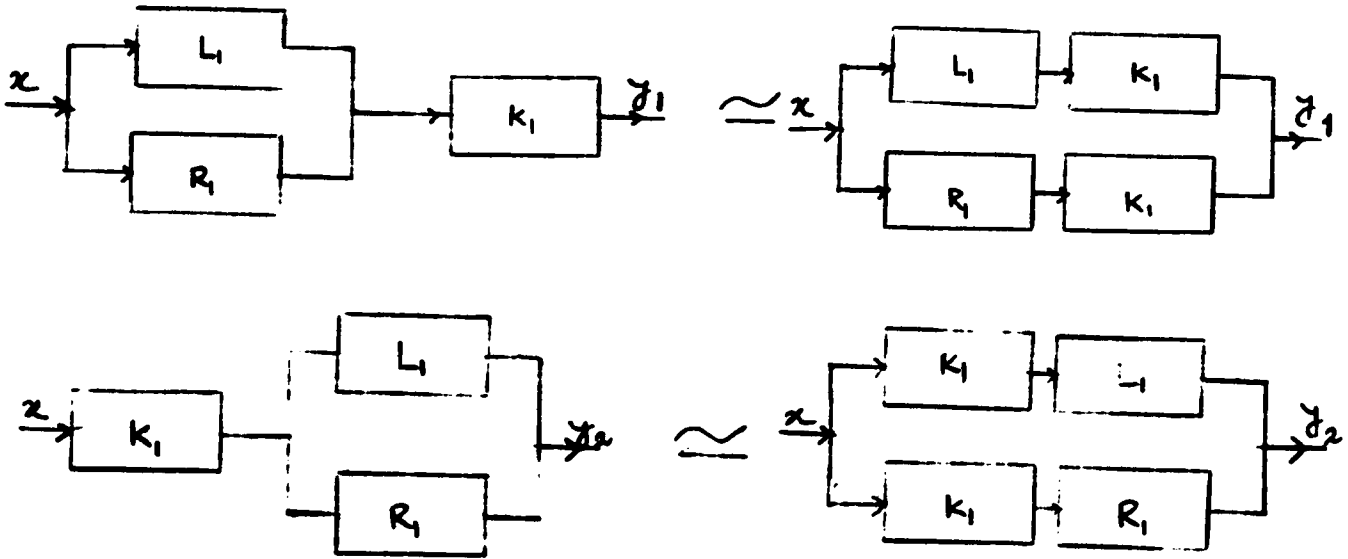


Fig. 1.3: Left and Right Distributive Laws.

From the above considerations we see that the class  $A$  of linear operators satisfies the following relations with respect to addition and ordinary multiplication,

For all  $K_1, L_1, R_1 \in A$  and  $a, b \in F$  (the field of reals or the complex numbers) we have,

- L1:  $K_1 + L_1 = L_1 + K_1$
- L2:  $K_1 + (L_1 + R_1) = (K_1 + L_1) + R_1$
- L3:  $K_1 + L_1 = K_1 + R_1 \implies L_1 \approx R_1$  1.15
- L4:  $a(K_1 + L_1) = aK_1 + aL_1$
- L5:  $(a+b) K_1 = aK_1 + bK_1$
- L6:  $a(b K_1) = (ab) K_1$
- L7:  $1 \cdot K_1 = K_1$

These are precisely the postulates of a linear vector space. Thus the set  $A$  of all linear operators on  $L^2$ ; with addition, multiplication by a scalar and the equivalence relations, constitutes a linear vector space over the complex field. That is we now

have the linear space  $(A, \Psi_1, \Psi_2, \simeq)$ , which for abbreviation will be denoted simply by  $A$ .

From the equivalence relation on the space  $A$  we have,

$L_1 \simeq R_1 \implies L_1 + (-1)R_1 \simeq \theta$ , the null vector is  $\theta(t, \tau) = 0$  a. e on  $(I \times I)$ . It can be easily proved that this null Kernel is unique in  $A$ , and that to every element  $K_1 \in A$  there corresponds a unique element  $(-1)K_1 = -K_1$  such that  $K_1 - K_1 \simeq \theta$ .

In addition to satisfying the postulates of linear vector space over the complex field, the class  $A$  of linear operators satisfies also a ring structure, with respect to addition "+" and product by composition  $\odot$ .

Definition: 1.2

A set  $R$  is a ring, if it is an additive Abelian group, and is also a semigroup under the operation " $\odot$ " and further it satisfies the left and right distributions under addition.

We have already seen that, the following relations are satisfied  $\forall K_1, L_1, R_1 \in A$ :

- R1: Additive closure,  $K_1 + L_1 \in A$
- R2: Additive Associativity,  $K_1 + (L_1 + R_1) = (K_1 + L_1) + R_1$
- R3: Additive identity,  $K_1 + \theta = \theta + K_1 = K_1$
- R4: Additive inverse,  $K_1 + (-K_1) = (-K_1) + K_1 = \theta$
- R5: Additive commutativity,  $K_1 + L_1 = L_1 + K_1$  1.16
- R6: Closure under the operation " $\odot$ "  

$$K_1 \odot L_1 \in A, \forall K_1, L_1 \in A$$
- R7: Associativity  

$$K_1 \odot (L_1 \odot R_1) = (K_1 \odot L_1) \odot R_1$$

R8: Left and Right distribution,

$$(L) \quad K_1 \odot (L_1 + R_1) = K_1 \odot L_1 + K_1 \odot R_1$$

$$(R) \quad (L_1 + R_1) \odot K_1 = L_1 \odot K_1 + R_1 \odot K_1$$

Thus according to the definition 1.2, the class of linear operators  $A$  constitutes a ring (in general noncommutative) with respect to the two important operations  $\Psi_1$  and  $\Psi_3$ . This ring may be denoted by  $R_A = (A, \Psi_1, \Psi_3, "-" \theta)$  where  $A$  is a non empty set - the set of linear operators as defined before,  $\Psi_1$  and  $\Psi_3$  are binary operations on  $A$ , "-" is a unary operation on  $A$ , and  $\theta$  is the zero of the ring  $R_A$ .

Definition: 1.3

$R$  is a ring with zero divisors if for any two elements  $a$  and  $b$  in  $R$  satisfying  $a \odot b = \theta$  does not necessarily imply, that, if  $a \neq \theta$ ,  $b = \theta$ .

It can be easily verified that the ring  $R_A$  is a ring with zero divisors, i.e the cancellation law does not hold in this ring. For example let  $I = [0 T]$ ,  $T < \infty$ , and let  $K_1(t, \tau) = \tau$  and  $L_1(t, \tau) = (t - \frac{2T}{3})\tau$ , then

$$\begin{aligned} K_1 \odot L_1 &= \int_0^T K_1(ts) L_1(s\tau) ds \\ &= \int_0^T s(s - \frac{2T}{3}) \tau ds = \theta \end{aligned}$$

whereas  $K_1$  and  $L_1$  belonging to  $L^2(I \times I)$  are both different from  $\theta$ . so far we have not defined any identity element in this ring. In order to carry out algebraic operations it is convenient to introduce an identity operator  $e$  defined by

$$ex = x \quad \forall x \in L^2(I)$$

That is,  $e$  carries each  $L^2$  function into itself. If the measure of the set  $I$  is infinite then there exists no  $L^2$  Kernel corresponding to this identity operator  $e$ , so it does not belong to the class  $A$ . Similarly for each real or complex number  $a$  we may define the operation,

$$(ae)x = a(ex) = ax \quad \forall x \in L^2(I).$$

According to this definition for each  $K_1 \in A$

$$eK_1 = K_1e = K_1 \tag{1.18}$$

$$\text{ie } e(K_1x) = K_1(ex) = K_1x \quad \forall x \in L^2(I).$$

An important property of the class  $A$  of operators is that, the  $n$ th iterated Kernel corresponding to any Kernel  $K_1 \in A$ , also belongs to  $A$ . Let  $n$  be any positive integer, then we define the  $n$ th iterated Kernel  $K_1^{(n)}$  by

$$K_1^{(n)} = \int_I K_1^{(n-m)}(t, s) K_1^{(m)}(s, \tau) ds \quad \text{for all } 1 \leq m \leq n$$

$$K_1^{(n)}(t, \tau) = \int_I \dots \int_I K_1(t, s_1) K_1(s_1, s_2) \dots K_1(s_{n-1}, \tau) ds_1 ds_{n-1} \tag{1.19}$$

By Schwarz inequality it can be easily proved that

$$\|K_1^{(n)}\| \leq (\|K_1\|)^n \quad \text{for all } n \geq 1. \tag{1.20}$$

Thus  $K_1^{(n)} \in A$ .

A simple application of this algebra is found in the formal inversion of the equation,

$$x = y + \lambda Kx \tag{A_1}$$

It is easily verified that if  $|\lambda| < (\|K\|)^{-1}$  then  $(e-\lambda K)^{-1}$  is defined and  $x$  is given by,

$$\begin{aligned} x &= (e-\lambda K)^{-1} y \\ &= y + \lambda \sum_{n=1}^{\infty} \lambda^{n-1} K^{(n)} y \\ &= y + \lambda \Gamma_{\lambda} y \end{aligned} \tag{A_2}$$

where  $K^{(n)}$  is the  $n$ th iterated Kernel corresponding to the Kernel  $K$  and  $\Gamma_{\lambda}$  is the resolvent of the Kernel  $K$  corresponding to the scalar  $\lambda$ . Substituting  $x$  from  $A_2$  in,  $(e-\lambda K)x$  it is easily verified that  $A_1$  is satisfied.

**Remark:**

In this section we have presented only some salient properties of linear integral operators useful to the engineer. The theory of linear operators [5] is so highly developed that even a brief presentation of its most interesting results may take hundreds of pages. This section is only intended to help in distinguishing the basic differences between the algebra of linear and nonlinear operators.

### 1.3: NONLINEAR OPERATORS

#### CLASSIFICATION

The classification of nonlinear operators is no doubt a difficult task. It is however very desirable to construct a few operators that represent a wide class of nonlinear systems. The discussion of these selected but general nonlinear operators should provide answers to problems that may arise in the study of any concrete problem.

The following classes of operators appear to be sufficiently general to cover a wide variety of problems arising in the study of non-linear systems:

(i) Uryson operator  $A_U$ : The class of Uryson operators is described by the following functional series,

$$A_U x(t) = \sum_{s=0}^N \int_{I^s} \cdots \int U_s [t; \tau_1, \dots, \tau_s; x(\tau_1) \dots x(\tau_s)] d\tau_1 \dots d\tau_s, t \in I \quad 1.21$$

defined for almost all  $t \in I$ . The sequence of functions  $\{U_s\}$  may be assumed Lebesgue measurable in  $t$  and  $\tau_i$  ( $i = 1, 2, \dots, s$ ) and continuous in  $x(\tau_i)$ , ( $i = 1, 2, \dots, s$ ).  $N$  may be finite or infinite.  $I^s$  is defined as the Cartesian product of  $s$  copies of the set  $I$ , where  $I$  may be a linear or an  $n$ -dimensional Lebesgue measurable set of finite or infinite measure.

(ii) Hammerstein operator  $A_H$ : The class of Hammerstein operators is described by a series of the following form,

$$A_H x(t) = \sum_{s=0}^N \int_{I^s} \cdots \int K_s (t; \tau_1, \dots, \tau_s) g(\tau_1, x(\tau_1)) \dots g(\tau_s, x(\tau_s)) d\tau_1 \dots d\tau_s \quad t \in I \quad 1.22$$

where the Kernels  $\{K_s\}$  are assumed to be Lebesgue measurable in  $t$  and  $\tau_i$ , ( $i = 1, \dots, s$ ) and the function  $g(\cdot, x(\cdot))$  may be assumed to be measurable in  $t$  and continuous in  $x$ . Again  $N$  may be finite or infinite and  $I$  as before.

(iii) Lyapunov operator  $A_{L_s}$ : The class of Lyapunov operators is described by the following series,

$$A_{L_s} x = \sum_{\alpha_0 \dots \alpha_s; \beta_0 \dots \beta_s} L_{\alpha_0 \dots \alpha_s; \beta_0 \dots \beta_s} (x; y) \quad 1.23$$

where,  $L_{\alpha_0 \dots \alpha_s; \beta_0 \dots \beta_s} (x; y)$  is defined by the following integral

power form of degree  $m_1 = \sum_{i=0}^s \alpha_i$  in  $x$  and of degree  $m_2 = \sum_{i=0}^s \beta_i$  in  $y$ .

$$L_{\alpha_0 \dots \alpha_s \beta_0 \dots \beta_s}(x; y) = \int_{I^s} \dots \int K_{\alpha_0 \dots \alpha_s; \beta_0 \dots \beta_s}(t, \tau_1 \dots \tau_s) x^{\alpha_0}(t), \\ \dots x^{\alpha_s}(\tau_s) \dots y^{\beta_0}(t) \dots y^{\beta_s}(\tau_s) d\tau_1 \dots d\tau_s.$$

$$t \in I. \quad 1.24$$

$\alpha_i$  ( $i=0, \dots, s$ ) and  $\beta_i$  ( $i=0, \dots, s$ ) are any two sets of non-negative integers,  $y$  is any fixed element in some Banach space  $E$  and  $x$  is its variable element. The set  $I^s$  is, as defined previously.

(iv) Volterra-Frechet operator  $A_V$ : The class of Volterra-Frechet operators is described by the following series,

$$A_V x = \sum_{s=0}^N \int_{I^s} \dots \int K_s(t; \tau_1 \dots \tau_s) x(\tau_1) \dots x(\tau_s) d\tau_1 \dots d\tau_s$$

$$t \in I. \quad 1.25$$

Here also, the Kernels  $\{K_s\}$  may be assumed to be Lebesgue measurable in  $t$  and  $\tau_i$  ( $i = 1, \dots, s$ ). The integer  $N$  may be finite or infinite. The set  $I^s$  is as defined previously.

It is noticed that the class  $A_V$  is contained in  $A_H$  which in turn is contained in  $A_U$ . We propose the first two classes of operators as extensions of the more familiar Uryson and Hammerstein operators [3] as noted below.

$$A_U^* x = \int_I K(t, \tau; x(\tau)) d\tau \quad 1.26$$

and,

$$A_H^* x = \int_I K(t, \tau) g(\tau, x(\tau)) d\tau. \quad 1.27$$

## CONTINUITY, BOUNDEDNESS AND COMPACTNESS OF OPERATORS.

In the study of nonlinear systems with the use of the operators defined above, it is necessary to have complete knowledge of such properties as continuity, boundedness and compactness of these operators on a given function space. Therefore the following definitions are important.

### Definition 1.4: Continuity.

An operator  $A$  acting from a Banach space  $E_1$  into another Banach space  $E_2$  is said to be continuous at the point  $x_0 \in E_1$  if it transforms every sequence of elements  $\{x_n\} \in E_1$  which converges strongly in  $E_1$  to an element  $x_0 \in E_1$ , into a sequence of elements  $\{Ax_n\} \in E_2$  which also converges strongly in  $E_2$  to an element  $Ax_0 \in E_2$ .

An operator is said to be continuous on the subset  $M \subset E_1$  if it is continuous at each point of the set  $M$ .

### Definition 1.5: Boundedness.

An operator is said to be bounded if it transforms every bounded (in the sense of norm) subset of  $E_1$  into a set which is bounded (in the sense of norm) in  $E_2$ .

### Remark

In contrast with linear operators, one cannot conclude for nonlinear operators that a bounded operator is continuous or conversely that a continuous operator is bounded. If, however, the space  $E_1$  is of finite dimension then continuity implies boundedness.

### Definition 1.6: Compactness.

An operator  $A$  is said to be compact if it transforms every bounded set into a compact set.

**Definition 1.7: Complete Continuity.**

An operator is completely continuous if it is continuous and compact.

**Definition 1.8: Strong Continuity.**

An operator is said to be strongly continuous if it transforms every weakly convergent sequence into a strongly convergent one.

**Definition 1.9: Weak Continuity**

An operator is said to be weakly continuous if it transforms every weakly convergent sequence into a weakly convergent one.

**Definition 1.10: Lipschitz Condition**

An operator  $A$  acting from the Banach space  $E_1$  into the Banach space  $E_2$  is said to satisfy on the subset  $M \subset E_1$  the Lipschitz condition with a constant  $\alpha$  if

$$\|Ax_1 - Ax_2\| < \alpha \|x_1 - x_2\|, \quad \forall x_1, x_2 \in M. \quad 1.28$$

In general  $\alpha$  may be a function of  $M$ .

**Remark:** operators which satisfy the Lipschitz condition are obviously bounded and continuous.

We will mainly consider the development of an algebra of operators of type  $A_H$  and  $A_v$ . Before we can do so, we must study the continuity and boundedness properties of these operators considered acting on some Banach space  $E$ . We will prove results for  $L^2(I)$  spaces which immediately apply to the space  $C(I)$  with the appropriate norm used. The results can be generalized to  $L^p(I)$  spaces also. The reasons for restricting our discussion to the two classes of operators  $A_H$  and  $A_v$  are:

(i) These operators are sufficiently general to cover a wide class of nonlinear physical systems.

(ii) The study of such properties as continuity, boundedness and compactness of these operators is comparatively easier than that of Uryson operator.

CONTINUITY AND BOUNDEDNESS OF  $A_H$  AND  $A_v$

We will prove the continuity and boundedness of the operator  $A_H$  and obtain the corresponding result for  $A_v$  as corollary. To prove the continuity and boundedness of the operator  $A_H$  we will make use of the following theorem [3]

**Theorem 1.4:**

If the operator  $G$  defined by  $Gx = g(t, x(t))$ , maps every function  $x \in L^q$  into a function in  $L^p$ , ( $\frac{1}{p} + \frac{1}{q} = 1$ ;  $p, q > 1$ ) then  $G$  is continuous and bounded and the following inequality holds,

$$|g(t, x(t))| \leq |z(t)| + \beta |x(t)|^{q/p} \tag{1.29}$$

where  $0 < \beta < \infty$  and  $z \in L^p$ .

We now present the following theorem on the boundedness of the operators  $A_H$  and  $A_v$ .

**Theorem 1.5:**

If  $N < \infty$  and the operator  $G$  satisfies conditions of theorem 1.4 and the set of Kernels  $\{K_s\}_{s \geq 0}^N \in L^q(I^{s+1})$  and are measurable in all the variables on  $I^{s+1}$ , then  $A_H$  acts within  $L^q$  and is bounded.

**Proof:**

Let us rewrite the operator  $A_H$  as

$$A_H x(t) = \sum_{s=1}^N A_{H_s} x(t), \tag{1.30a}$$

with  $t \in I$ , and  $x \in L^q(I)$ ,

where

$$A_{H_s} x(t) = \int_{I^s} \cdots \int K_s(t, \tau_1, \dots, \tau_s) g(\tau_1; x(\tau_1)) \cdots g(\tau_s; x(\tau_s)) d\tau_1 \cdots d\tau_s, \\ t \in I. \quad 1.30b$$

Let us prove that  $A_{H_s}$  acts within  $L^q$ . It is clear that by Holder's inequality we can write,

$$|A_{H_s} x(t)| \leq \left( \int_{I^s} \cdots \int |K_s(t; \tau_1 \dots \tau_s)|^q d\tau_1 \cdots d\tau_s \right)^{1/q} \left( \int_I |g(\tau; x(\tau))|^p d\tau \right)^{s/p}. \\ \text{a.e in } I. \quad 1.31$$

By Fubini's theorem [2] the function on the left of the above inequality is defined almost everywhere on  $I$  and is measurable in  $t$ . Now since by hypothesis the operator  $G$  satisfies the inequality 1.29, we have,

$$\left( \int_I |g(t, x(t))|^p dt \right)^{1/p} \leq \left\{ \|z\|_p + \beta (\|x\|_q)^{q/p} \right\}. \quad 1.32$$

According to the hypothesis of the present theorem, since  $K_s \in L^q(I^{s+1})$ , the function  $A_{H_s} x(t) \in L^q(I)$ . Thus by raising either side of the inequality 1.31 to the power  $q$  and integrating the resulting function over  $I$  and finally performing the  $q$ th in root of the resulting quantity we obtain,

$$\|A_{H_s} x\|_q \leq \|K_s\|_q \left[ \|z\|_p + \beta (\|x\|_q)^{q/p} \right]^s \quad 1.33 \\ (s = 0, 1, \dots, N).$$

This inequality ensures that each of the terms in 1.30a belongs to  $L^q$ . Hence by Minkowski's inequality the left hand side of 1.30a also belongs to  $L^q$  for each finite  $N$ .

Thus

$$\begin{aligned} \|A_H x\|_q &\leq \sum_{s=0}^N \|A_{H_s} x\|_q \\ &\leq \sum_{s=0}^N \|K_s\|_q (\|z\|_p + \beta(\|x\|_q)^{q/p})^s, \end{aligned} \quad 1.34$$

and  $A_H$  maps  $L^q$  into  $L^q$  ( $q > 1$ ) and therefore bounded.

In particular if  $K_s \in L^2(I^{s+1})$  and  $G: L^2 \rightarrow L^2$ , then the inequality 1.34 reduces to,

$$\|A_H x\|_2 \leq \sum_{s=0}^N \|K_s\|_2 (\|z\|_2 + \beta\|x\|_2)^s. \quad 1.35$$

If  $g(t; x(t)) = \theta$  a.e. on  $I$  for  $x(\cdot) = \theta$  a.e. on  $I$ , then 1.34 and 1.35 become respectively,

$$\|A_H x\|_q \leq \sum_{s=0}^N \|K_s\|_q (\beta^s (\|x\|_q)^{\frac{sq}{p}}). \quad 1.36$$

and

$$\|A_H x\|_2 \leq \sum_{s=0}^N \|K_s\|_2 (\beta^s \|x\|_2^s). \quad 1.37$$

This completes the proof that for  $N$  finite the operator  $A_H$  acts within  $L^q$  ( $q > 1$ ) and is bounded. Q.E.D.

Corollary:

In the case of Volterra-Frechet operator  $A_v$ , if the Kernels  $K_s \in L^p(I^{s+1})$  and  $x \in L^q(I)$  then by a similar procedure as above, it can be shown that  $A_v$  maps  $L^q$  into  $L^p$  ( $q, p > 1$ ) and the inequality corresponding to 1.34 is given by,

$$\|A_v x\|_p \leq \sum_{s=0}^N \|K_s\|_p (\|x\|_q)^s \quad 1.38$$

which in the case of  $L^2$  spaces becomes,

$$\|A_v x\|_2 \leq \sum_{s=0}^N \|K_s\|_2 (\|x\|_2)^s. \quad 1.39$$

Throughout our discussion we will consider our operators to be acting on  $L^2$  spaces.

In the case of strong nonlinearities, it may be necessary to let  $N = \infty$ . In this situation the boundedness of the operators of either of the types  $A_H$  or  $A_V$  does not follow simply from the facts as mentioned above. The question of convergence of the resulting infinite series must be settled now.

If  $R$ , defined as

$$R = \frac{1}{\lim \sqrt{\|K_n\|}} \tag{1.40}$$

is greater than zero, then the operator  $A_H$  is bounded if in addition to satisfying all the conditions of theorem 1.5, it satisfies the following additional conditions:

- (i)  $\|x\| \leq \frac{R - z}{\beta}$
  - (ii)  $R > |z|$
- 1.41

The Volterra-Frechet operator  $A_V$  is bounded in a sphere  $S_R$  of radius  $R$  as given by 1.40, where  $S_R \subset L^2$ .

As mentioned in the remark following definition 1.10, boundedness of nonlinear operators does not necessarily imply continuity or conversely. So we will prove the continuity of these operators in the following theorem.

**Theorem 1.6:**

Let the operator  $G$  satisfy a Lipschitz condition with constant  $c > 0$  in addition to satisfying the inequality 1.29 corresponding to  $L^2$  space, and let  $R$ , as defined by 1.40, lie in the interval  $(0, \infty)$ . Under these assumptions, each of the operators  $A_H$  and  $A_V$  satisfies a Lipschitz condition in a sphere  $S_\rho$  of

radius  $\rho$  in the space  $L^2$  with constants,

$$\alpha_H(\rho) = \sum_{s=0}^{\infty} s c \|K_s\| (\|z\| + \beta\rho)^{s-1} < \infty, \quad 1.42$$

and 
$$\alpha_V(\rho) = \sum_{s=0}^{\infty} s \|K_s\| \rho^{s-1} < \infty, \quad 1.43$$

respectively. In the case of operator of type  $A_H^{\rho} = \frac{R-z}{\beta}$  and in the case of  $A_V$ ,  $\rho = R$ , and both are completely continuous and bounded in  $S_{\rho}$ .

Proof:

Let  $x_1$  and  $x_2$  be any two elements belonging to  $L^2(I)$  such that  $\|x_1\|, \|x_2\| \leq \rho$  then for any finite  $s$ ,

$$\|A_{H_s} x_1 - A_{H_s} x_2\| \leq s \|K_s\| \left( \sup_{x \in S_{\rho}} \|g(\tau; x(\cdot))\| \right)^{s-1} \|g(\tau; x_1(\cdot)) - g(\tau; x_2(\cdot))\| \quad 1.44$$

By hypothesis on the operator  $G$ , we have

$$\begin{aligned} \sup_{x \in S_{\rho}} \|Gx\| &= \sup_{x \in S_{\rho}} \|g(\cdot, x(\cdot))\| \\ &\leq \|z\| + \beta\rho \end{aligned} \quad 1.45$$

By substituting 1.45 into 1.44 and using the Lipschitz condition for  $G$ , we obtain,

$$\|A_{H_s} x_1 - A_{H_s} x_2\| \leq s c \|K_s\| (\|z\| + \beta\rho)^{s-1} \|x_1 - x_2\| \quad 1.46$$

Hence,

$$\|A_H x_1 - A_H x_2\| \leq \left\{ \lim_{N \rightarrow \infty} \sum_{s=0}^N s c \|K_s\| (\|z\| + \beta\rho)^{s-1} \right\} \|x_1 - x_2\| \quad 1.47$$

If  $\rho$  is chosen as assumed in the theorem the series inside the bracket converges to a limit  $\alpha_H(\rho)$  and 1.47 becomes,

$$\|A_H x_1 - A_H x_2\| \leq \alpha_H(\rho) \|x_1 - x_2\| \quad \forall x_1, x_2 \in S_{\rho} \quad 1.48$$

with  $\rho = \frac{R - \|z\|}{\beta}$ .

In the case of operators of type  $A_v$  it can be similarly shown that,

$$\|A_v x_1 - A_v x_2\| \leq \alpha_v(\rho) \|x_1 - x_2\|, \quad \forall x_1, x_2 \in S_\rho, \quad 1.49$$

with  $\rho = R$ .

The values of  $R$  in the two cases need not be related at all.

The continuity of the operators  $A_H$  and  $A_v$  follows from the inequalities 1.48 and 1.49 respectively. Q. E. D.

Remarks:

It is important to notice that in case  $N$  is finite both  $\alpha_H(\rho)$  and  $\alpha_v(\rho)$  tends to infinity only when  $\rho \rightarrow \infty$ , but in case  $N$  is infinite

$$\alpha_H(\rho) = \infty, \quad \forall \rho > \frac{R - |z|}{\beta}$$

and  $\alpha_v(\rho) = \infty, \quad \forall \rho > R$ .

Thus in the former case, the domains  $D(A_H)$  and  $D(A_v)$  of the operators  $A_H$  and  $A_v$  respectively, are the whole of  $L^2$  space and in the later case, these are only some suitable closed and bounded subsets of the space  $L^2$ .

THE ALGEBRA:

Let us denote by  $B(H, G, F, \dots)$  a nonempty set of continuous and compact operators of either of the class  $A_H$  or  $A_v$  acting on suitable closed and bounded subsets of the space  $L^2$ .  $H, G, F, \dots$  etc. are the elements of  $B$ . Those algebraic properties of the set  $B$ , which are most important for engineering work, are stated here without making any attempt at a rigorous proof.

It may be shown that  $B$  is an additive Abelian group.

Let us define a map  $\Psi_1: B \times B \rightarrow B$  by  $\Psi_1(H, G) = H + G$  ie  $\Psi_1$  assigns to any two elements of  $B$  a new element which also belongs to  $B$ . The closure property with respect to addition holds if, and

only if, the domain of the resulting operator  $R$  is defined as the intersection of the domains of  $H$  and  $G$ ; that is  $R$  is defined on a subset  $D(R)$  of  $L^2$  such that

$$D(R) \subseteq D(H) \cap D(G)$$

1.50

and  $Rx = (H + G)x = Hx + Gx \quad \forall x \in D(R).$

It is the second equality that requires proof.

$A_1$ : Additive closure:

$$\text{Let } R_1 = H + G \text{ with } H, G \in B$$

Then  $R_1 \in B$  provided  $D(R_1) = D(H) \cap D(G).$

In case  $H$ , and  $G$  both belong to  $B_N \subset B$ , where  $B_N$  is defined as the set of all operators of type  $A_H$  and  $A_V$  consisting of finite series (i.e.  $N < \infty$ ) only, then  $D(R_1)$  is the whole of  $L^2$  space.

$A_2$ : Associativity:

$$R_2 = H + (G + F) = (H + G) + F$$

and  $R_2 \in B$  provided  $D(R_2) \subseteq D(H) \cap D(G) \cap D(F).$

$A_3$ :  $\forall H \in B, -H \in B$  and  $D(H) = D(-H)$

because  $\forall x \in D(H), \|Hx\| = \|-Hx\|$ .

$A_4$ : It can be proved that there exists a unique null element  $\phi \in B$  such that for all  $x \in L^2$   $\phi x = \theta$ , where  $\theta$  is the null of the  $L^2$  space. By  $A_3$ ,  $H + (-H) \in B$ , hence  $\forall x \in D(H), Hx + (-Hx) = Hx - Hx = \theta \in L^2$ . Since,  $D(H) \subset L^2$ ,  $\phi \in B$ , and  $H + \phi = \phi + H = H \quad \forall H \in B.$

This shows that  $B$  is an additive Abelian group.

We can define another operation on the set  $B$ . Let us denote by  $F$  the field of real or complex numbers and let us define by  $\Psi_2$  a map such that  $\Psi_2: F \times B \rightarrow B$  by  $\Psi_2(a, H) = a.H, \forall H \in B$  and  $\forall a \in F$ . The set  $B$  is closed with respect to the operation  $\Psi_2$ , because  $\forall x \in D(H), (a.H)x = a.Hx$  and  $\|(a.H)x\| = |a| \|Hx\|$ .

Therefore  $a.H \in B \quad \forall H \in B$  and  $\forall a, \in F$  and  $D(a.H) = D(H)$ .

It is easily verified that with respect to this operation, the following relations hold true.

$$A_5: \quad a.(G+F) = a.G + a.F \quad \forall G, F \in B \text{ and } \forall a \in F \text{ with } D(a.G+a.F) = D(G) \cap D(F)$$

$$A_6: \quad (a+b).G = a.G + b.G, \quad \forall G \in B \text{ and } \forall a, b \in F, \text{ with } D((a+b).G) = D(G)$$

$$A_7: \quad a.(b.G) = (a.b).G \quad \forall G \in B \text{ and } \forall a, b \in F \text{ with } D(a.b.G) = D(G).$$

$$A_8: \quad 1.H = H, \quad \forall H \in B \text{ and } 1 \in F.$$

We can also define an equivalence relation " $\simeq$ " on the set  $B$  such that

$H+G = H+F \Rightarrow G \simeq F \quad \forall H, G, F \in B$  and that  $D(G) = D(F)$  and  $\forall x \in D(G), Gx = Fx$ . This equality must be understood in the sense of the norm of the space  $L^2$ .

For example in the case of Volterra-Frechet operator two operators are equivalent if, and only if, each of the corresponding kernels are pairwise equivalent in the  $L^2$  sense.

Let  $G$  and  $F$  both belong to  $B$  and

$$\text{Let } Gx \triangleq \sum_s \int_{I^s} \dots \int K_s(t, \tau_1, \dots, \tau_s) x(\tau_1) \dots x(\tau_s) d\tau_1 \dots d\tau_s \quad 1.51$$

with  $t \in I$  and  $x \in D(G)$

$$\text{and } Fx \triangleq \sum_s \int_{I^s} \dots \int L_s(\tau_1, \dots, \tau_s) x(\tau_1) \dots x(\tau_s) d\tau_1 \dots d\tau_s \quad 1.52$$

with  $t \in I$  and  $x \in D(F)$

then for  $G$  to be equivalent to  $F$  it is necessary and sufficient that  $D(G) = D(F)$  and that

$$\sum_{s=0}^{\infty} \|K_s - L_s\|^2 = 0 \quad 1.53$$

which implies that  $K_s$  be equal to  $L_s$  almost

everywhere on  $I^{s+1}$  for all  $s = (0, 1, 2, \dots)$ . In the case of operators of type  $A_H$ , the equivalence relation is not so straight-forward.

In case,  $K_s = L_s$  a.e on  $I^{s+1}$  for all  $s = 0, 1, 2, \dots$  then  $g$  must be equal to  $f$  for all  $x \in L^2$  and for almost all  $\tau \in I$ . Conversely if  $g = f$ ,  $\forall x \in L^2$  and almost all  $\tau \in I$  then for the equivalence of the corresponding operators it is necessary and sufficient that  $K_s = L_s$  a.e on  $I^{s+1}$  for each  $s$ . However it is important to note that an operator  $G$  could be equivalent to an operator  $F$ , both belonging to the class  $A_H$ , without actually any of the Kernels of  $G$  being equivalent to any of the corresponding Kernels of  $F$ . This is simply due to the presence of the corresponding zero memory operators  $g$  and  $f$ .

All the relations  $A_1 - A_8$  resulting from the two operations  $\Psi_1$  and  $\Psi_2$  defined on  $B$  are precisely the postulates of a linear vector space in which an equivalence relation " $\simeq$ " is also defined. Thus  $B(H, G, F, \dots, \Psi_1, \Psi_2, \simeq, \emptyset)$  is a linear vector space whose elements are the set of all continuous bounded nonlinear operators defined on suitable subsets of the space  $L^2$ .

Another important operation that can be defined on the set  $B$  is "the product by composition". Let  $H$  and  $G \in B$ , and let  $\Psi_3 : B \times B \rightarrow B$  be defined as

$$\Psi_3(H, G) = H \odot G = R_3 \tag{1.54}$$

so that,

$$R_3 x = (H \odot G)x = H(Gx), \quad \forall x \in D(R_3)$$

$$\text{where } D(R_3) = D(G) \cap \{x, x \in D(G) : Gx \in D(H)\}. \tag{1.55}$$

Thus,  $B$  is closed with respect to this operation provided  $D(R_3)$  is chosen as defined by 1.55. We note the following properties of the set  $B$  with respect to the operation  $\Psi_3$ :

$$A_9: \forall H, G \in B, H \circ G \in B$$

$$\text{with } D(H \circ G) = D(G) \cap \{x, x \in D(G) : Gx \in D(H)\} \quad 1.56$$

$$A_{10} \quad \forall H, G, F \in B$$

$$H \circ (G \circ F) = (H \circ G) \circ F = H \circ G \circ F$$

$$\text{with } D(H \circ G \circ F) = D(F) \cap \{x, x \in D(F) : Fx \in D(G)\} \cap \{x, x \in D(F) : (G \circ F)x \in D(H)\}. \quad 1.57$$

In general the set  $B$  does not satisfy the left distributivity property but it does satisfy the right distributivity.

$$A_{11}: H \circ (G+F) \neq H \circ G + H \circ F. \quad 1.58$$

The equality holds only when  $H$  is linear.

$$A_{12}: (G+F) \circ H = G \circ H + F \circ H$$

$$\text{with } D((G+F) \circ H) = D(G \circ H) \cap D(F \circ H) \quad 1.59$$

$$= D(H) \cap \{x, x \in D(H) : Hx \in D(G)\} \cap \{x, x \in D(H) : Hx \in D(F)\}$$

with respect to the operation  $\Psi_3$ ,  $B$  is a semigroup and it would be a Ring if  $A_{11}$  were true. We note that the set of all bounded linear operators defined on any Banach space forms a Ring, which makes the study of linear operators comparatively simpler.

With respect to the operation  $\Psi_3$ , we may define an identity element  $I$  by,

$$H \circ I = I \circ H = H \quad 1.60$$

so that

$$(H \circ I)x = (I \circ H)x = Hx \quad \forall x \in D(H).$$

with this new element included,  $B$  is an algebra closed under the operations of addition, multiplication by scalars, and product by composition. This algebra may now be denoted by  $B(H, F, G, \dots$

$\Upsilon_1, \Upsilon_2, \Upsilon_3, \simeq, \phi, I)$ , where,  $H, F, G$  etc. are the elements of  $B$ ,  $\Upsilon_1, \Upsilon_2, \Upsilon_3$  are the operations as defined, " $\simeq$ " is a relation, and  $\phi$  and  $I$  are the two special elements of  $B$ .

In this algebra cancellation law does not hold since

$$H \circ G \simeq H \circ F \not\Rightarrow G \simeq F. \quad 1.61$$

As we have seen, the negation of this implication is true even in the case of linear operators. Therefore, the element  $I$  may not be unique. Also there may not exist inverses for the elements of  $B$  since  $y = Hx$ , may not have a solution for  $x$ , for an arbitrary  $y \in L^2$ .

An important point associated with the operation  $\Upsilon_3$ , defined on  $B$ , is the domain of the combined operator  $H \circ G$  for all  $H$  and  $G$  in  $B$ . The entries in the following table indicate the domain of the operator  $H \circ G$ . It will be clear that  $D(H \circ G)$  is a function of the domains of the individual operators and the nature of the leading operator  $G$ .

Table of  $D(H \circ G)$ .

$G$	$D(G) = D(H)$	$D(G) \supset D(H)$	$D(G) \subset D(H)$	$D(G) \cap D(H) = \phi$
Reducing $G D(G) \subset D(G)$	$D(H \circ G) = D(H)$	$D(G) \supset D(H \circ G) \supseteq D(H)$	$D(H \circ G) = D(G)$	-
Expanding $G D(G) \supset D(G)$	$D(H \circ G) \subset D(H)$	$D(H \circ G) \subset D(H)$	$D(H \circ G) \subseteq D(G)$	-
$G D(G) = D(G)$	$D(H \circ G) = D(H)$	$D(H \circ G) = D(H)$	$D(H \circ G) = D(G)$	-

The situation represented by the last column of the table may arise in case the zeroth order terms in our operators are present and are not identical.

**Remark:**

In this chapter we have studied some important properties of the two classes of operators  $A_H$  and  $A_V$ . These will be useful in the following chapter. It remains to study the continuity and boundedness of the operators of type  $A_U$ . This is a more difficult problem and is not attempted here.

Proof of complete continuity of the first order Uryson operator in the spaces  $C$  and  $L^P$  are given in the book of Krasnosel'skii [3]. From the complexity of his proof for this smaller class, it is quite conceivable that the proof of complete continuity of the larger class will be too complex.

Some simpler properties, such as functional derivatives of these operators which will be required in the problems of synthesis in chapter III, are studied there.

The algebra developed in this chapter is meant to be useful in system engineering where a complete system is built out of many subsystems.

**CHAPTER II**  
**NONLINEAR FUNCTIONAL EQUATIONS IN THE ANALYSIS**  
**OF FEEDBACK CONTROL SYSTEMS**

## 2.1-1: INTRODUCTION

In the preceding chapter we presented the algebra of nonlinear integral operators of general character. It is interesting to note that many physical problems of apparently different nature can be conveniently formulated in terms of functional operators presented. The main interest in this chapter lies in the study of nonlinear feedback control systems where the plant is described by anyone of those operators or their suitable combination. Before considering this problem, some examples of this class of operators will be presented.

In the analysis of feedback control systems to be considered in section 2.2 of this chapter, the operator  $A$  may be assumed to be an element of the algebra  $B$  or its subalgebra  $B_N$ .

## 2.1-2: SOME EXAMPLES OF NONLINEAR PHYSICAL SYSTEMS.

(a) examples of first order Hammerstein operator.

(i) As an example of a first order operator of Hammerstein type equation 1.27, let us consider the following variable coefficient nonlinear differential equation.

$$y^{(n)}(t) + \sum_{k=1}^n a_K(t) y^{(n-k)}(t) + f(t, y(t) \dots y^{(m)}(t)) = x(t)$$

2.1

with,  $t \in I: [a, \beta]$  and  $y^{n-r}(a) \triangleq Y_{n-r}$

( $r = 1, 2, \dots, n$ ), and  $m \leq n$ .

It will be shown that we can reduce this equation to its corresponding nonlinear volterra integral equation, where the operator involved turns out to be of Hammerstein type.

Let us write,

$$y^{(n)}(t) = U_n(t) \tag{2.2}$$

Then it can be easily shown, that  $y^{(n-k)}(t)$  is given by,

$$y^{(n-k)}(t) = \sum_{r=1}^k Y_{n-r} \frac{(t-a)^{k-r}}{(k-r)!} + \int_a^t \frac{(t-s)^{k-1}}{(k-1)!} U_n(s) ds \tag{2.3}$$

for all  $(k = 1, 2, \dots, n)$

Particularly  $y(t)$  is given by,

$$y(t) = \sum_{r=1}^n Y_{n-r} \frac{(t-a)^{n-r}}{(n-r)!} + \int_a^t \frac{(t-s)^{n-1}}{(n-1)!} U_n(s) ds \tag{2.4}$$

Substituting (2.3) in (2.1) we obtain,

$$U_n(t) = x(t) - f(t, y(t), y^{(1)}(t), \dots, y^{(m)}(t)) - \tag{2.5}$$

$$\sum_{k=1}^n \sum_{r=1}^n a_k(t) Y_{n-r} \frac{(t-a)^{k-r}}{(k-r)!} - \int_a^t \sum_{k=1}^n a_k(t) \frac{(t-s)^{k-1}}{(k-1)!} U_n(s) ds$$

Let us define,

$$K(t, s) = \sum_{k=1}^n (-1) a_k(t) \frac{(t-s)^{k-1}}{(k-1)!} \tag{2.6}$$

and

$$g(t) = (x(t) - \sum_{k=1}^n \sum_{r=1}^k a_k(t) Y_{n-r} \frac{(t-a)^{k-r}}{(k-r)!})$$

Substituting these quantities in equation 2.5 we obtain,

$$U_n(t) = [g(t) - f(t, y(t), y^{(1)}(t), \dots, y^{(m)}(t))] + \int_a^t K(t, s) U_n(s) ds \tag{2.7}$$

If the function  $f$ , were absent from the system 2.1 we would obtain a linear integral equation of Volterra type with  $U_n(t)$  as the unknown function. It is obvious that if the coefficients  $a_k(t)$  ( $k=1, \dots, n$ ) are continuous then  $K(t, s)$  is also continuous. By the usual method of solution of linear Volterra integral equations [7, 8], the solution of equation 2.7 is given by,

$$U_n(t) = [g(t) - f(t, y, y^{(1)} \dots y^{(m)})] + \int_a^t R(t, s) [g(s) - f(s, y(s) \dots y^{(m)}(s))] ds \quad 2.8$$

where,

$$R(t, s) = \sum_{n=1}^{\infty} K^{(n)}(t, s)$$

and 
$$K^{(n)}(t, s) = \int_s^t K^{(n-1)}(t, \tau) K(\tau, s) d\tau \quad 2.9$$

$$K^{(1)}(t, s) = K(t, s)$$

(N. B. It can be shown [8] that if the Kernel  $K$  is continuous then the above series converges absolutely and uniformly everywhere on the triangle  $a \leq s \leq t \leq \beta$ .) By substituting 2.8 in equation 2.4, we obtain,

$$y(t) = \sum_{r=1}^n Y_{n-r} \frac{(t-a)^{n-r}}{(n-r)!} + \int_a^t \left[ \frac{(t-\tau)^{n-1}}{(n-1)!} + \int_{\tau}^t \frac{(t-s)^{n-1}}{(n-1)!} R(s, \tau) ds \right] [g(\tau) - f(\tau, y(\tau) \dots y^{(m)}(\tau))] d\tau \quad 2.10$$

which reduces to the form,

$$y(t) = L_0(t) - \int_a^t L(t, s) f(s, y(s), y^{(1)}(s); \dots y^{(m)}(s)) ds \quad 2.11$$

where

$$L_0(t) = \sum_{r=1}^n Y_{n-r} \frac{(t-a)^{n-r}}{(n-r)!} + \int_a^t L(t, s) g(s) ds \quad 2.12$$

and

$$L(t, s) = \frac{(t-s)^{n-1}}{(n-1)!} + \int_s^t \frac{(t-\xi)^{n-1}}{(n-1)!} R(\xi, s) d\xi . \quad 2.13$$

Equation 2.11 is a nonlinear integro differential equation of Volterra type. If the nonlinear term were absent equation 2.12, written as

$$y(t) = I_0(t) + \int_a^t L(t, s) x(s) ds \quad 2.14$$

where,

$$I_0(t) = \sum_{r=1}^n Y_{n-r} \frac{(t-a)^{n-r}}{(n-r)!} - \sum_{k=1}^n \sum_{r=1}^k Y_{n-r} L(t, s) a_k(s) \frac{(s-a)^{k-r}}{(k-r)!} ds \quad 2.15$$

provides an explicit input output relation. If all the initial conditions are assumed to be zero then  $I_0(t) \equiv 0$  and in this case equation 2.14 becomes

$$y(t) = \int_a^t L(t, s) x(s) ds. \quad 2.16$$

This is just the transient part of the solution of the corresponding linear differential equation. It can be shown easily that  $L(t, s)$  satisfies the following differential equation,

$$L^{(n)}(t, s) + \sum_{k=1}^n a_k(t) L^{(n-k)}(t, s) = \delta(t-s) \quad 2.17$$

where  $\delta(t)$  is the familiar impulse function. In the absence of the nonlinear part the same equation 2.1 can be shown to lead to a linear Fredholm integral equation of the second kind in case the terminal conditions,  $y(\beta), \dots, y^{(n-1)}(\beta)$  are also specified.

In any case the integral equation corresponding to the differential equation 2.1 appears in the following more general form,

$$y(t) = z(t) + \lambda \int_I K(t, s) f(s, y(s), \dots, y^{(m)}(s)) ds \quad 2.18$$

where  $\lambda$  is a parameter, and  $I$  is either  $[a, t]$  or  $[a, \beta]$  depending on whether it is an initial value or boundary value problem. In many physical situations, the function  $f$  is independent of the derivatives of  $y$ , in which case equation 2.18 becomes

$$y(t) = z(t) + \lambda \int_I K(t, s) f(s, y(s)) ds. \quad 2.19$$

Remark:

(i) Looking at the equation 2.7, it appears logical to consider the possibility of differential equations with infinite order. If we let  $n \rightarrow \infty$   $U_n(t)$  may converge to a limit function  $U(t)$ , which will represent the solution of a differential equation of infinite order. This seems possible if the two series (equation 2.6) converge to certain limit functions. In any case the form of the final equations 2.18 and 2.19 remain the same.

(ii) If in equation 2.1, the leading coefficient is taken to be a function  $a_0(t)$  then equation 2.7 becomes,

$$a_0(t) U_n(t) = [g(t) - f(t, y, y^{(1)} \dots y^{(m)})] + \int_a^t K(t, s) U_n(s) ds. \quad 2.20$$

equations 2.6 remaining the same. If  $a_0(t) \neq 0$  on,  $a \leq t \leq \beta$  then there is no change whatsoever, however if  $a_0(t)$  assumes zero values at certain points in  $[a, \beta]$  then equation 2.20 represents a singular integro differential equation.

(iii) Another important point in favour of integral equations is that the interval  $I$ , may be finite or infinite with  $a \rightarrow -\infty$ , but this does not seem to have any meaning in the case of differential equations unless  $y^{(n)}(-\infty) = 0$  for all  $n$ .

(iv) In the case of differential equations, of the form discussed here, the solution  $y(t)$  has to be necessarily continuous or at most piecewise continuous, but in the case of integral equations the solution may be allowed to be unbounded on sets of measure zero, and may only be demanded to belong to any of the  $L^p$  spaces.

(v) The solutions of equations of the form 2.19 when the upper limit is variable (i.e. of Volterra type) can be discussed under much less restrictive conditions on the functions  $K(t, s)$ ,  $f(t, u)$  and  $z(t)$ . It is sufficient to suppose that

$$\left. \begin{aligned} & \left| K(t, s) f(s, y_1(s)) - K(t, s) f(s, y_2(s)) \right| \leq h(t, s) |y_1 - y_2| \\ \text{and } & \left| \int_a^t K(t, s) f(s, z(s)) ds \right| \leq v(t), \end{aligned} \right\} 2.21$$

where  $h(t, s)$ , and  $v(t)$  are any two  $L^2$  functions. If these conditions are satisfied, then it can be proved that with  $I = [a, t]$ , equation 2.19 possesses an unique  $L^2$  solution, for all  $\lambda \in \mathbb{R}^1$ . The solutions of equations of the form 2.18 can also be studied under similar conditions.

Examples of nonlinear oscillatory systems.

(ii) An important example of a nonlinear differential equation that can be easily reduced to either of the forms 2.18 or 2.19 is given below.

Let us consider the following electric circuit with a nonlinear inductance (Fig: 2.1, 2.2), in which  $\phi$  and  $i$  are approximately

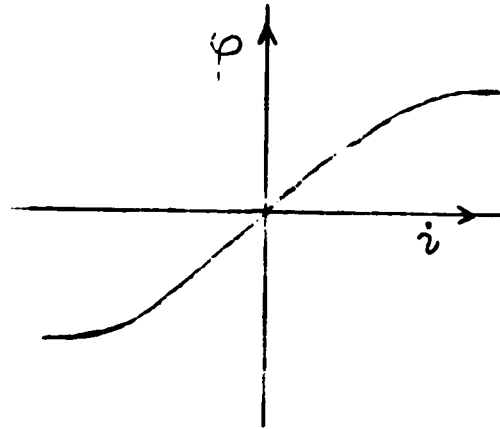
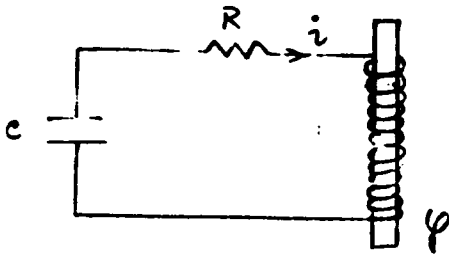


Fig. 2.1: Nonlinear R, L, C Circuit:

Fig.2.2: Inductance Characteristics:

related by the following equation,

$$i = a\varphi + b\varphi^3, \quad a, b > 0 \tag{2.22}$$

The relevant differential equation is,

$$\frac{d\varphi}{dt} + R(a\varphi + b\varphi^3) + \int_0^t \left( \frac{a\varphi + b\varphi^3}{c} \right) dt = 0 \tag{2.23}$$

which on differentiating once reduces to

$$\frac{d^2\varphi}{dt^2} + \omega^2\varphi = \lambda f(\varphi, \dot{\varphi}) \tag{2.24}$$

$$\text{where, } \omega^2 = \frac{a}{c}, \quad \text{and } f(\varphi, \dot{\varphi}) = (R(a+3b\varphi^2) + \frac{b}{c}\varphi^3) \tag{2.25}$$

and  $\lambda = 1$ .

Equation 2.24 can be easily reduced to the following integro differential equation,

$$\varphi(t) = z(t) + \int_0^t L(t,s) f(\varphi(s), \dot{\varphi}(s)) ds, \tag{2.26}$$

where

$$z(t) = \varphi(0) + \varphi^{(1)}(0)t + \omega \int_0^t \sin\omega(t-\xi) [\varphi(0) + \varphi^{(1)}(0)\xi] d\xi \tag{2.27}$$

and

$$L(t,s) = (t-s) + \int_s^t \omega \cdot \sin\omega(t-\xi) \cdot (\xi-s) d\xi.$$

Apparently, equation 2.26 looks more complex than equation 2.24, however, when higher order differential equations are considered the reduction always leads to the same form. Moreover, in proving the existence and uniqueness of solutions of differential equations the integral form is known to be best suited. We mention a few more important examples of nonlinear oscillatory systems, which appear as differential equations of the form 2.24 and can always be reduced to an equation of the form 2.26.

(iii) Vanderpol equation:

$$\ddot{x} + x - \mu(1-x^2)\dot{x} = 0, \quad \mu > 0 \quad 2.28$$

(iv) Rayleigh equation:

$$\ddot{x} + ax - (b - c\dot{x}^2)\dot{x} = 0, \quad a, b, c > 0 \quad 2.29$$

Both these equations arise in self oscillatory systems, in which the dissipation coefficient can change its sign depending on the amplitude of oscillation as in (iii) or on its derivative as in (iv). Thus small oscillations will expand and large oscillations will die down.

(v) Let us consider two rotating shafts connected through a nonlinear elastic bar. Let  $J_1$  and  $J_2$  be their moments of inertia, and  $\theta_1$  and  $\theta_2$  their angles of rotation, and  $f(\theta) = f(\theta_1 - \theta_2)$  the nonlinear characteristic of the coupling. The differential equation for this system is

$$\ddot{\theta} + a f(\theta) = 0, \quad a = \left( \frac{J_1 + J_2}{J_1 J_2} \right) \quad 2.30$$

All these examples involve differential equations of the same form as 2.24.

(vi) In the study of phenomenon of wave propagation, the following differential equation has been widely studied.

$$(a(x) \phi^{(1)}(x))^{(1)} + b(x) \phi(x) = 0, \quad x \in I \quad 2.31$$

This equation is immediately reduced to the following integral equation,

$$\phi(x) = \phi(0) + a(0) \phi^{(1)}(0) h(x) + \int_0^x [h(x) - h(\xi)] b(\xi) \phi(\xi) d\xi \quad 2.32$$

where  $h(x) = \int_0^x \frac{1}{a(\sigma)} d\sigma$  with  $a(x) \neq 0 \quad \forall x \in I$ .

The differential equation 2.31 represents a wide variety of physical phenomenon, such as:-

propagation of electric waves along transmission lines, in which case  $\phi$  is the voltage difference between the lines,  $1/a(x)$ ,  $b(x)$  are the distributed series impedance and the distributed shunt impedance respectively.

the one dimensional motion of periodic sound waves in fluid with  $\phi$  interpreted as the pressure in the fluid at the point  $x$  and  $\psi$  as the velocity at that point, where  $\psi^{(1)}(x) = b(x) \phi(x)$ . Here  $a(x) = i/\omega\rho$  and  $b = \omega/i\rho c^2$ , where,  $\rho$  is the density of the fluid and  $c$  the velocity of sound in fluid.

In quantum mechanics 2.31 represents the time independent schrodinger's equation for a particle. In this case  $a(x) = \hbar/im$  and  $b(x) = \frac{i}{2\hbar} (E - v(x))$ , where  $E$  is the total energy of the particle and  $v(x)$  is the field in which the particle is moving, and  $\phi$  the wave function. In case  $v$  depends both on  $x$  and  $\phi$ , the equation 2.31 becomes a nonlinear equation and the corresponding integral equation becomes

$$\phi(x) = \phi(0) + a(0) \phi^{(1)}(0) h(x) + \int_0^x (h(x) - h(\xi)) b(\xi) \phi(\xi) d\xi \quad 2.32'$$

This equation may be used to study the effects of nonlinearities on wave motion.

(b) From the previous considerations it appears that the first order Uryson operator, (equation 1.26) may be encountered in systems where there is a complex interaction within the system itself. A simple example may be constructed from the following differential equation,

$$\left. \begin{aligned} (a(x, \varphi(x)) \varphi^{(1)}(x))^{(1)} + b(x, \varphi(x)) = 0, \quad x \in I, \text{ and} \\ a(x, \varphi(x)) \neq 0 \quad \forall x \in I. \end{aligned} \right\} 2.33$$

This equation can be reduced to the following integral equation,

$$\varphi(x) = \varphi(0) + \int_0^x R(x, \eta, \varphi(\eta)) d\eta \quad 2.34$$

where, 
$$R(x, \eta, \varphi(\eta)) = \frac{a(0, \varphi(0)) \varphi^{(1)}(0)}{a(\eta, \varphi(\eta))} - \int_{\eta}^x \frac{b(\xi, \varphi(\xi))}{a(\xi, \varphi(\xi))} d\xi$$

(c) Examples of nonlinear integral operators of more complex nature can be encountered in systems whose present state depends on all the preceding states through which the systems has passed.

As an example let us consider an elastic bar Fig. 2.3 subjected to a torsional couple  $m$  with the corresponding angle of torsion  $\theta$ .

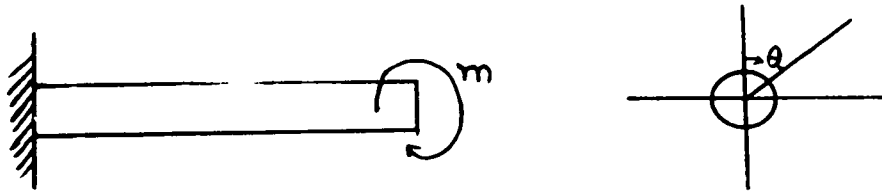


Fig. 2.3. Elastic Bar Under Torsion.

To a first approximation,  $\theta$  and  $m$  is related by,  $\theta = c m$  where  $c$  is a constant of the material. The instantaneous deflection  $\theta$  is an instantaneous function of torsional couple exerted on the bar. Accurate experiments reveal that this is only an approximation; more correct description of the phenomenon requires the consideration of all the preceding torsional moments to which the bar was subjected to. In this case, the equation  $\theta = c m$ , may be replaced by the more accurate functional equation,

$$\theta(t) = c m(t) + F \left[ m(\tau), t \right] \quad 2.35$$

where,  $F$  represents the hereditary nature of the system. If  $F$  is assumed to be an analytic function of the element  $m(\cdot)$  within the interval  $-u \leq m \leq u$ , then equation 2.35 can be expanded into a functional series of the following form,

$$\theta(t) = c m(t) + \sum_{s=1}^{\infty} \int_{-\infty}^t \dots \int_{-\infty}^t K_s(t, \tau_1, \dots, \tau_s) m(\tau_1) \dots m(\tau_s) d\tau_1 \dots d\tau_s \quad 2.36$$

where the kernels  $\{K_s\}$  are the hereditary co-efficients of the bar. This gives us an example of volterra-Frechet operator equation 1.25.

It is expected that magnetic hysteresis may also be accurately represented by a functional equation of the form,

$$B(t) = \mu H(t) + F \left[ H(\cdot), t \right] \quad 2.37$$

which relates magnetic induction with the magnetic field.

If the loop is continuous then equation 2.37 can be expanded into the form 2.36. It is clear that the first term in the series on the right hand side of equation 2.36 represents the linear part of the heridity.

Let us now consider some examples of control systems where various operators, as classified in section 1.3 chapter I, may be encountered.

(d) (i) Let us consider the example, Figure 2.4, of a linear system B preceded by a variable nonlinear zero memory operator G.

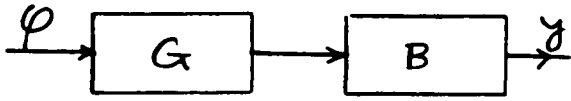


Fig: 2.4: Linear System Preceded by a Nonlinear Zero Memory System.

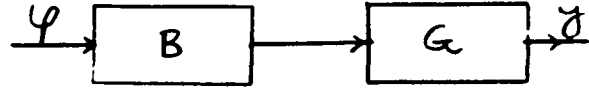


Fig. 2.5: Same in Reverse Order.

$$\text{Let, } Bx(t) = \int_I K(t, \tau) x(\tau) d\tau, \quad t \in I \subset \mathbb{R}^1 \quad 2.38$$

and,

$$G \varphi(t) = g(t, \varphi(t)).$$

$$\text{Then } y = BG\varphi = \int_I K(t, \tau) g(\tau, \varphi(\tau)) d\tau. \quad 2.39$$

This is an example of basic Hammerstein operator. This operator may represent a detector followed by a linear time variable filter.

(ii) If the two operators in Fig. 2.4 are interchanged such that B precedes G, Fig. 2.5, then y is given by

$$y = GB\varphi = g(t, \int_I K(t, \tau) \varphi(\tau) d\tau) \quad 2.40$$

If  $g(t, u)$  is assumed to be analytic in a certain disk D in the complex plane, then for all  $u \in D$  we have,

$$y = GB\varphi = \sum_{s=0}^{\infty} \int_I \left[ \frac{\partial^s g(t, u)}{\partial u^s} \right]_{u=0} K(t, \tau_1) \dots K(t, \tau_s) \varphi(\tau_1) \dots \varphi(\tau_s) d\tau_1 \dots d\tau_s. \quad 2.41$$

This is a special case of an operator of type  $A_V$  as defined by equation 1.25, chapter I. The situation represented by Fig. 2.5 is quite common in control and communication engineering.

(iii) An example of a higher order Hammerstein operator is given by the following configuration.

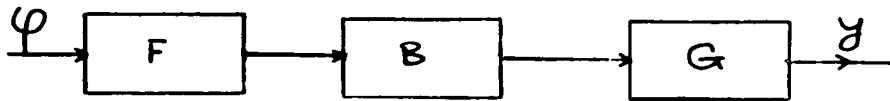


Fig. 2.6: Higher Order Hammerstein Operator:

Here  $F$  and  $G$  are two nonlinear zero memory operators and  $B$  is a linear operator.

$$\begin{aligned}
 y &= GBF\varphi \\
 &= g(t; \int_I K(t, \tau) f(\tau, \varphi(\tau)) d\tau)
 \end{aligned} \tag{2.42}$$

Provided  $g$  satisfies the analyticity property as before, this gives,

$$\begin{aligned}
 y &= \sum_{s=0}^{\infty} \int_I^s \dots \int_I^s \left[ \frac{\partial^s g(t, u)}{\partial u^s} \right]_{u=0} K(t, \tau_1) K(t, \tau_s) f(\tau_1, \varphi(\tau_1)) \dots \\
 &\quad f(\tau_s, \varphi(\tau_s)) d\tau_1 \dots d\tau_s,
 \end{aligned} \tag{2.43}$$

which is a special case of the operator of type  $A_H$  equation 1.22 chapter I.

Another example of Hammerstein operator is provided by a combination of a Volterra-Frechet operator preceded by a nonlinear zero memory operator,

$$\begin{aligned}
 y &= A_H \varphi = A_V G \varphi \\
 &= \sum_{s=0}^{\infty} \int_I^s \dots \int_I^s K_s(t, \tau_1 \dots \tau_s) g(\tau_1, \varphi(\tau_1)) \dots g(\tau_s, \varphi(\tau_s)) \\
 &\quad d\tau_1 \dots d\tau_s.
 \end{aligned} \tag{2.44}$$

Remark

It is important to note that an Higher order Hammerstein operator cannot be represented by a Volterra-Frechet operator without introducing products of multiple impulse functions, into the corresponding Kernels. This is true even if one assumes for maximum generality that the Kernels of the operator are Lebesgue measurable .

(iv) A higher order Uryson operator, equation 1.21, may be constructed by a basic first order Uryson operator, equation 1.26, (both of chapter I) followed by a nonlinear zero memory operator, Fig. 2.7.

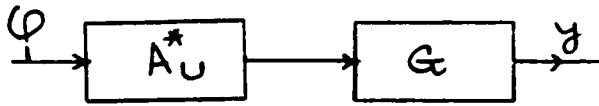


Fig. 2.7: Higher Order Uryson Operator.

Thus,

$$y = GA_U^* \varphi = g(t, \int_I K(t, \tau, \varphi(\tau)) d\tau) \tag{2.45}$$

If  $g$  satisfies the usual conditions, we can expand the right hand side and obtain a special case of an higher order Uryson operator,

$$i.e \quad y = A_U \varphi = GA_U^* \varphi = \sum_{s=0}^{\infty} \int_{I^s} \left( \frac{\partial^s g}{\partial u^s} \right)_{u=0} K(t, \tau_1; \varphi(\tau_1)) \dots K(t, \tau_s; \varphi(\tau_s)) d\tau_1 \dots d\tau_s \cdot \tag{2.46}$$

$t \in I.$

In this section we have presented some basic examples of nonlinear operators each of which belongs to either of the classes of operators as described in section 1.3 chapter I. In the study of control

systems, the plant may, in general, be described by either of the operators  $A_U$ ,  $A_H$ , or  $A_V$ . For better performance and stability reasons, it is sometime absolutely necessary to provide feedback loops around the plant. This leads to complicated nonlinear integral equations and to the questions of the existence and uniqueness of their solutions. This problem is considered in the following section. The principle of contraction mapping and topological fixed point principles are employed to solve feedback problems.

In subsections 2.2-1, 2.2-2 and 2.2-3 we study the analysis of systems which are represented by integral operators with upper limit fixed, and in subsection 2.2-4, with upper limit variable.

## 2.2 NONLINEAR FUNCTIONAL EQUATIONS AND THEIR RESOLVENTS.

### 2.2.1 : The Principle of Contraction Mapping.

The following theorem will be used in the solution of a class of feedback control problems.

#### Theorem 2.1

Let  $T$  be a closed sphere in a Banach space  $E$  and let  $A$  be an operator defined on  $T$  and suppose  $A$  satisfies the Lipschitz condition

$$\| A \varphi_1 - A \varphi_2 \| \leq a \| \varphi_1 - \varphi_2 \| \quad \forall \varphi_1, \varphi_2 \in T \quad 2.47$$

where  $a < 1$ . Let  $AT \subset T$  (i.e  $A$  transforms  $T$  into itself).

(i) Then the equation

$$\varphi = A \varphi \quad 2.47a$$

has a unique solution  $\varphi^*$  in  $T$  and the solution can be computed by the method of successive approximations by the formula

$$\varphi_n = A \varphi_{n-1} \quad (n = 1, 2, \dots) \quad 2.47b$$

where  $\varphi_0$  is any arbitrary element in  $T$ .

(ii) The rate of convergence of the process is governed by the inequality,

$$\|\varphi_n - \varphi^*\| \leq \frac{a^n}{1-a} \|\varphi_1 - \varphi_0\| \quad (n = 1, 2, \dots) \quad 2.48$$

Proof :

(i) Let  $\varphi_0 \in T$  and let  $\{\varphi_n\}$  be the sequence given by 2.47b. By 2.47, for an arbitrary  $n$  we have

$$\begin{aligned} \|\varphi_{n+1} - \varphi_n\| &= \|A\varphi_n - A\varphi_{n-1}\| \leq a \|\varphi_n - \varphi_{n-1}\| \dots \\ &\dots \leq a^n \|\varphi_1 - \varphi_0\| \quad 2.49 \end{aligned}$$

Hence,

$$\begin{aligned} \|\varphi_{n+k} - \varphi_n\| &\leq a^n \|\varphi_k - \varphi_0\| \\ &\leq a^n \left\| \sum_{s=1}^k (\varphi_s - \varphi_{s-1}) \right\| \\ &\leq a^n \sum_{s=1}^k \|\varphi_s - \varphi_{s-1}\| \\ &\leq a^n \|\varphi_1 - \varphi_0\| \sum_{s=1}^k a^{s-1} \end{aligned}$$

Since  $a$  lies in the interval  $0 \leq a < 1$  we have

$$\|\varphi_{n+k} - \varphi_n\| \leq \frac{a^n}{1-a} \|\varphi_1 - \varphi_0\| \quad 2.50$$

$$(n = 0, 1, 2, \dots)$$

Hence the sequence  $\{\varphi_n\}$  forms a cauchy sequence in the Banach space  $E$ .

Since the Banach space  $E$  is complete with respect to the metric induced by its norm, the sequence  $\varphi_n$  converges in the norm

to a limit  $\varphi^*$  in E. Obviously  $\varphi^*$  is in T.

Now we show that  $\varphi^*$  is a solution of equation 2.47a.

$$\begin{aligned}\|A \varphi^* - \varphi^*\| &= \|A \varphi^* - \varphi_{n+1} + \varphi_{n+1} - \varphi^*\| \\ &\leq \|A \varphi^* - \varphi_{n+1}\| + \|\varphi_{n+1} - \varphi^*\| \\ &= \|A \varphi^* - A \varphi_n\| + \|\varphi_{n+1} - \varphi^*\| \\ &\leq a \|\varphi^* - \varphi_n\| + \|\varphi_{n+1} - \varphi^*\|\end{aligned}$$

since the expression on the left is independent of  $n$  and is true for all  $n$ , we let  $n$  approach to infinity on the right. Then both terms on the right tend to zero and the desired result follows.

$$\|A \varphi^* - \varphi^*\| = 0 \text{ ie. } \varphi^* = A \varphi^* \text{ a.e..}$$

The uniqueness of the solution of the equation 2.47a is established very simply.

Let  $\psi^*$  be another solution of 2.47a.

Then

$$\begin{aligned}\|\varphi^* - \psi^*\| &= \|A \varphi^* - A \psi^*\| \\ &\leq a \|\varphi^* - \psi^*\|\end{aligned}$$

since  $a < 1$ , it is impossible to satisfy the inequality unless  $\varphi^* = \psi^*$ . This proves the uniqueness of the solution.

(ii) In order to obtain the rate of convergence expressed by inequality 2.48 we let  $k \rightarrow \infty$  in 2.50 and obtain

$$\|\varphi^* - \varphi_n\| \leq \frac{a^n}{1-a} \|\varphi_1 - \varphi_0\| \quad \text{Q. E. D.}$$

This theorem, originally proved by Banach himself, establishes the so called principle of contraction mapping in complete

metric spaces. The fixed point principle of Schauder [6] is similar to the principle of contraction mapping as presented above. Because, both principles require that  $A T \subset T$ . But these principles are independent of each other.

Schauder's principle states that if a continuous operator  $A$  transforms a closed convex set  $T$  of a Banach space  $E$  into a compact subset of  $T$  then there exists a point  $\varphi \in T$  such that  $\varphi = A\varphi$ . Schauder's principle and the principle of contraction mapping are particular cases of the following fixed point theorem as pointed out by Krasnoselskii [3].

Let  $T$  be a closed, convex, and bounded set of a Banach space  $E$ . Let  $A$  and  $B$  be operators defined on  $T$  and assume  $A$  and  $B$  satisfy the following conditions:-

(i)  $A\varphi + B\psi \in T \quad \forall \varphi, \psi \in T$

(ii) The operator  $A$  satisfies the lipschitz condition,

$$\|A\varphi_1 - A\varphi_2\| \leq \alpha \|\varphi_1 - \varphi_2\|$$

(iii) The operator  $B$  is continuous and compact.

Under these conditions there exists a point  $\varphi^* \in T$  such that

$$A\varphi^* + B\varphi^* = \varphi^*$$

## 2.2-2: Resolvent of Nonlinear Operators and its Properties.

In many feedback control problems we are required to solve certain integral equations of either of the following kinds.

(i)  $\varphi = x + \lambda A\varphi$

(ii)  $\varphi = \lambda A\varphi$ .

where the operator  $A$  may be anyone of the types  $A_U, A_H, A_L$  or  $A_v$  or their suitable combination belonging to the algebra  $B$  as in chapter I.

This is very easily seen from the following consideration.

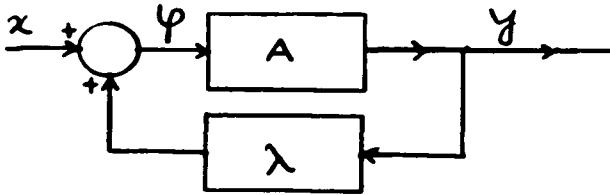


Fig. 2.8: A Feedback Configuration.

In figure 2.8, the operator  $A$  may be any one of the types mentioned and  $\lambda$  is the static gain of an amplifier connected in the feedback loop. The relevant equations of the feedback system are:

$$\begin{aligned}\varphi &= x + \lambda y \\ y &= A\varphi.\end{aligned}$$

$$\text{therefore} \quad \varphi = x + \lambda A\varphi \quad 2.51$$

In the case of regulators we arrive at the second equation,

$$\varphi = \lambda A\varphi. \quad 2.52$$

In the following theorem we will study the properties of the resolvent operator corresponding to the operator equations 2.51 and 2.52. The measure of the set on which the functions  $x(\cdot)$  and  $y(\cdot)$  are defined may be finite or infinite.

Theorem: 2.2

Let  $A$  be a nonlinear operator (one of the types mentioned in section 1.3) defined on some Banach space  $E$  and let  $A$  satisfy on the sphere  $T_\rho$  of radius  $\rho$  about the origin  $\theta$  in  $E$ , the Lipschitz condition

$$\|A\varphi_1 - A\varphi_2\| \leq \alpha(\rho) \|\varphi_1 - \varphi_2\|, \quad \forall \varphi_1, \varphi_2 \in T_\rho,$$

where  $\alpha(\rho) > 0$  is a number as defined before.

Under this condition the equation  $\varphi = x + \lambda A\varphi$  has a unique solution  $\varphi^* \in T_\rho$  provided, in addition, the following inequalities are satisfied:-

$$(a) \quad |\lambda \alpha(\rho)| < 1$$

$$(b) \quad \|x\| \leq (1 - |\lambda \alpha(\rho)|) \rho.$$

Two cases may arise, (i)  $A\theta = \theta$ , (ii)  $A\theta \neq \theta$ .

Proof: (i)  $A\theta = \theta$

Let us define an operator  $B$  such that

$$\varphi = B\varphi. \quad 2.53$$

where  $B\varphi = x + \lambda A\varphi \quad 2.54$

so  $\|B\varphi_1 - B\varphi_2\| \leq |\lambda| \|A\varphi_1 - A\varphi_2\|$   
 $\leq |\lambda \alpha(\rho)| \|\varphi_1 - \varphi_2\| \quad 2.55$

$$\forall \varphi_1, \varphi_2 \in T_\rho$$

Therefore if the condition (a) is satisfied ie if  $|\lambda \alpha(\rho)| < 1$  then the operator  $B$  satisfies the Lipschitz condition with a constant less than unity.

Again for  $\forall \varphi_1, \varphi_2 \in T_\rho$

$$\|B\varphi\| \leq \|x\| + |\lambda| \|A\varphi - A\theta\|$$

$$\|B\varphi\| \leq \|x\| + |\lambda \alpha(\rho)| \|\varphi\|$$

$$< \|x\| + |\lambda \alpha(\rho)| \rho \quad 2.56$$

Thus, if  $\|x\| \leq (1 - |\lambda \alpha(\rho)|) \rho$  then the operator  $B$  as defined by 2.54 transforms the sphere  $T_\rho$  into itself,  $\forall \varphi \in T_\rho$ .

Hence the operator  $B$  satisfies the principles of contraction mapping and therefore by the theorem 2.1, there exists a unique

solution  $\varphi^* \in T_\rho$  for the equation  $\varphi = x + \lambda A\varphi$ , and the solution is computed by the previous method.

$$(ii) \quad A\theta \neq \theta$$

If condition (a) as stated in the theorem 2.2 is satisfied, inequality 2.55 still remains valid.

In order that B transforms the sphere  $T_\rho$  into itself the condition (b) is modified to

$$(b') \quad \|x + A\theta\| \leq (1 - |\lambda a(\rho)|) \rho.$$

The rest follows as in (i)

Q.E.D.

Let us now present the following definition for the resolvent operator corresponding to the equation 2.51.

Definition 2.1

An operator  $R_\lambda$  is defined as the resolvent of the operator equation 2.51 corresponding to the parameter  $\lambda$  if it transforms each element  $x \in E$  satisfying the inequality (b) or (b') in theorem 2.2 on to the solution  $\varphi^*$  of the equation 2.51.

Definition 2.2

Resolvent set: The resolvent set of the operator A is the set of all complex numbers  $\lambda$  such that  $R_\lambda$  exists. This set will be denoted by  $\rho(A)$ . The spectral set of the operator A to be denoted by  $\sigma(A)$  is the complement of the set  $\rho(A)$  in the complex  $\lambda$  plane.

By definition 2.1,  $R_\lambda$  is the resolvent of the equation,

$$\varphi = x + \lambda A\varphi \quad \forall \lambda \in \rho(A)$$

if  $\varphi$  is given by the equality,

$$\varphi = R_\lambda x$$

2.57

$\forall x \in E$  such that,

$$\|x\| \leq (1 - |\lambda a(\rho)|) \rho \triangleq \xi(\rho)$$

CONTINUITY OF THE OPERATOR  $R_\lambda$ .

Theorem 2.3

The resolvent operator  $R_\lambda$  corresponding to the operator equation  $\varphi = x + \lambda A\varphi$ , is continuous in the sphere  $\xi(\rho) \subset E$  of radius less than  $(1 - |\lambda a(\rho)|)\rho$  whenever the operator  $A$  is continuous in  $T \subset E$ .

Proof

Substituting 2.57 in the integral equation  $\varphi = x + \lambda A\varphi$  we have

$$R_\lambda x = x + \lambda A R_\lambda x. \quad 2.58$$

$\forall \lambda \in \rho(A)$  and  $x \in \xi(\rho)$ .

Hence

$$R_\lambda x_1 = x_1 + \lambda A R_\lambda x_1 \quad 2.59a$$

$$R_\lambda x_2 = x_2 + \lambda A R_\lambda x_2 \quad \forall x_1, x_2 \in \xi(\rho) \quad 2.59b$$

Subtracting 2.59b from 2.59a and taking the norm on either side we obtain,

$$\|R_\lambda x_1 - R_\lambda x_2\| \leq \|x_1 - x_2\| + |\lambda a(\rho)| \|R_\lambda x_1 - R_\lambda x_2\|$$

Therefore,

$$\|R_\lambda x_1 - R_\lambda x_2\| \leq \frac{\|x_1 - x_2\|}{(1 - |\lambda a(\rho)|)} \quad \forall x_1, x_2 \in \xi(\rho) \quad 2.60$$

This inequality proves that  $R_\lambda$  is a continuous operator in the sphere  $\xi(\rho)$ .

Moreover, from the identities 2.59 we have

$$\|x_1 - x_2\| \leq \|R_\lambda x_1 - R_\lambda x_2\| + |\lambda a(\rho)| \|R_\lambda x_1 - R_\lambda x_2\|$$

which yields the following inequality.

$$\|R_\lambda x_1 - R_\lambda x_2\| \geq \frac{\|x_1 - x_2\|}{(1 + |\lambda a(\rho)|)} \quad \forall x_1, x_2 \in \xi(\rho) \quad 2.61$$

From 2.60 and 2.61 we have,

$$\frac{\|x_1 - x_2\|}{(1 + |\lambda a(\rho)|)} \leq \|R_{\lambda} x_1 - R_{\lambda} x_2\| \leq \frac{\|x_1 - x_2\|}{(1 - |\lambda a(\rho)|)} . \quad 2.62$$

This inequality remains valid for both (i)  $A\theta = \theta$  and (ii)  $A\theta \neq \theta$ .

If  $A\theta = \theta$ ,  $R\theta = \theta$  and it follows from 2.62 that

$$\frac{\|x\|}{(1 + |\lambda a(\rho)|)} \leq \|R_{\lambda} x\| \leq \frac{\|x\|}{(1 - |\lambda a(\rho)|)} \quad 2.63$$

$$\forall x \in \xi(\rho)$$

If, on the other hand,  $A\theta \neq \theta$  then the inequality 2.63 becomes,

$$\frac{\|g\|}{(1 + |\lambda a(\rho)|)} \leq \|R_{\lambda} g\| \leq \frac{\|g\|}{(1 - |\lambda a(\rho)|)} \quad 2.63'$$

where  $g = x + \lambda A\theta$ , with  $\|x + \lambda A\theta\| < (1 - |\lambda a(\rho)|) \rho$ .

This case arises when the system  $A$  has an output even in the absence of any input. For example in the case of operators of type  $A_V$  and  $A_H$ , the zero-input responses are  $\lambda K_O(t)$  and  $\lambda A G\theta$  respectively.

#### Theorem 2.4

The resolvent operator  $R_{\lambda}$  is a continuous function of the parameter  $\lambda$  provided  $\lambda \in \rho(A)$ , and  $x \in \xi(\rho)$ .

#### Proof:

Let  $\lambda$  and  $\mu$  both belong to the resolvent set  $\rho(A)$ . Then the resolvents  $R_{\lambda}$  and  $R_{\mu}$  are defined provided

$$\|x\| \leq [1 - (|\lambda a(\rho)| \vee |\mu a(\rho)|)] \rho . \quad 2.64$$

If these conditions are satisfied then we have from 2.58,

$$\begin{aligned} R_{\lambda} x &= x + \lambda A R_{\lambda} x, \\ R_{\mu} x &= x + \mu A R_{\mu} x, \quad \text{and hence,} \end{aligned}$$

$$\begin{aligned} \|R_{\lambda} x - R_{\mu} x\| &= \|\lambda A R_{\lambda} x - \mu A R_{\mu} x\| \\ &= \|\lambda A R_{\lambda} x - \mu A R_{\lambda} x + \mu A R_{\lambda} x - \mu A R_{\mu} x\| \\ &\leq |\lambda - \mu| \|A R_{\lambda} x\| + |\mu a(\rho)| \|R_{\lambda} x - R_{\mu} x\| \end{aligned}$$

Therefore,

$$\|R_\lambda x - R_\mu x\| \leq \frac{|\lambda - \mu| \|AR_\lambda x\|}{(1 - |\mu a(\rho)|)} \quad 2.65$$

If  $A\theta = \theta$ ,  $R\theta = \theta$ , Hence it follows that

$$\|R_\lambda x - R_\mu x\| \leq \frac{|(\lambda - \mu) a(\rho)| \|R_\lambda x\|}{(1 - |\mu a(\rho)|)} \quad 2.66$$

Substituting the value of  $\|R_\lambda x\|$  from the inequality 2.63, into the inequality 2.66 we have,

$$\|R_\lambda x - R_\mu x\| \leq \frac{a(\rho) \|x\|}{(1 - |\mu a(\rho)|)(1 - |\lambda a(\rho)|)} |\lambda - \mu| \quad 2.67$$

This proves that the resolvent  $R_\lambda$  is a continuous operator function of the parameter  $\lambda$ , in the real case the amplifier gain.

(ii)  $A\theta \neq \theta$

We now prove the continuity of the resolvent  $R_\lambda$  with respect to  $\lambda$  for the zero input response case (ie  $A\theta \neq \theta$ ).

In this case the inequality 2.64 is to be replaced by

$$\|x + \lambda A\theta\|, \|x + \mu A\theta\| \leq [1 - (|\lambda a(\rho)| \vee |\mu a(\rho)|)] \rho \quad 2.68$$

Let us rewrite the equation

$$\begin{aligned} \varphi &= x + \lambda A\varphi \quad \text{as} \\ \varphi &= x + \lambda A\theta + \lambda A_1 \varphi \end{aligned} \quad 2.69$$

where  $A_1 \theta = \theta$ .

Let us define  $R_\lambda$  as before, such that

$$\varphi = R_\lambda x. \quad 2.70$$

Substituting 2.69 in 2.58 we obtain for all  $\lambda$  and  $\mu \in \rho(A)$  the following pair of equations.

$$\left. \begin{aligned} R_\lambda x &= x + \lambda A\theta + \lambda A_1 R_\lambda x \\ \text{and } R_\mu x &= x + \mu A\theta + \mu A_1 R_\mu x, \end{aligned} \right\} \quad 2.71$$

for all  $x$  satisfying the condition 2.68.

From 2.71 we obtain,

$$\|R_\lambda x - R_\mu x\| \leq \frac{|\lambda - \mu| \|A\theta\| + |\lambda - \mu| \|A_1 R_\lambda x\|}{(1 - |\mu a(\rho)|)} \quad 2.72$$

Since  $A_1 \theta = \theta$ , 2.72 becomes,

$$\|R_\lambda x - R_\mu x\| \leq \frac{|\lambda - \mu| \|A\theta\| + |(\lambda - \mu)a(\rho)| \|R_\lambda x\|}{(1 - |\mu a(\rho)|)} \quad 2.73$$

From 2.71a we obtain

$$\|R_\lambda x\| \leq \frac{\|x\| + |\lambda| \|A\theta\|}{(1 - |\lambda a(\rho)|)} \quad 2.74$$

Substituting the value of  $\|R_\lambda x\|$  from 2.74 into the inequality 2.73 we obtain,

$$\|R_\lambda x - R_\mu x\| \leq \frac{a(\rho) \|x\| + \|A\theta\|}{(1 - |\mu a(\rho)|)(1 - |\lambda a(\rho)|)} |\lambda - \mu| \quad 2.75$$

for all  $x$  satisfying the inequalities 2.68. Q.E.D.

### Remarks

Another obvious property of the resolvent  $R_\lambda$  is that if  $A$  is odd (ie  $A(-\varphi) = -A\varphi$ ) then  $R_\lambda$  is also odd.

The resolvent operator  $R_\lambda$  corresponding to the value of the parameter  $\lambda$  and the operator equation  $\varphi = x + \lambda A\varphi$ , is the so called Banach space inverse of the operator  $(I - \lambda A)$ , ie

$$R_\lambda (I - \lambda A) = (I - \lambda A) R_\lambda = I.$$

### 2.2-3 EXISTENCE OF LOCAL AND NONLOCAL SOLUTIONS.

The conditions on the existence theorems for equations of the type

$$\varphi = x + \lambda A\varphi \quad 2.76$$

as shown in this section frequently consist of upperbounds on the

absolute value of  $\lambda$ . Such theorems are usually referred to as local theorems relative to  $\lambda$ . Nonlocal theorems relative to  $\lambda$  are those existence theorems in which  $\lambda$  can take on values in some other regions, for example, in a system of intervals.

The principle of contraction mapping is applicable to proofs of both local and nonlocal existence theorems.

Let  $A$  be defined in the sphere  $T_\rho \subset E$  and let  $x \in E$ . Furthermore let us assume that  $A$  satisfies the Lipschitz condition,

$$\|A\varphi_1 - A\varphi_2\| < a(\rho) \|\varphi_1 - \varphi_2\|; \varphi_1, \varphi_2 \in T_\rho.$$

As shown in this section, for local existence theorems in the sphere  $T_\rho$  of the solution of the above equation, two conditions are imposed on  $\lambda$ ; one of these assures the inequality  $|\lambda a| < 1$ , the second one also restricts  $\lambda$  from above and requires that  $x + \lambda A\varphi \in T_\rho$  whenever  $\varphi \in T_\rho$ .

Nonlocal existence theorems, are obtained in the case when the operator  $A$  can be split into a linear and a nonlinear part, i.e.  $A = A_1 + A_2$ , where  $A_1$  is linear.

Examples of such operators are: The Volterra operator  $A_V$ ; the operators of type  $A_H$  with the operator  $G$  containing a linear term in  $\varphi$ ; the operators of type  $A_L$  with a term of first degree in  $\varphi$  and of degree zero in  $\Psi$ .

If  $\lambda \notin \sigma(A_1)$ , then equation 2.76 can be written as

$$\varphi = \Pi_\lambda x + \lambda \Pi_\lambda A_2 \tag{2.77}$$

$$\text{where } \Pi_\lambda = (I - \lambda A_1)^{-1}. \tag{2.78}$$

It is easily seen that the solution of equation 2.77 is the solution of equation 2.76.

The operator defined by the right hand side of equation 2.76 will satisfy a Lipschitz condition in  $T_\rho$  with a constant less than unity if  $A_2$  satisfies a Lipschitz condition with a constant

$\beta$  such that

$$|\lambda \beta| \|\Gamma_\lambda\| < 1$$

If further  $\|x\| \leq \left( \frac{1 - |\lambda \beta| \|\Gamma_\lambda\|}{\|\Gamma_\lambda\|} \right) \rho$ ,

then the principles of contraction mapping are satisfied and therefore, the equation 2.77 will have a unique solution  $\varphi^* \in T_\rho$ . This solution can be computed by the method described in section 2.2-1.

That is,

choose,  $\varphi_0 = \Gamma_\lambda x$

and compute,

$$\varphi_n = \Gamma_\lambda x + \lambda \Gamma_\lambda A_2 \varphi_{n-1}.$$

} 2.79

There is no particular reason for choosing  $\varphi_0$  as  $\Gamma_\lambda x$ , one may choose any element  $y \in T_\rho$ . Therefore, the conditions in nonlocal existence theorems are the following two: (i)  $\lambda \notin \sigma(A_1)$  (ii) the non-linear part  $A_2$  satisfies a Lipschitz condition with a suitable constant. The resolvent  $\Gamma_\lambda$  (2.78) of the linear operator  $A_1$  corresponding to the parameter  $\lambda$  is similar to Neumann series [7] arising in the theory of linear integral equations.

Here,

$$\Gamma_\lambda = I + \lambda \left( \sum_{n=1}^{\infty} \lambda^{n-1} A_1^{(n)} \right) \tag{2.80}$$

where the expression inside the bracket is usually known as the Neumann series, and  $A_1^{(n)} = \underbrace{A_1 \odot A_1 \odot \dots \odot A_1}_{n \text{ times}}$  is called the nth

iterated kernel defined by,

$$A_1^{(n)}(t, \tau) = \int_I A_1^{(n-k)}(t, \xi) A_1^{(k)}(\xi, \tau) d\xi; \quad k = 1, 2, \dots, n-1;$$

with  $t, \tau \in I$ .

and  $A_1^{(1)}(t, \tau) = A_1(t, \tau)$ .

Remarks:

If in a feedback loop of the system A in figure 2.8, we add a linear operator K, then the equation 2.51 may have either of the two forms,

$$(i) \quad \varphi = x + \lambda KA\varphi$$

$$(ii) \quad \varphi = x + \lambda AK\varphi$$

In this situation the inequalities (a) and (b) in theorem 2.2 need to be replaced by  $a_i, b_i$  and  $a_{ii}, b_{ii}$  respectively,

where,

$$(a_i) \quad |\lambda a(\rho)| \|K\| < 1$$

$$(b_i) \quad \|x\| \leq (1 - |\lambda a(\rho)| \|K\|) \rho$$

$$(a_{ii}) \quad |\lambda a(\rho)| \|K\| < 1$$

$$(b_{ii}) \quad \|x\| \leq (1 - |\lambda a(\rho)| \|K\|) \frac{\rho}{\|K\|}.$$

} 2.81

and  $\varphi_i \in T_\rho$  and  $\varphi_{ii} \in \frac{T_\rho}{\|K\|}$ .

It is clear from these inequalities that by adjusting the amplifier and the linear system K it is possible to increase the sensitivity of the overall system. In particular, if the linear operator K is chosen such that  $\|K\| = 1$  then the inequalities 2.81 reduce to the original form thus presenting the possibility of improving the dynamic response of the system without affecting its static gain.

2.2-4: NONLINEAR CAUSAL SYSTEMS.

A large class of nonlinear systems can be described by the volterra functional polynomial of some finite degree with the upper limits of integrations variable (in case of control systems the upper limit is the present time) and the lower limits fixed say at " $a_0$ " where " $a_0$ " may take the value  $-\infty$ , if the system has

infinite memory. It will be shown in the following theorem, that, in this case, the parameter  $\lambda$  appearing in the feedback system can take values on any arbitrary finite interval on the real line. Let us assume that the plant  $A$  in the forward loop is described by,

$$A \varphi = \sum_{s=0}^{N < \infty} \int_{a_0}^t \dots \int_{a_0}^t K_s(t, \tau_1, \dots, \tau_s) \varphi(\tau_1) \dots \varphi(\tau_s) d\tau_1 \dots d\tau_s \quad 2.82$$

$$a_0 \leq \tau_i \leq t \leq b_0; i=1, \dots, s,$$

and the equation to be solved is the integral equation  $\varphi = x + \lambda A \varphi$  arising from the corresponding feedback system. It is clear that this equation is quite complicated and there exists no direct method of solution. The only possibility is to construct a sequence of approximations and prove that the sequence converges to the solution of the equation. This leads us to the consideration of the following theorem.

Theorem 2.5:

Let the set of kernels  $\{K_s\}$  be Lebesgue measurable and Lebesgue square integrable real valued functions in  $I^{s+1}$  for all  $s \in J_N^+$  (the set of non-negative integers, including  $N$ ) and let  $\{K_s\}$  be symmetric in the variables  $\tau_i, i = 1, 2, \dots, s$ .

Let the functional  $H_N(t, \tau; \varphi(\cdot))$  defined by,

$$H_N(t, \tau; \varphi(\cdot)) = \sum_{r=1}^N \int_{a_0}^t \dots \int_{a_0}^t K_r(t, \tau; \tau_1 \dots \tau_{r-1}) \varphi(\tau_1) \dots \varphi(\tau_{r-1}) d\tau_1 \dots d\tau_{r-1}$$

$$t, \tau \in I: a_0 \leq \tau \leq t \leq b_0,$$

satisfy the inequality,

$$\sup_{\varphi \in S_\alpha} |H_N(t, \tau; \varphi(\cdot))| \leq H(t, \tau) \text{ in every sphere } S_\alpha \text{ of finite radius } \alpha \text{ in the space } L^2 \text{ such that } H(t, \tau) \in L^2(I^2)$$

under these hypothesis, the feedback equation

$$\varphi = x + \lambda A \varphi \tag{2.83}$$

as described before, has a unique  $L^2$  solution for every  $x \in L^2$ .

Proof:

The equation  $\varphi = x + \lambda A \varphi$  can be rewritten as,

$$\varphi = (x + \lambda K_0) + \lambda A_1 \varphi \tag{2.84}$$

where  $K_0(t) \in L^2(I)$  is the zero input response of the operator  $A$  and  $A_1$  represents the rest of  $A$ .

Let  $\varphi_0 = 0$  and let us consider the sequence

$$\varphi_{n+1} = x + \lambda A \varphi_n \quad (n = 0, 1, 2, 3, \dots) \tag{2.85}$$

since  $\varphi_0 = 0$ ,  $\varphi_1 = x + \lambda K_0 \triangleq g(t)$ .

since both  $x$  and  $K_0$  belong to  $L^2(I)$ ,  $g(t) \in L^2(I)$  for each finite  $\lambda$ . Now by the corollary to the theorem 1.5 we know that the operator  $A$  maps every element of the space  $L^2$  into an element which also belongs to  $L^2$ . Therefore, the sequence  $\{\varphi_n\}$  is an  $L^2$  sequence. Let  $\|\varphi_n\| = a_n$  for each  $n \in J^+$ . Since  $\{a_n\}$  is a bounded sequence of real numbers,  $\sup_n a_n$  is defined and since the sphere  $s_a$  can be chosen as the closure of  $T_a \triangleq \bigcup_{n=1}^{\infty} S_{a_n}$ , the hypothesis made on the functional  $H_N$  can be satisfied.

Now,

$$[\varphi_1 - \varphi_0]^2 = [g(t)]^2 \text{ and since } g \in L^2(I) \text{ we have,}$$

$$\int_{a_0}^t [g(\xi)]^2 d\xi = M^2 < \infty \tag{2.85'}$$

where  $M$  is some real number.

similarly,

$$\begin{aligned} \varphi_2 - \varphi_1 &= \lambda(A_1 \varphi_1 - A_1 \varphi_0) \\ &= \lambda \sum_{s=1}^N \int_{a_0}^t \dots \int_{a_0}^t K_s(t, \tau_1 \dots \tau_s) \left[ \prod_{i=1}^s \varphi_1(\tau_i) - \prod_{j=1}^s \varphi_0(\tau_j) \right] d\tau_1 \dots d\tau_s \end{aligned} \tag{2.86}$$

It is easily verified that we can express the quantity inside the bracket as,

$$\left( \prod_{i=1}^s \varphi_1(\tau_i) - \prod_{j=1}^s \varphi_0(\tau_j) \right) = \sum_{r=1}^s (\varphi_1(\tau_r) - \varphi_0(\tau_r)) \prod_{i=r+1}^s \varphi_1(\tau_i) \prod_{j=1}^{r-1} \varphi_0(\tau_j), \quad 2.87$$

where the terms like  $\prod_{i=s+1}^s \varphi_1(\tau_i)$  and  $\prod_{j=1}^0 \varphi_0(\tau_j)$

in which the upper limit in the product is smaller than the lower limit must be interpreted as one.

By substituting 2.87 into 2.86 and recalling that each of the kernels  $K_s$  is symmetric in the variables  $\tau_i; i = 1, 2, \dots, s$ , we obtain,

$$\begin{aligned} (\varphi_2 - \varphi_1) &= \lambda \sum_{s=1}^N \int_{a_0}^t \dots \int_{a_0}^t K_s(t, \tau, \tau_1, \tau_{s-1}) \left\{ \sum_{r=1}^s \prod_{i=r}^{s-1} \varphi_1(\tau_i) \prod_{j=1}^{r-1} \varphi_0(\tau_j) \right\} (\varphi_1(\tau) - \varphi_0(\tau)) \\ &\qquad \qquad \qquad d\tau \, d\tau_1 \dots d\tau_{s-1}. \quad 2.88 \\ &= \lambda \int_{a_0}^t \left( \sum_{s=1}^N \int_{a_0}^t \int_{a_0}^t K_s(t, \tau, \tau_1 \dots \tau_{s-1}) \left\{ \sum_{r=1}^s \prod_{i=r}^{s-1} \varphi_1(\tau_i) \prod_{j=1}^{r-1} \varphi_0(\tau_j) \right\} d\tau_1 \dots d\tau_{s-1} \right) \times \\ &\qquad \qquad \qquad \times (\varphi_1(\tau) - \varphi_0(\tau)) \, d\tau \quad 2.89 \end{aligned}$$

The interchange of the order of integration is justified by the fact that the kernels  $K_s$  are measurable in all the variables and symmetric with respect to  $\tau, \tau_1 \dots \tau_{s-1}$ . Writing the expression inside the first bracket as  $H_N(t, \tau; \varphi_1(\cdot), \varphi_0(\cdot))$  we have,

$$(\varphi_2 - \varphi_1) = \lambda \int_{a_0}^t H_N(t, \tau; \varphi_1(\cdot), \varphi_0(\cdot)) (\varphi_1(\tau) - \varphi_0(\tau)) \, d\tau \quad 2.90$$

Since both  $\varphi_1$  and  $\varphi_0 \in s_a$ , it follows from the hypothesis on the functional  $H_N$  that,

$$\begin{aligned}
 [\varphi_2 - \varphi_1]^2 &\leq (\lambda N)^2 \int_{a_0}^t [H(t, \tau)]^2 d\tau \int_{a_0}^t [\varphi_1(\tau) - \varphi_0(\tau)]^2 d\tau \\
 &\leq M^2 (\lambda N)^2 \int_{a_0}^t [H(t, \tau)]^2 d\tau
 \end{aligned}
 \tag{2.91}$$

Let  $B^2(t) \triangleq \int_{a_0}^t [H(t, \tau)]^2 d\tau$ ; since  $H(t, \tau)$  is measurable in both  $t$  and  $\tau$  and is in  $L^2$ ,  $B^2(t)$  is defined almost everywhere on  $I$  and that,

$$\int_{a_0}^{b_0} B^2(t) dt \leq b^2 < \infty.$$

Thus,

$$\begin{aligned}
 [\varphi_2 - \varphi_1]^2 &\leq M^2 (\lambda N)^2 B^2(t) \\
 [\varphi_3 - \varphi_2]^2 &\leq M^2 (\lambda N)^4 B^2(t) \int_{a_0}^t B^2(\xi) d\xi \\
 [\varphi_4 - \varphi_3]^2 &\leq M^2 (\lambda N)^6 B^2(t) \int_{a_0}^t B^2(\xi_1) \int_{a_0}^{\xi_1} B^2(\xi_2) d\xi_2 \cdot d\xi_1.
 \end{aligned}$$

$$\text{Let } \int_{a_0}^t B^2(\xi) d\xi = h(t). \tag{2.92}$$

then,  $h(t) = B^2(t)$

and

$$\begin{aligned}
 [\varphi_4 - \varphi_3]^2 &\leq M^2 B^2(t) (\lambda N)^6 \int_{a_0}^t h(\xi) \partial h(\xi) \\
 &\leq M^2 B^2(t) (\lambda N)^6 \frac{h^2(t)}{2} \\
 [\varphi_5 - \varphi_4]^2 &\leq \frac{M^2 B^2(t) (\lambda N)^8}{2} \int_{a_0}^t h^2(\xi) \partial h(\xi) \\
 &\leq \frac{M^2 B^2(t) (\lambda N)^8}{3!} h^3(t)
 \end{aligned}$$

In general, for  $n = 0, 1, 2, 3, \dots$ , we have,

$$\begin{aligned} [\varphi_{n+2} - \varphi_{n+1}]^2 &\leq (\lambda NM)^2 B^2(t) \frac{(h(t)(\lambda N)^2)^n}{n!} \\ &\leq (\lambda NM)^2 B^2(t) \frac{(\lambda Nb)^{2n}}{n!} \end{aligned}$$

which yields,

$$|\varphi_{n+2} - \varphi_{n+1}| \leq (\lambda NM) B(t) \frac{(\lambda Nb)^n}{\sqrt{n!}} \text{ a. e.} \quad 2.93$$

on I.

It follows from equation 2.93 that the infinite series,

$$\varphi^* = \varphi_1 + \sum_{s=0}^{\infty} (\varphi_{s+2} - \varphi_{s+1}) \quad 2.94$$

converges absolutely, whenever  $g(t)$  and  $B(t)$  are finite, since neglecting the first term it admits the majorant

$$Z(t) = (\lambda NM)B(t) \sum_{s=0}^{\infty} \frac{(\lambda Nb)^s}{\sqrt{s!}} \quad 2.95$$

which is always convergent.

Obviously,  $\|\varphi_{n+p} - \varphi_n\| \rightarrow 0$  for any fixed positive integer  $p \geq 1$ .

Therefore  $\{\varphi_n\}$  is a Cauchy sequence and since  $L^2$  is a complete metric space  $\varphi_n$  converges to a limit in  $L^2$ .

Since the  $n$ th partial sum of the series is  $\varphi_n$  and the series admits a majorant whenever  $B(t)$  is finite, it is even true, that  $\varphi_n$  converges almost uniformly to the function  $\varphi^*$ .

$$\text{That is, } \varphi_n \rightarrow \varphi^* \text{ (a. u.)}$$

This however does not imply that  $\varphi^*$  is the solution of the equation 2.83. So we must prove that the limit function  $\varphi^*$  is actually a solution of the equation. That is, we must prove that,

$$(\varphi^* - x - \lambda A \varphi^*) = 0 \text{ a. e.}$$

If we define  $R_n$  by,  $\varphi^* = \varphi_n + R_n$ . 2.96

then,

$|R_n| \leq (\lambda N M) B(t) \sum_{s=n-1}^{\infty} \frac{(\lambda N b)^s}{\sqrt{s!}}$ . This follows from equations 2.94 and 2.95. Since  $B(t) \in L^2$ ,  $R_n \in L^2$  and it follows from the above inequality that  $\|R_n\| \rightarrow 0$ .

From equations 2.85, 2.94 and 2.96 we have,

$$\begin{aligned} (\varphi^* - x - \lambda A \varphi^*) &= R_n + \lambda (A \varphi_{n-1} - A \varphi^*) \\ &= R_n + \lambda \int_{a_0}^t H_N(t, \tau; \varphi_{n-1}(\cdot), \varphi^*(\cdot)) (\varphi_{n-1}(\tau) - \varphi^*(\tau)) d\tau \end{aligned} \quad 2.97$$

Since each of the members of the sequence  $\{\varphi_n\}$  belongs to  $L^2$  and the series 2.94 converges almost uniformly,  $\varphi$  also belongs to the class  $L^2$ . So by hypothesis of the theorem on the functional  $H_N$ , we have,

$$\begin{aligned} [\varphi^* - x - \lambda A \varphi^*]^2 &\leq 2R_n^2(t) + 2(\lambda N)^2 \int_{a_0}^t [H(t, \tau)]^2 d\tau \int_{a_0}^t [\varphi^*(\tau) - \varphi_{n-1}(\tau)]^2 d\tau \\ &\leq 2R_n^2(t) + 2(\lambda N)^2 B^2(t) \int_{a_0}^t [R_{n-1}]^2 d\tau. \end{aligned} \quad 2.98$$

hence,

$$\|\varphi^* - x - \lambda A \varphi^*\|^2 \leq 2\|R_n\|^2 + 2(\lambda N b)^2 \|R_{n-1}\|^2. \quad 2.99$$

The right hand side of this inequality tends to zero as  $n \rightarrow \infty$ .

Therefore,  $\varphi^* - x - \lambda A \varphi^* = 0$ , which proves that  $\varphi^*$  satisfies the equation 2.83 almost everywhere. So  $\varphi^*$  is a solution of equation 2.83. It remains to prove that  $\varphi^*$  is the only  $L^2$  solution of equation 2.83, (modulo the set of almost everywhere null functions). Let us assume that  $\psi^*$  is another  $L^2$  solution of equation 2.83. Then,

$$(\psi^* - \varphi^*) = \lambda (A \psi^* - A \varphi^*)$$

and

$$(\Psi^* - \Phi^*) = \lambda \int_{a_0}^t H_N(t, \tau; \Psi^*(\cdot), \Phi^*(\cdot)) (\Psi^*(\tau) - \Phi^*(\tau)) d\tau.$$

where  $H_N$  is as defined before, since both  $\Psi^*$  and  $\Phi^*$  belong to  $L^2$   $H_N$  satisfies the hypothesis of the theorem and therefore,

$$|\Psi^* - \Phi^*| \leq |(\lambda N)| \int_{a_0}^t |H(t, \tau) (\Psi^*(\tau) - \Phi^*(\tau))| d\tau. \quad \text{a.é.}$$

$$\begin{aligned} \text{ie } [\Psi^* - \Phi^*]^2 &\leq (\lambda N)^2 \int_{a_0}^t [H(t, \tau)]^2 d\tau \cdot \int_{a_0}^t [\Psi^*(\tau) - \Phi^*(\tau)]^2 d\tau \\ &\leq (\lambda N)^2 B^2(t) \int_{a_0}^t [\Psi^*(\tau) - \Phi^*(\tau)]^2 d\tau \end{aligned} \quad 2.100$$

Substituting inequality 2.100 into itself  $n$  times and then setting

$\int_{a_0}^b |\Psi^*(\tau) - \Phi^*(\tau)|^2 d\tau = \rho^2$ , the following inequality is obtained.

$$\int_{a_0}^t [\Psi^* - \Phi^*]^2 d\tau \leq (\lambda N \rho)^2 \frac{(\lambda N b)^{2n}}{n!} \quad 2.101$$

In the limit, as  $n \rightarrow \infty$ , the right hand side tends to zero thus implying that  $\Psi^* = \Phi^*$  a. e.

This completes the proof of the theorem.

Remarks:

(a) By suitable assumptions on the properties of the operator  $G$  it is possible to prove a similar theorem for the operators of type  $A_H$ .

(b) It is to be noted that if the zero-input response of the open loop system is nonzero i.e. if  $K_0(t) \neq \theta$  then, even if the loop is closed, the error  $\Phi$  remains nonzero. If however  $K_0(t) = \theta$  and  $x = \theta$  then  $\Phi = \theta$ . This is because the present equation  $\Phi = \lambda A_1 \Phi$  corresponding to the nonhomogeneous equation studied in the previous

theorem has only the null solution; which is easily proved by substituting the null vector for the vector  $\varphi^*$  in the inequality 2.101.

(c) the steady state solution of the feedback system is stable. This follows from the fact that the equation of variation of the system  $\varphi = x + \lambda A \varphi$  is a linear homogeneous volterra integral equation the solution of which is the null vector.

The proof is quite simple. For let  $\varphi_0$  be the steady state solution of the equation  $\varphi = x + \lambda A \varphi$ .

Then,

$$\varphi_0 = x + \lambda A \varphi_0 \quad \text{with } x, \varphi_0 \in L^2. \quad 2.102$$

Let  $x$  be the fixed  $L^2$  input to the system.

Let us assume that the system has been perturbed by application of an impulse at the input. Let the corresponding change in the steady state solution  $\varphi_0$  of the system be denoted by  $\mu h$  where  $h \in L^2$  and  $\mu$  is a real number lying in the interval  $[0, 1]$ .

Under these conditions equation 2.102 reduces to

$$(\varphi_0 + \mu h) = x + \lambda A(\varphi_0 + \mu h). \quad 2.102'$$

Taking the derivative on either side of this equation with respect to  $\mu$  and assuming that the kernels are symmetric in  $\tau$ 's we have,

$$h(t) = \lambda \int_{a_0}^t \text{grad } A_{\varphi_0} h(\tau) d\tau \quad 2.103$$

where,

$$\text{grad } A_{\varphi_0} = \sum_{s=1}^N \int_{a_0}^t \dots \int_{a_0}^t K_s(t, \tau, \tau_1, \dots, \tau_{s-1}) \varphi_0(\tau_1) \dots \varphi_0(\tau_{s-1}) d\tau_1 \dots d\tau_{s-1} \quad 2.104$$

Let us denote  $\text{grad } A_{\varphi_0}$  by  $R(t, \tau)$

Then the equation 2.103 reduces to the following linear homogeneous volterra integral equation,

$$h(t) = \lambda \int_{a_0}^t R(t, \tau) h(\tau) d\tau \quad 2.105$$

which has the solution  $h = \Theta$  a.e.

This proves the proposition made in the previous remark.

(d) It is important to mention that in all the above functional equations we have used  $t$  as the independent variable. In dynamical problems, such as control,  $t$  signifies time, but since similar kind of equations may also arise in problems of mechanics,  $t$  may be interpreted as required.

### EQUATIONS OF THE FIRST KIND:

In many control problems, it is often required to analyze the input data from the record of the output data which is available to the analyser.

Examples are, data telemetering, wind gust velocity prediction in aircraft control etc.

Let us consider the following system Fig. 2.9. in which the plant characteristics is known, and the output  $y(\cdot)$  as a function of time is being continuously recorded, at the station. The problem is to determine the input  $x(\cdot)$  that gave rise to the recorded output  $y(\cdot)$ .

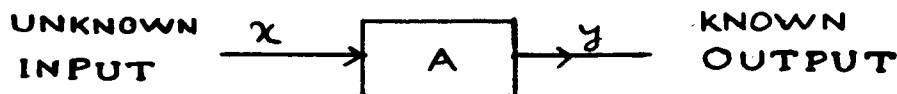


Fig 2.9: Data Telemetering:

The equation to be solved is,

$$y = Ax \tag{2.106}$$

Let us assume that  $A$  is a continuous functional operator possessing at least one continuous Gateaux derivative (see definition 3.1, chapter III) at every point  $x$  in some Banach space  $E$ .

Let us also assume that  $x_0(\cdot)$  and  $y_0(\cdot)$  are the steady state inputs and outputs respectively. It is observed over a given interval of time that the output  $y_0(\tau)$  has changed to  $y_0(\tau) + H(\tau)$ ,  $\tau \in I: (t_0 \leq \tau \leq t)$ . The problem is to determine the corresponding change in the input over the interval  $I$ . Let this undetermined change in the input state be denoted by  $x_0^h$ .

Then the equation 2.106 can be written as:

$$(y_0 + \beta H) = A(x_0 + \beta h) \text{ with } \beta \in [0, 1] \quad 2.107.$$

Similarly as before, the corresponding equation of variation is given by,

$$H(t) = \int_{t_0}^t \text{grad } A_{x_0} h(\xi) d\xi \quad 2.108$$

where,  $\text{grad } A_{x_0} \triangleq B(x_0 | t, \xi)$  is in general a nonlinear functional of  $x_0$  and an ordinary function of  $t$  and  $\xi$ . Therefore, the equation 2.108 can be written as

$$H(t) = \int_{t_0}^t B(x_0 | t, \xi) h(\xi) d\xi. \quad 2.108'$$

This is a linear volterra integral equation of the first kind which must be solved for  $h(\cdot)$ . This can be solved very conveniently if we assume that the function  $B(x_0 | t, \xi)$  is continuous in  $\xi$ .

$$\text{Let us put } g(t) = \int_{t_0}^t h(\xi) d\xi. \quad 2.109$$

Then equation 2.108' reduces to,

$$H(t) = \int_{t_0}^t B(x_0 | t, \xi) dg(\xi)$$

which on integrating by parts reduces to

$$H(t) = B(x_0 | t, t)g(t) - \int_{t_0}^t B_{\xi}^{(1)}(x_0 | t, \xi)g(\xi) d\xi, \quad 2.110$$

where,  $B_{\xi}^{(1)}(x_0 | t, \xi) = \frac{\partial B(x_0 | t, \xi)}{\partial \xi}$ .

If  $B(x_0 | t, t) \neq 0$  then the equation 2.110 can be written as,

$$\frac{H(t)}{B(x_0 | t, t)} = g(t) - \int_{t_0}^t \frac{B_{\xi}^{(1)}(x_0 | t, \xi)}{B(x_0 | t, t)} g(\xi) d\xi. \quad 2.111$$

This is a volterra integral equation of the second kind and by the usual method of solving such equations  $g(\cdot)$  can be determined. This gives the unknown function  $h(\cdot) = g^{(1)}(\cdot)$ .

It may turn out that  $B(x_0 | t, t) \equiv 0$ , in that case equation 2.110 reduces to,

$$H(t) = - \int_{t_0}^t B_{\xi}^{(1)}(x_0 | t, \xi) g(\xi) d\xi$$

which is again an equation of the first kind. By the same procedure as above, we can convert this equation to

$$\frac{H(t)}{B_{\xi}^{(1)}(x_0 | t, t)} = -G(t) + \int_{t_0}^t \frac{B_{\xi}^{(2)}(x_0 | t, \xi)}{B_{\xi}^{(1)}(x_0 | t, t)} G(\xi) d\xi. \quad 2.112$$

If  $B_{\xi}^{(1)}(x_0 | t, t) \neq 0$  then we have again an equation of the second kind which can be solved for  $G(\cdot)$  and obtain  $h(\cdot) = G^{(2)}(\cdot)$ . This may be repeated until we obtain an equation of the second kind. But it is important to notice that it requires that  $B(x_0 | t, \xi)$  be continuous in  $\xi$  and possess requisite number of derivatives with respect to  $\xi$ .

Another method of solving the equation 2.108' consists of taking the derivative on either side with respect to  $t$ . This requires that: (a)  $H(\cdot)$  be differentiable on  $t_0 \leq \tau \leq t$  and (b)  $B(x_0 | t, \xi)$  be differentiable with respect to the first variable  $t$  unlike the previous case in which it was required that  $B(x_0 | t, \xi)$  be differentiable with respect to  $\xi$ . If the system  $B$  changes continuously with time and if the function  $h$  is bounded then both the conditions (a) and (b) are

satisfied, and on taking the first derivative of  $H$ , we obtain,

$$H^{(1)}(t) = B(x_0 | t, t) h(t) + \int_{t_0}^t B_t^{(1)}(x_0 | t, \xi) h(\xi) d\xi. \quad 2.113$$

If  $B(x_0 | t, t) \neq 0$ , then equation 2.113 provides the equation of the second kind which we are seeking, and can be solved by the usual method; otherwise the procedure may be repeated until the desired equation of the second kind is obtained.

In some instances, the above method breaks down if there is a singularity in the kernel  $B(x_0 | t, \xi)$  at  $\xi = t$ . In that case we obtain a singular integral equation. Any discussion of this singular case requires the exact knowledge of the nature of singularity.

In control problems singularity may arise only if the system remains completely insensitive, over some interval of time, to any change in the input condition.

This is illustrated in the figure 2.10.

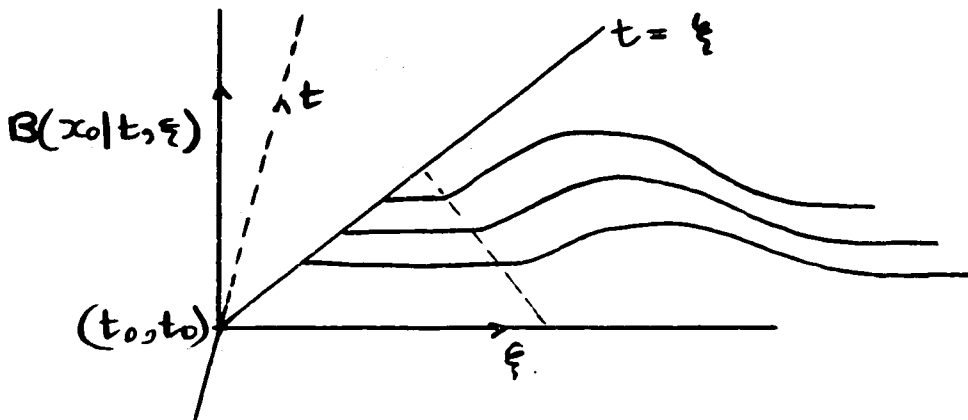


Fig. 2.10: A Plot of grad  $A_{x_0}$ .

such a system proves to be useless at least in data telemetering work but is not very uncommon in chemical process control. The system  $A$  can be anyone of the types  $A_U, A_H, A_L$  or  $A_V$  satisfying the appropriate continuity conditions.

Remarks:

(a) In sections 2.2-2 and 2.2-3 we treated control problems in which the terminal conditions are specified. This gives rise to functional equations in which the upper limit of integration is fixed. It was observed that the conditions that the functional must satisfy for the existence and uniqueness of solutions is quite restrictive. In contrast in sections 2.2-4 where causal systems are treated with the upper limits of integrations variable, the conditions are not so restrictive.

(b) By similar considerations, as in remark (c) following theorem 2.5, it can be shown that in the case of functionals with fixed upper limits, the equation 2.105 becomes

$$h(t) = \lambda \int_{a_0}^{b_0} R(t, \tau) h(\tau) d\tau \quad 2.114$$

which is a homogeneous Fredholm equation [7] and its solution  $h(\cdot) = 0$  if  $\lambda$  is not an eigen value of the kernel  $R$ . If however  $\lambda$  corresponds to an eigen value of the kernel  $R$  there may exist several solutions, the number of solutions depending on the index of the eigen value  $\lambda$ . Obviously for stability reasons the feedback loop gain  $\lambda$  must be chosen away from the spectral set of  $R$ .

Similarly the equation 2.108' becomes

$$H(t) = \int_{t_0}^{t_1} B(x_0 | t, \xi) h(\xi) d\xi \quad 2.115$$
$$t \in I: [t_0, t_1]$$

which is a Fredholm equation of the first kind. For a fixed  $x_0(\cdot)$ , we can very easily reduce this equation to a corresponding equation with symmetric kernel by simply multiplying both sides of equation 2.115 by  $B(x_0 | t, \tau)$  and then integrating the resulting function with respect to  $t$  over  $[t_0, t_1]$ .

This gives,

$$\int_{t_0}^{t_1} B(x_0 | t, \tau) H(t) dt = \int_{t_0}^{t_1} d\xi h(\xi) \int_{t_0}^{t_1} B(x_0 | t, \tau) B(x_0 | t, \xi) dt \quad 2.116$$

writing

$$\begin{aligned} H'(x_0 | \tau) &\triangleq \int_{t_0}^{t_1} B(x_0 | t, \tau) H(t) dt \quad \text{and} \\ L(x_0 | \tau, \xi) &\triangleq \int_{t_0}^{t_1} B(x_0 | t, \tau) B(x_0 | t, \xi) dt, \end{aligned} \quad 2.117$$

$\tau, \xi \in [t_0, t_1]$

we have,

$$H'(x_0 | t) = \int_{t_0}^{t_1} L(x_0 | t, \tau) h(\tau) d\tau. \quad 2.118$$

The kernel  $L$  is symmetric and even positive. If we assume that for every  $x_0 \in L^2$ ,  $L(x_0 | t, \tau) \in L^2(I \times I)$  and  $H'(x_0 | t) \in L^2(I)$  then we may conveniently apply Hilbert schmidt theorem [7] to solve for  $h(\cdot)$ . The necessary and sufficient condition for the existence of an unique  $L^2$  solution is that  $L$  be closed and that

$$\sum_i |\lambda_i(x_0) (H', \varphi_i)|^2 < \infty. \quad 2.119$$

Here,  $\{\lambda_i(x_0)\}$  and  $\{\varphi_i\}$  are the eigenvalues and the eigen functions of the kernel  $L$  and  $h$  is given by,

$$h = \text{l.i.m} \sum_{i=1}^n \lambda_i(x_0) (H', \varphi_i) \varphi_i. \quad 2.120$$

### CONCLUDING REMARKS:

In this chapter we were mainly concerned with the problem of analysis of nonlinear control systems. The class of systems considered are quite general, therefore the results obtained can be applied to a fairly large class of problems.

Though the class of systems considered is quite general, the work done is very little; this is mainly because we restricted

our operators to be acting on some  $L^p$  space (for some  $p$ ) and demanded the solution to belong to some  $L^p$  space (for some  $p$ ) also.

By extending the domain of the operators to some measure space and demanding measurable functions as the solution to equations of the form previously dealt with. we would have exploited the full potentiality of the operators defined.

For the existence of a solution, Schauder's principle [19] demands only that the operator be continuous and compact and that it transforms a closed, convex subset of a Banach space into the same subset.

The introduction of Orlicz spaces [20], which is a generalization of the classical Banach spaces, makes it possible to consider completely continuous operators acting on this larger class of spaces.

Krasnosel'skii and Rutickii [21] have proved the existence theorems for solutions of equations of Uryson and Hammerstein types:-

$$y(t) = x(t) + \lambda \int_I K(t, \tau; y(\tau)) d\tau. \quad 2.121$$

and 
$$y(t) = \lambda \int_I K(t, \tau) g(\tau; y(\tau)) d\tau. \quad 2.122$$

Their methods require the construction of suitable Orlicz spaces, for example in the case of the second equation it is required to construct two spaces  $L^M$  and  $L^N$  (where  $M$  is a convex function complementary to  $N$ ) such that  $g$  carries each function  $y \in L^N$  into one which belongs to  $L^M$  and then  $K$  carries the resulting function into one which belongs to  $L^N$ . If there exist two such

spaces satisfying the schauder's fixed point principle then the equation 2.122 has a solution in  $L^N$  for every value of  $\lambda$ . It is shown [ 21,22 ] by Krasnosel'skii and others that in the case of Hammerstein operators of type 2.122, even exponential nonlinearities may be treated if the problem is formulated in the frame work of Orlicz spaces.

In future it may be possible to prove similar theorems for the larger class of Uryson and Hammerstein operators introduced in this work.

In the following two chapters we will be mainly concerned with synthesis of Nonlinear control systems.

**CHAPTER III**

**"OPTIMUM SYNTHESIS OF NONLINEAR SYSTEMS"**

### 3.1: INTRODUCTION.

In the previous two chapters only problems of analysis were considered. The problem of synthesis, (which may be roughly described as designing a suitable system (operator) knowing the class of inputs and the corresponding class of outputs such that the overall performance is optimum in some sense), is always a difficult engineering problem. The degree of complexity of the problem depends upon the degree of accuracy of the performance demanded. For high degree of accuracy in performance, more complex systems must be taken into consideration.

We will consider the following two typical design problems of optimal control systems.

(i) The plant input and the desired output defined over a given interval of time  $I$  are specified in terms of either their stochastic characteristics or in terms of their higher order moments with respect to time. The class of Nonlinear operators to be considered is specified, (for example it may be mentioned which of the types  $A_U$ ,  $A_H$ ,  $A_L$  or  $A_V$  will be considered). The problem is to design the parameters of the control systems (for example the kernels of the operators) such that the difference between the desired output and the actual output is minimized in some sense.

We consider this problem for operators of type  $A_H$  and  $A_V$  in section 3.2.

(ii) The plant is fully specified (for example the functional representing the plant has known properties). The desired output of the plant is also specified.

The problem is to design a control signal with or without constraints such that the overall cost of control of the plant is minimized in some sense.

This problem is considered in section 3.3. The class of systems considered is quite general; it is only required that the functional representing the system possesses two continuous Gateaux derivatives at every point in the space in which it acts.

### 3.2: INTEGRAL EQUATIONS IN THE OPTIMUM SYNTHESIS OF A CLASS OF NONLINEAR TIME-VARIANT SYSTEMS.

In this section we consider the optimum synthesis of time-varying (including the time invariant system as the special case) nonlinear systems subject to an arbitrary nonstationary nonGaussian (including Gaussian as the special case) input process. In any optimization procedure, it is necessary to know a priori the class of systems on which the optimization is to be performed. We will restrict ourselves mainly to the two classes of operators  $A_H$  and  $A_v$  and denote them by  $A$ .

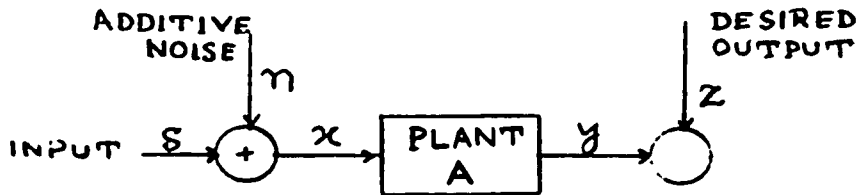


Fig. 3.1: Optimum Synthesis of Nonlinear Systems.

Let  $\{n(t), t \in I\}$  and  $s(t), t \in I$  be the nonstationary noise and signal processes respectively and  $\{x(\sigma, t)\}$  defined as  $x(\sigma, t) = n(\sigma, t) + s(t)$  be the input to the system  $A$ , Fig. 3.1. It will be assumed that  $x(\sigma, t)$  is Borel measurable in  $\sigma \in B_I$  and Lebesgue measurable in  $t \in M_I$ , where  $B_I$  is the  $\sigma$ -ring of Borel sets in the function space  $\Omega_I$  and  $M_I$  is the class of Lebesgue measurable sets on  $I$ . The interval  $I$  is the period during which the system operates on the element  $x(\sigma, t)$ .

Let  $y(\sigma, t)$  be the corresponding actual output and  $z(t)$  be the corresponding desired output.  $z(t)$  may be taken as some desired functional of the signal component  $s(t)$  of the total input. For example  $z(t)$  may be taken as  $s(t-\alpha)$ ,  $\alpha > 0$  in the case of prediction, or any other desired functional of  $s(t)$ . For convenience of discussion of the physical realizability of the optimum system, we will rewrite our operators  $A_H$  and  $A_V$  as,

$$A_H x = \sum_{s=0}^{N \ll \infty} \int_{I^s} \dots \int K_s(t; \tau_1, \dots, \tau_s) g(t-\tau_1, x(\sigma, t-\tau_1)) \dots g(t-\tau_s, x(\sigma, t-\tau_s)) \cdot d\tau_1 \dots d\tau_s$$

$t \in I: [0 T] \quad 3.1$

and,  $N \ll \infty$

$$A_V x = \sum_{s=0}^{N \ll \infty} \int_{I^s} \dots \int L_s(t; \tau_1, \dots, \tau_s) x(\sigma, t-\tau_1) \dots x(\sigma, t-\tau_s) d\tau_1 \dots d\tau_s$$

$t \in I: [0 T]. \quad 3.2$

where  $0 \leq N < \infty$  and  $K_s(t, \tau_1, \dots, \tau_s) \equiv 0$  for  $\tau_i < 0$  ( $i = 1, 2, \dots, s$ ) for all  $s \geq 1$ . This last condition is imposed by the fact that physical systems cannot respond to future inputs. This is known as the so-called causality relation. The interval  $I$  is assumed to be of finite Lebesgue measure.

It is assumed that the operator  $A$  is continuous and that it maps  $L^2(\Omega_I \times I, B_I \times M_I, \mu_B \cdot m)$  into  $L^2(\Omega_I \times I, B_I \times M_I, \mu_B \cdot m)$  where,  $\Omega_I \times I$  is the product space,  $B_I \times M_I$  is the product  $\sigma$ -field generated by the  $\sigma$ -ring  $B_I$  of Borel cylinders on  $\Omega_I$  and the  $\sigma$ -ring  $M_I$  of Lebesgue measurable sets on  $I$ .  $(\mu_B \cdot m)$  is the product measure with  $\mu_B$  the Borel measure on  $\Omega_I$  and  $m$  the Lebesgue measure on  $I$ .

The above conditions are satisfied by the operator  $A$  if we assume that  $x(\sigma, t) \in L^2(\Omega_I \times I, B_I \times M_I, \mu_B \cdot m)$  and  $L_s \in L^2(I^{s+1})$  for all  $0 \leq s \leq N < \infty$ , since according to measure theory a continuous function of a measurable function is a measurable function. In the case of operator  $A_H$  we need the additional condition that  $G$ -defined by,  $Gx = g(t; x(\sigma, t))$  be continuous in  $x$  and that  $G: L^2 \rightarrow L^2$ . Since the class of Volterra-Frechet operators is a subclass of the class of Hammerstein operators,

we consider the later class, and denote this class by  $H_N$ . Since the operator  $G$  affects the optimal choice of the Kernels of the Hammerstein operator, we will assume that  $G$  has been chosen to be the optimal zero memory operator. For the optimal choice of the Kernels, we have the following, theorem,

Theorem 3.1

For almost all  $t \in I$ , let  $x(\sigma, t) \in L^2(\Omega_I, B_I, \mu_B)$  be the input to the system  $A \in H_N$ . Let  $\{K_s\} \in L^2(I^{s+1})$  for all  $s \in J_N^+$  (the set of positive integers up to  $N$ ) and let  $y(\sigma, t) \in L^2(\Omega_I, B_I, \mu_B)$  be the output of the system  $A$ . Let  $A$  be assumed continuous in the sense that for every  $\epsilon > 0 \exists$  a  $\delta(\epsilon) > 0 \ni \langle x_1 - x_2, x_1 - x_2 \rangle_{\mu_B} < \delta(\epsilon)$  implies  $\langle Ax_1 - Ax_2, Ax_1 - Ax_2 \rangle_{\mu_B} < \epsilon$ . Let  $G$  be the fixed optimal zero memory operator continuous exactly in the same sense.

Then a necessary and sufficient condition that  $A \in H_N$ , be the optimum system is that the corresponding Kernels  $\{K_s\}$ , ( $s=0, 1, 2, \dots, N$ ) satisfy the following system of integral equations.

$$\phi_s(t, t-\theta_1, \dots, t-\theta_s) = \sum_{r=0}^N \int_{I^r} \dots \int \Psi_{s,r}(t-\theta_1, \dots, t-\theta_s; t-\tau_1 \dots t-\tau_r) K_r(t, \tau_1 \dots \tau_r) d\tau_1 \dots d\tau_r \quad 3.3$$

for almost all  $t \in I$ , and almost all  $(\theta_1 \dots \theta_s) \in I^s$ , ( $s = 1, 2, \dots, N$ ).

Proof:

Let us define the functional  $F^2(A, t) \geq 0$  by,

$$F^2(A, t) = \int_{\Omega_I} |Ax(\sigma, t) - z(t)|^2 d\mu_B, \quad t \in I. \quad 3.4$$

This expression gives the ensemble average of the square of the error as a functional of the operator  $A \in H_N$  and as an ordinary function of  $t \in I$ . We are interested to choose an element  $A$  from  $H_N$  such that  $F^2(A, t)$  attains its infimum. That is we are interested in the quantity,

$$\text{Inf } F^2(A, t)$$

$$A \in H_N \subset B_N$$

such that  $\forall B \in H_N$  and all  $\beta$  in  $F(R/C)$  (the field of real or complex numbers), and for almost all  $t \in I$ ,

$$F^2(A + \beta B, t) \wedge F^2(A, t) = F^2(A, t).$$

For abbreviation, we will write the equation 3.4 as

$$F^2(A, t) = \langle Ax-z, Ax-z \rangle_{\mu_B} \quad 3.4'$$

Let us assume that  $A \in H_N$  be the optimum system for any finite  $N$ .

Let  $G^2(D, t)$  be the error functional corresponding to the operator  $D \in H_N$  where  $D = A + \beta B$ , with  $B \in H_N$  and  $\beta$  is any real number.

Then,

$$G^2(D, t) = F^2(A + \beta B, t) \quad 3.5$$

$$= F^2(A, t) + \beta^2 \langle Bx, Bx \rangle_{\mu_B} - 2\beta \text{Re} \langle z-Ax, Bx \rangle_{\mu_B} \quad 3.6$$

If we assume our stochastic processes to be real valued functions and the Kernels corresponding to the operators, all real valued Lebesgue measurable Kernels, then the equation 3.6 becomes simply,

$$G^2(D, t) = F^2(A, t) + \beta^2 \langle Bx, Bx \rangle_{\mu_B} - 2\beta \langle z-Ax, Bx \rangle_{\mu_B} \quad 3.7$$

The last term in the equation 3.7 can be written explicitly as,

$$\langle z-Ax, Bx \rangle_{\mu_B} = \sum_{s=0}^N \int_{I_s} \int L_s(t; \theta_1 \dots \theta_s) [\phi_s(t; t-\theta_1, \dots, t-\theta_s) - \sum_{r=0}^N \int_{I_r} \int \psi_{s,r}(t-\theta_1, \dots, t-\theta_s; t-\tau_1, t-\tau_r) K_r(t; \tau_1 \dots \tau_r) d\tau_1 \dots d\tau_r] d\theta_1 \dots d\theta_s \quad 3.8$$

for: almost every  $t \in I$ , where,  $\{K_s\}$  are the Kernels corresponding to the optimum system  $A$  and  $\{L_s\}$  are the Kernels corresponding to the operator  $B$ .

The functions  $\phi_s$  and  $\psi_{s,r}$  are defined by,

$$\Phi_s(t; t-\theta_1 \dots t-\theta_s) = \langle z(t), g(t-\theta_1, x(t-\theta_1)) \dots g(t-\theta_s, x(t-\theta_s)) \rangle_{\mu_B} \quad 3.9$$

and

$$\Psi_{sr}(t-\theta_1 \dots t-\theta_s; t-\tau_1 \dots t-\tau_r) = \langle g(t-\theta_1, x(t-\theta_1)) \dots g(t-\tau_r, x(t-\tau_r)) \rangle_{\mu_B} \quad 3.10$$

The result of the theorem will follow if we can show that  $F^2(A, t)$  is minimum if, and only if  $\langle z-Ax, Bx \rangle_{\mu_B}$  (as in 3.8) is zero almost everywhere except on subsets of sets of Lebesgue measure zero; for any arbitrary choice of the operator  $B \in H_N$  (i.e the corresponding set of Kernels  $\{L_s\}$ ).

If  $\langle z-Ax, Bx \rangle_{\mu_B} = 0$  a.e. on  $I$ , then

$$G^2(D, t) = F^2(A, t) + \beta^2 \langle Bx, Bx \rangle_{\mu_B} \geq F^2(A, t) \quad \text{a.e on } I$$

for any  $\beta$ . Hence  $F^2(A, t)$  is minimum and  $A$  is the optimum system.

This proves the sufficiency condition of the thorem.

For the proof of the necessary condition let us note the following.

Let us rewrite 3.7 as,

$$G^2(D, t) = F^2(A, t) - 2\beta \left( \langle z-Ax, Bx \rangle_{\mu_B} - \frac{\beta}{2} \langle Bx, Bx \rangle_{\mu_B} \right) \quad 3.11$$

Let  $\langle z-Ax, Bx \rangle_{\mu_B} \neq 0$ , then, since  $B$  is an arbitrary element of  $H_N$ , we could choose the operator  $B$  ie the corresponding Kernels  $\{L_s\}$

such that  $\langle z-Ax, Bx \rangle_{\mu_B} > 0$ . Then since  $\langle Bx, Bx \rangle_{\mu_B} \geq 0$  and  $\beta$  is any arbitrary positive real number,  $\beta$  can be chosen so small that

the term inside the bracket becomes positive. This immediately leads

to the contradiction that  $F^2(A, t)$  is minimum. Hence for any arbitrary

$B \in H_N$  it is necessary and sufficient that,

$$\langle z-Ax, Bx \rangle_{\mu_B} = 0, \quad \forall B \in H_N. \quad 3.12$$

Since  $B$  is arbitrary,  $\{L_s\}$  are all arbitrary except that  $B \in H_N$  and  $L_s \in L^2(I^{s+1})$ . Therefore, setting  $L_s(t, \theta_1 \dots \theta_s)$  equal to the conjugate of the quantity inside the bracket (equation 3.8), we have,

$$\sum_{s=0}^N \int \dots \int_{I^s} |\phi_s(t; t-\theta_1, \dots, t-\theta_s) - \sum_{r=0}^N \int \dots \int_{I^r} \psi_{s,r}(t-\theta_1, \dots, t-\theta_s; t-\tau_1, \dots, t-\tau_r) \times \\ \times K_r(t, \tau_1, \dots, \tau_r) d\tau_1 \dots d\tau_r|^2 d\theta_1 \dots d\theta_s \\ = 0 \text{ a.e. on } I. \quad 3.13$$

for every  $N < \infty$ .

This implies that

$$\phi_s(t; t-\theta_1, \dots, t-\theta_s) = \sum_{r=0}^N \int \dots \int_{I^r} \psi_{s,r}(t-\theta_1, \dots, t-\theta_s; t-\tau_1, \dots, t-\tau_r) \\ K_r(t, \tau_1, \dots, \tau_r) d\tau_1 \dots d\tau_r \quad 3.14$$

for almost all  $t \in I$  and almost all  $(\theta_1, \dots, \theta_s) \in I^s$ , where,  $s = 0, 1, 2, \dots, N$ .

Q.E.D.

N.B. In the case of operators of type  $A_v$ , the equations 3.9 and 3.10 reduce to

$$\phi_s(t, t-\theta_1, \dots, t-\theta_s) = \langle z(t), x(\sigma, t-\theta_1), \dots, x(\sigma, t-\theta_s) \rangle_{\mu_B}$$

and

$$\psi_{s,r}(t-\theta_1, \dots, t-\theta_s; t-\tau_1, \dots, t-\tau_r) = \langle x(\sigma, t-\theta_1), \dots, x(\sigma, t-\theta_s), x(\sigma, t-\tau_1), \dots, x(\sigma, t-\tau_r) \rangle_{\mu_B}$$

respectively.

Corollary:

The minimum of the error functional  $F^2(A, t)$  resulting from the optimum choice of the operator  $A$ , whose corresponding kernels  $\{K_r\}$  satisfy the system of integral equations 3.14 is,  $F_m^2(A, t)$  where

$$F_m^2(A, t) = \langle z, z \rangle_{\mu_B} - \langle Ax, z \rangle_{\mu_B} \quad 3.15$$

$$= \xi_z(t, t) - \sum_{s=0}^N \int \dots \int_{I^s} \phi_s(t; t-\theta_1, \dots, t-\theta_s) K_s(t, \theta_1, \dots, \theta_s) \times \\ \times d\theta_1 \dots d\theta_s \quad 3.16$$

Proof:

$$\begin{aligned}
 F^2(A, t) &= \langle Ax-z, Ax-z \rangle_{\mu_B} \\
 &= \langle Ax-z, Ax \rangle_{\mu_B} + \langle z, z \rangle_{\mu_B} - \langle Ax, z \rangle_{\mu_B} \quad 3.16
 \end{aligned}$$

The necessary and sufficient condition for optimality was found to be,

$$\langle z-Ax, Bx \rangle_{\mu_B} = 0 \quad \forall B \in H_N \text{ for any finite } N.$$

Since  $A \in H_N$  the first term in equation 3.16 is zero for almost all  $t \in I$ . Hence it follows that if  $A$  is the optimum operator then

$$F_m^2(A, t) = \langle z, z \rangle_{\mu_B} - \langle Ax, z \rangle_{\mu_B} \quad \text{which is}$$

the desired result. Equation 3.16 follows from writing out all the terms explicitly and interchanging the order of integration on  $\Omega_I$  with respect to  $\mu_B$ , with those of Lebesgue integration, and denoting,

$$\langle z, z \rangle_{\mu_B} \text{ by } \xi_z(t, t)$$

$$\text{and } \langle z(t) \cdot x(\sigma t - \theta_1) \dots x(\sigma t - \theta_s) \rangle_{\mu_B} \text{ by}$$

$$\Phi_B(t; t - \theta_1, \dots, t - \theta_s).$$

The interchange of the order of integration is justified by Fubini's theorem [1, 2].

Q. E. D

Remarks:

The preceding theorem is quite general in the sense that it covers the problem of optimization of a large class of time variable nonlinear systems. The input is not required to be deterministic but can be some arbitrary nonstationary stochastic process.

It is clear that the volterra functional series does not admit delay functions in its Kernels even though the Kernels are assumed

to be Lebesgue measurable. On the other hand the Hammerstein functional series (proposed in this thesis) does admit delta functions of all possible orders through inclusion of the operator G. In addition the Kernels are also assumed Lebesgue measurable. This theorem utilizes both these classes of operators.

### Solution of Equations 3.14

In general it will be very difficult to solve the system of integral equations 3.14 by any available analytical method. However computer techniques may be employed. For this purpose, the method of steepest descent may be utilized to obtain an approximate solution. Before considering the steepest descent method, we will slightly modify our problem.

Instead of minimizing the functional  $F^2(A, t)$  equation 3.4', let us consider the quantity,

$$\inf_{A \in H_N} \int_I F^2(A, t) dm(t)$$

This is completely equivalent since if

$$F^2(A, t) = F^2(A + \epsilon B, t) \wedge F^2(A, t) \text{ then,}$$

$$\int_I F^2(A, t) dm(t) \leq \int_I F^2(A + \epsilon B, t) dm(t) \quad \forall B \in H_N$$

for every  $\epsilon \neq 0$ . The converse need not be true, since we may be interested only in minimizing,

$$\int_I F^2(A, t) dt = \int_I \langle Ax-z, Ax-z \rangle_{\mu_B} dm(t) = \langle Ax-z, Ax-z \rangle_{\mu} \quad 3.17$$

where  $\mu = \mu_B \times m$ , is the product of Borel and Lebesgue measure on the product space  $L^2(\Omega_I \times I, B_I \times M_I, \mu_B \times m)$ .  $M_I$  is the class of Lebesgue measurable sets on I, and m the Lebesgue measure.

Thus,

$$\begin{aligned} & \langle Ax-z, Ax-z \rangle_{\mu} \\ = & (\Psi K, K) - (\Phi, K) - (K, \Phi) + \langle z, z \rangle_{\mu} \end{aligned} \quad 3.18$$

where,

$$\left. \begin{aligned} \langle Ax, Ax \rangle_{\mu} &= (\Psi K, K) \\ \langle Ax, z \rangle_{\mu} &= (K, \phi) \\ \langle z, Ax \rangle_{\mu} &= (\phi, K) \end{aligned} \right\} \quad 3.19$$

and

$$(\Psi K, K) = \sum_{s=0}^{N<\infty} \sum_{r=0}^{N<\infty} \int_{I^{s+r+1}} \Psi_{s,r}(t-\theta_1, \dots, t-\theta_s; t-\tau_1, t-\tau_r) K_s(t; \theta_1 \dots \theta_s) K_r^*(t; \tau_1 \dots \tau_r) d\theta_1 \dots d\theta_s d\tau_1 \dots d\tau_r dt \quad 3.20$$

$$(K, \phi) = \sum_{s=0}^{N<\infty} \int_{I^{s+1}} \phi_s^*(t; t-\theta_1 \dots t-\theta_s) K_s(t; \theta_1 \dots \theta_s) d\theta_1 \dots d\theta_s dt. \quad 3.21$$

$$(\phi, K) = \sum_{s=0}^{N<\infty} \int_{I^{s+1}} \phi_s(t; t-\theta_1 \dots t-\theta_s) K_s^*(t; \theta_1 \dots \theta_s) d\theta_1 \dots d\theta_s dt. \quad 3.22$$

where \* means the complex conjugate.

It is clear that  $\Psi$  is a linear operator on a Hilbert space  $H$ , defined as the direct sum of  $N+1$  copies of the Hilbert spaces  $H_m$ ,

$$\text{ie } H = \sum_{m=1}^{N+1} \oplus H_m. \quad 3.23$$

The elements of  $H_m$  are Lebesgue measurable and square integrable functions defined on  $I^m$ . Hence the elements of  $H$  are  $N+1$  tuples.

An element  $K$  of  $H$  has the form,

$$K = [K_0(t), K_1(t, \tau_1), \dots, K_N(t, \tau_1 \dots \tau_N)] .$$

$\Psi$  maps an element  $K \in H$ , into an element  $L \in H$  such that,  $L = \Psi K$  and

$$L_s(t, \tau_1 \dots \tau_s) = \sum_{r=0}^N \int_{I^r} \Psi_{s,r}(t-\tau_1 \dots t-\tau_s; t-\theta_1 \dots t-\theta_r) K_r(t, \theta_1 \dots \theta_r) d\theta_1 \dots d\theta_r, \quad 3.24$$

with,  $(t, \tau_1 \dots \tau_s) \in I^{s+1}$ ,  $(s = 0, 1, 2, \dots, N)$ .

Similarly, the inner product on this space is defined by,

$$(L, K) = \sum_{s=0}^N \int_{I^{s+1}} \dots \int L_s(t, \theta_1, \dots, \theta_s) K_s^*(t, \theta_1, \dots, \theta_s) d\theta_1 \dots d\theta_s dt. \quad 3.25$$

In particular,

$$(K, K) = \sum_{s=0}^N \int_{I^{s+1}} \dots \int |K_s(t, \theta_1, \dots, \theta_s)|^2 d\theta_1 \dots d\theta_s dt. \quad 3.26$$

From 3.22, and 3.24 it is clear that  $\phi$  and  $\Psi$  satisfy, the following linearity properties,

$$(aK + bL, \phi) = a(K, \phi) + b(L, \phi) \quad 3.27$$

$$\Psi(aK + bL) = a\Psi K + b\Psi L \quad 3.28$$

To avoid notational difficulties we will assume that the stochastic processes under consideration are real valued functions and the kernels  $\{K_s\}$  are also real valued. This will imply that the Hilbert space  $H$  is real.

We are interested in minimizing the expression 3.18, over  $H_N$  and therefore over  $H$ . Since the last term in this expression is independent of  $K$ , and  $H$  is real we need to consider only,

$$Q(K) = (\Psi K, K) - 2(\phi, K). \quad 3.29$$

Extension to complex Hilbert space is not difficult. It is clear from the expression 3.19 that whether  $H$  is real or complex, the map  $\Psi$  is at least positive semi-definite and self adjoint.

$$\text{ie } (\Psi K, L) \geq 0 \quad \forall K, L \in H \quad 3.30$$

$$\text{and } (\Psi K, L) = (K, \Psi^* L) = (K, \Psi L) \quad 3.31$$

In this case, the operator is in fact positive definite.

That is,

$$(\Psi K, K) > 0 \quad \forall K \neq \theta \in H.$$

Let  $(\Psi K, K) = 0$  and  $K \neq \theta$  then it implies that,

$$(\Psi K, K) = \left\langle \left( \sum_r \int_{I^r} \dots \int K_r(t, \tau_1, \dots, \tau_r) g(t - \tau_1, x(\sigma, t - \tau_1)) \dots g(t - \tau_r, x(\sigma, t - \tau_r)) \right. \right. \\ \left. \left. d\tau_1 \dots d\tau_r \right)^2 \right\rangle_{\mu} = 0 \quad 3.32$$

If the stochastic process  $x(\sigma, t)$  is non-singular that is the measure  $\mu_B$  induced by the process on the function space  $\Omega_I$ , is a nondegenerate Borel measure then there exists an element  $\sigma_1 \in B_I \ni (g(t-\tau_1, x(\sigma_1, t-\tau_1)) \times \dots \times g(t-\tau_r, x(\sigma_1, t-\tau_r))) = K_r^*(t, \tau_1 \dots \tau_r)$  a.e with  $\mu_B(\sigma_1) > 0$ . In that case 3.32 becomes,

$$(\Psi K, K) = \mu_B(\sigma_1) \cdot \int_I \left\{ \sum_r \int_{I^r} |K_r(t, \tau_1 \dots \tau_r)|^2 d\tau_1 \dots d\tau_r \right\}^2 dt = 0$$

This implies that  $K = \theta$  a.e, which contradicts the assumption that  $K \neq \theta$ . Hence  $(\Psi K, K) > 0 \quad \forall K \neq \theta \in H$ .

The Hilbert space  $H$  on which  $\Psi$  is defined can be decomposed into the form  $H = H^1 \oplus H^0$  such that  $(H^1, H^0) = 0$  and an element  $L \in H$  has the representation  $L = L_1 + L_0$  where  $L_1 \in H^1$  and  $L_0 \in H^0$ .

The subspace  $H^0$  may be defined as the null space of the operator  $\Psi$  such that  $\Psi L = \theta \quad \forall L \in H^0$  and  $L \neq \theta$ .

with this preparation, we now consider the problem of minimization of the quadratic functional  $Q(K)$ . A geometrical interpretation of the method of steepest descent on the Hilbert space  $H$  can be given as follows.

Let us define a set  $A_r$  by

$$A_r = \{ K \in H \ni Q(K) \leq r \}$$

where  $r$  is some real number and  $A_r \subset H$ .

For each value of  $r \in \mathbb{R}^1$ ,  $\partial A_r$  defines the surface of a closed ellipsoid in  $H$ . Therefore  $\{A_{r_i}\}$ ,  $r_i \in \mathbb{R}^1$ , defines a family of ellipsoids with centre at the minimum point.

Let  $K_0 \in H^1$  be given as the initial approximation then  $K_0$  lies on the surface of some ellipsoid of the family  $\{A_{r_i}\}$ . The essential idea of the method lies in the fact that from the point  $K_0 \in \partial A_{r_0}$  one moves along the direction of the gradient at this point, ie along the normal to  $\partial A_{r_0}$  at  $K_0$  and reaches a point  $K_1$  where  $Q(K_1)$  is the least on this

normal, that is, a point  $K_1$  where the normal line is tangent to some ellipsoid of the family  $\{A_{r_i}\}$ . Let  $\partial A_{r_1}$  be the surface of the corresponding ellipsoid with  $K_1 \in \partial A_{r_1}$ . The next step is to move from this point along the normal to  $\partial A_{r_1}$ , at the point  $K_1$ . This process is repeated until the desired accuracy is reached.

It is important to note that solving the integral equation  $\Psi K = \phi$  is entirely equivalent to minimizing the functional,  $(\Psi K, K) - 2(\phi, K)$ . So the quadratic functional to be minimized is,

$$Q(K) = (\Psi K, K) - 2(\phi, K) \quad 3.33$$

Let us choose  $K_0 \notin H^0$  as the initial approximation and let  $L$  be any arbitrary element of  $H$  and  $\beta \in R^1$ .

Then,

$$Q(K_0 + \beta L) = Q(K_0) + 2\beta (\Psi K_0 - \phi, L) + \beta^2 (\Psi L, L), \quad 3.34$$

and the variation of the functional  $Q(K)$  at the point  $K = K_0$  and for any  $L \in H$  is

$$\left(\frac{dQ}{d\beta}\right)_{\beta=0} = 2(\Psi K_0 - \phi, L). \quad 3.35$$

By Schwarz inequality, the expression  $(\Psi K_0 - \phi, L)$  attains its maximum if  $L$  is chosen to be equal to  $\Psi K_0 - \phi$ , which may be called the gradient of the functional  $Q(K)$  at  $K = K_0$ .

Let us denote this  $L$  by  $L_0$  that is,

$$L = L_0 = \Psi K_0 - \phi.$$

with this choice for the value of  $L$ , equation 3.34 becomes,

$$f(\beta) = Q(K_0 + \beta L_0) = Q(K_0) + 2\beta (L_0, L_0) + \beta^2 (\Psi L_0, L_0).$$

Since  $\Psi$  is at most positive semidefinite, this quadratic polynomial in  $\beta$  will attain its minimum for some value of  $\beta$ . This value of  $\beta$  is obviously obtained from,

$$\frac{df}{d\beta} = 2(L_0, L_0) + \beta (\Psi L_0, L_0) = 0.$$

Therefore,

$$\beta = \beta_0 = -\frac{(L_0, L_0)}{(L_0, L_0)} = -\frac{\|L_0\|^2}{(\Psi L_0, L_0)}$$

with this choice of  $\beta_0$  equation 3.34 becomes,

$$Q(K_0 + \beta_0 L_0) = Q(K_0) - \frac{\|L_0\|^4}{(\Psi L_0, L_0)}$$

at the point  $K_1 = K_0 + \beta_0 L_0 \in H$ .

This is an improvement over the initial choice for  $K = K_0$ .

By repeating the above procedure, it can be shown that at the  $n$ th step we obtain,

$$\begin{aligned} \text{(i)} \quad L_{n-1} &= \Psi K_{n-1} - \phi \\ \text{(ii)} \quad \beta_{n-1} &= \frac{\|L_{n-1}\|^2}{(\Psi L_{n-1}, L_{n-1})} \quad h = 1 \quad 3.36 \end{aligned}$$

$$\text{(iii)} \quad K_n = K_{n-1} + \beta_{n-1} L_{n-1}$$

$$\begin{aligned} \text{and,} \quad Q(K_n) &= Q(K_{n-1}) - \frac{\|L_{n-1}\|^4}{(\Psi L_{n-1}, L_{n-1})} \\ &= Q(K_0) - \sum_{s=1}^n \frac{\|L_{s-1}\|^4}{(\Psi L_{s-1}, L_{s-1})} \quad 3.37 \end{aligned}$$

Equation, 3.37 shows, that at each step  $Q$  decreases steadily. If at certain stage (say  $n$ th step) the sequence terminates ie

$$L_m = 0 \text{ then, } \Psi K_m = \phi \text{ and } K_m \text{ is given by } K_m = K_{m-1} + \beta_{m-1} L_{m-1}.$$

such an ideal situation can be hardly expected and hence the question of convergence of the sequence  $\{Q(K_n)\}$  needs special attention.

Instead of considering the convergence of  $Q(K_n)$  let us consider the convergence of an equivalent quantity,  $(\Psi w_n, w_n)$  where  $w_n$  is the error vector defined as follows. If  $\bar{K} \in H$ , be the exact solution of the equation  $\Psi K = \phi$ , then the error vector at the  $n$ th step may be defined as,

$$w_n = K_n - \bar{K} \text{ with } w_n \in H. \quad 3.38$$

Thus, by 3.36 and 3.38, we obtain

$$\left. \begin{aligned} w_n &= \Psi K_n - \Psi \bar{K} = \Psi K_n - \phi = L_n, \\ \text{and } w_n &= w_{n-1} + \beta_{n-1} L_{n-1}. \end{aligned} \right\} \quad 3.39$$

Therefore,

$$\begin{aligned} (\Psi w_n, w_n) &= (\Psi w_{n-1} + \beta_{n-1} \Psi L_{n-1}, w_{n-1} + \beta_{n-1} L_{n-1}) \\ &= (\Psi w_{n-1}, w_{n-1}) \left\{ 1 - \frac{(L_{n-1}, L_{n-1})^2}{(\Psi L_{n-1}, L_{n-1})(\Psi w_{n-1}, w_{n-1})} \right\} \end{aligned} \quad 3.40$$

A theorem due to Greub and Rheinboldt [9] which is a generalization of a corresponding theorem proved by Kantorovich [10] states that if  $\Psi$  is a self adjoint linear operator on the Hilbert space  $H$  and if  $\Psi$  satisfies the condition,

$$d(K, K) \leq (\Psi K, K) \leq D(K, K) \text{ and } 0 < d$$

then,

$$(\Psi K, K)(\Psi^{-1} K, K) \leq \frac{(D+d)^2}{4Dd} (K, K)^2, \quad \forall K \in H. \quad 3.41$$

In our case,  $\Psi$  is a real and symmetric linear integral operator on  $H$  and further if we assume that  $\Psi$  is positive definite then all the hypothesis of the above theorem are satisfied and we can use the inequality 3.41 in equation 3.40 and obtain,

$$\begin{aligned} (\Psi w_n, w_n) &\leq (\Psi w_{n-1}, w_{n-1}) \left( \frac{D-d}{D+d} \right)^2 \\ &\leq (\Psi w_0, w_0) \left( \frac{D-d}{D+d} \right)^{2n}. \end{aligned}$$

since  $\left( \frac{D-d}{D+d} \right) < 1$ , this shows that  $(\Psi w_n, w_n)$  and therefore  $Q(K_n) - Q(\bar{K})$ , converges to zero with the geometric progression. In fact,

since  $\Psi$  is positive definite, the error vector  $w_n$  converges strongly to the null vector in  $H$ , since  $\|w_n\| \leq \sqrt{\frac{D}{d}} \left( \frac{D-d}{D+d} \right)^n \|w_0\|$ . 3.42

This in turn implies that the sequence  $\{K_n\}$  constructed by the method of steepest descent converges strongly to the optimal solution  $\bar{K}$ .

But if the operator  $\Psi$  is only positive semidefinite then the sequence  $K_n$  converges weakly i.e.  $\|\Psi K_n - \phi\| \rightarrow 0$  and  $\lim_n Q(K_n) = \text{Inf } Q(K)$ .

Recently an elegant proof of the weak convergence of a sequence constructed by the method of steepest descent has been given by Balakrishnan [11]. The weak convergence ensures that an approximate solution can always be computed even when there is no element  $K \in H$  that satisfies the equation  $\Psi K = \phi$ .

Remarks: (a) Theorem 3.1 is quite general in the sense that (i) optimization is considered over the whole class  $H_N$  of nonlinear time variant systems. (ii) the input to the plant can be an arbitrary nonstationary stochastic process and (iii) the interval  $I$  over which the stochastic process is defined is arbitrary.

The principal limitation is that it requires the complete knowledge of stochastic properties, in terms of multiple distributions, of the input process. In the case of time invariant nonlinear systems the necessity for multiple distributions may be replaced by that of multiple time averages provided the additional assumptions of stationarity and ergodicity is satisfied.

(b) Another possible method for solving the equation  $\Psi K = \phi$  is the Hilbert-schmidt method. For let  $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3| \dots$  be the eigenvalues arranged in the nondecreasing order and  $u_1, u_2, \dots, u_n, \dots$  the corresponding eigen functions belonging to the real symmetric Kernel  $\Psi$ , such that  $\bar{u}_s = \lambda_s \Psi \bar{u}_s$ .

If both  $\Psi$  and  $\phi$  belong to their appropriate  $L^2$  spaces and if  $\Psi K = \phi$  has a solution  $K$  of class  $L^2$  then by Hilbert-schmidt theorem [7]

we have,

$$\phi = \sum_{s=1}^{\infty} \frac{(K, u_s)}{\lambda_s} u_s$$

and the series on the right hand side converges almost uniformly to the given function  $\phi$ . If the series,  $\sum_{s=1}^{\infty} |\lambda_s (\phi, u_s)|^2$  converges then there exists an  $L^2$  function  $K_0$  to which the sequence  $K_n$  defined as,

$$K_n = \sum_{s=1}^n \lambda_s (\phi, u_s) u_s$$

converges in the limit in the mean. That is

$$K_0 = \text{l.i.m.}_n K_n$$

This follows from Riesz-Fischer theorem.

If further  $H^0$  is empty that is if  $\Psi$  is closed then  $K_0$  is the unique  $L^2$  solution of the equation  $\Psi K = \phi$ .

This method requires the knowledge of the eigenvalues and eigen functions of the operator which is not always easy to evaluate. However there exists [7] definite variational methods for their computations.

On the other hand steepest descent method requires only the evaluation of multiple integrals which can be performed by computers.

(b) It is important to note that in equation 3.23, we defined  $H$  as the direct sum of a finite number of Hilbert spaces, but in fact there is no theoretical limitation on the consideration of a countable number of spaces giving  $H = \sum_{m=1}^{\infty} \oplus H_m$ , so long,

$$\|K_0\|^2 + \sum_{n=1}^{\infty} \int_{I_{n+1}} \dots \int |K_n(t\tau_1 \dots \tau_n)|^2 d\tau_1 \dots d\tau_n dt < \infty .$$

### 3.3 FORMULATION OF AN OPTIMAL CONTROL PROBLEM AND ITS SOLUTION.

In the consideration of the second problem as stated in the introduction, the following definitions and remarks will be useful.

The operator  $A$  appearing in the following context may be

any one of those defined in chapter, I. Since the following work applies to even more general functional operators,  $A$  may be interpreted as  $Ax = F[x(\cdot), t]$ . For compactness  $F[x(\cdot), t]$  will often be written as simply  $Fx$ .

Definition 3.1: G-Derivative.

The operator  $A$  acting from the Banach space  $E_1$  into the Banach space  $E_2$  is said to possess a Gateaux derivative at the point  $x \in E_1$  if there exists a linear operator  $B$  from  $E_1$  into  $E_2$  such that

$$\lim_{\mu \rightarrow 0} \left\| \frac{A(x+\mu h) - A(x) - \mu B(h)}{\mu} \right\| = 0 \quad 3.43$$

$\forall h \in E_1$ , where  $\mu$  is a real number.

If  $B$  exists, it will be called the G-derivative of  $A$  at  $x$  and will be denoted by

$$B = A'_x = \text{grad } A_x$$

Definition 3.2: F-Derivative.

If the operator  $B$  satisfies the more restrictive condition that

$$\lim_{\|h\| \rightarrow 0} \frac{\|A(x+h) - A(x) - B(h)\|}{\|h\|} = 0 \quad 3.44$$

then  $A$  is said to be Frechet differentiable at  $x \in E_1$  and will also be denoted by

$$B = A'_x = \text{grad } A_x \quad 3.45$$

Remark

If  $Ax$  is F-differentiable at  $x_0 \in E_1$  then its G-derivative also exists at  $x_0 \in E_1$  and equals its F-derivative.

On the other hand, if  $Ax$  is G-differentiable at  $x_0 \in E_1$  and if  $A'_x$ , the G-derivative of  $Ax$  at  $x_0$ , is continuous at  $x_0$ , then its F-derivative exists and equals the G-derivative.

An example of an operator possessing both the G- and F-derivatives at any point  $x \in E_1$  is the operator of type  $A_V$  (Volterra Frechet operator).

The operator  $A$  corresponding to the functional

$$Ax = F [x(\cdot); t] = \sum_{s=0}^{\infty} \int_I \dots \int_I K_s(t; \tau_1 \dots \tau_s) x(\tau_1) \dots x(\tau_s) d\tau_1 \dots d\tau_s \quad 3.46$$

has the derivative  $A'_{x_0}$  at  $x_0 \in E$ , where  $A'_{x_0}$  is given as

$$A'_{x_0} = \sum_{s=1}^{\infty} \int_I \dots \int_I K_s(t, \xi, \tau_1, \dots, \tau_{s-1}) x_0(\tau_1) \dots x_0(\tau_{s-1}) d\tau_1 \dots d\tau_{s-1} \quad 3.47$$

$$\triangleq \bigwedge_V(t, \xi | x_0).$$

And for any  $h \in E_1$ ,  $A'_{x_0} h$  is a linear functional of  $h$ .

$$A'_{x_0} h = \int_I \bigwedge_V(t, \xi | x_0) h(\xi) d\xi. \quad 3.48$$

In deriving the equation 3.47 it is assumed that the set of Kernels  $\{K_s\}$  is symmetric in all its variables. This is not a restriction since the Kernels can be symmetrized.

Similarly, we can define the G and F-derivatives of the operators of type  $A_U$  and  $A_H$  (equations 1.21 and 1.22) under certain simplifying assumptions.

In the case of operator of type  $A_H$ , we assume that  $g(t, u)$  is measurable in  $t$  and continuous in  $u$  for all  $t \in I$  and  $|u| \leq r < \infty$  where  $r$  is some real number.

In this case the gradient of the operator  $A_H$  at the point  $x_0 \in T \rho \subset E_1$  is given by,

$$\text{grad } A_{x_0} = \sum_{s=1}^N \int_I \dots \int_I P_{\xi}^s K_s(t, \xi, \tau_1 \dots \tau_{s-1}) g^1(\xi, x_0(\xi)) g(\tau_1, x_0(\tau_1)) \dots$$

$$\dots g(\tau_{s-1}, x_0(\tau_{s-1})) d\tau_1 \dots d\tau_{s-1}$$

$$(t, \xi) \in I^2 \quad 3.49$$

where  $P_{\xi}^s$  is a cyclic permutation of the variables  $(\tau_1 \dots \tau_s)$  of order  $s$  with respect to the symbol  $\xi$  and  $g^1(t, u) = \frac{\partial K}{\partial u}(t, u)$ . In case the Kernels are symmetric

$$\text{grad } A_{x_0} = \sum_{s=1}^N \int_{I^{s-1}} \dots \int_{I^{s-1}} K_s(t, \xi, \tau_1 \dots \tau_{s-1}) g^1(\xi, x_0(\xi)) \dots g(\tau_{s-1}, x_0(\tau_{s-1})) d\tau_1 \dots d\tau_{s-1}. \quad 3.50$$

Denoting  $\text{grad } A_{x_0}$  by  $\bigwedge_H(t, \xi | x_0(\cdot))$ , the first derivative of the functional corresponding to the Hammerstein operator  $A_H$  at the point  $x_0 \in T_p$  is given by,

$$(\text{grad } A_{x_0}, h) = \int_I \bigwedge_H(t, \xi | x_0(\cdot)) h(\xi) d\xi \quad 3.51$$

for each  $x_0 \in T_p$  and  $\forall h \in E_1$

similarly in the case of operators of type  $A_U$ , the gradient of  $A_U$  is given by,

$$\text{grad } A_{x_0} = \sum_{s=1}^N \int_{I^{s-1}} \dots \int_{I^{s-1}} P_{\xi}^s U_s^{(1)}(t; \xi, \tau_1 \dots \tau_{s-1}; x_0(\xi), x_0(\tau_1) \dots x_0(\tau_{s-1})) d\tau_1 \dots d\tau_{s-1} \quad 3.52$$

for almost all  $(t, \xi) \in I^2$ .

$P_{\xi}^s$  is as defined before and  $U_s^{(1)}$  is defined as  $U_s^{(1)}(t, \tau_1 \dots \tau_s, u_1 \dots u_s) =$

$\frac{\partial U_s}{\partial u_1}$ , under the obvious assumption that each of the functions  $U_s$  are continuous in the variable  $(u_1 \dots u_s)$  in a suitable  $s$ -dimensional

Euclidian space, and the derivatives are integrable. The first derivative of the functional corresponding to the Uryson operator is given by,

$$(\text{grad } A_{x_0}, h) = \int_I \bigwedge_U(t, \xi | x_0(\cdot)) h(\xi) d\xi \quad 3.53$$

for each  $x_0 \in T_p$  and  $\forall h \in E_1$ .

Since most of our subsequent discussions are valid for all the three classes of operators we will omit the subscripts of the functional  $\bigwedge$ .

Another important example of a functional which has a derivative in both the Gateaux and Frechet sense at each  $x \in L^P$  is the functional  $\phi(x)$  defined as,

$$\phi(x) = \|x\|_p = \left( \int_I |x(t)|^p d\mu \right)^{1/p} \quad \forall x \in L^P. \quad 3.54$$

It can be easily shown that  $\text{grad } \phi_x$  at the point  $x \in L^P$  is,

$$\text{grad } \phi_x = (\|x\|_p)^{1-p} |x|^{p/q} \text{ sign } x, \quad (p, q > 1, \frac{1}{p} + \frac{1}{q} = 1) \quad 3.55$$

and the corresponding differential of the functional  $\phi$  at an arbitrary point  $x_0 \in L^P$  for any  $h$  is,

$$\begin{aligned} \phi^{(1)}(x_0, h) &= (\text{grad } \phi_{x_0}, h) \\ &= (\|x_0\|_p)^{1-p} (|x_0|^{p/q} \text{ sign } x_0, h). \end{aligned} \quad 3.56$$

Clearly this is also a linear functional in  $h$ .

In the case of  $L^2$  space,

$$\phi^{(1)}(x, h) = (\|x\|_2)^{-1} (x, h) \quad 3.57$$

In the theory of nonlinear analysis another important concept which is often encountered is concerned with potential operators.

Definition 3.3: Potential operators.

A real differentiable functional  $Gx$  with domain  $D(G) \subseteq E_1$  is called the potential of an operator  $Fx$ , also acting in  $E_1$ , if  $Fx$  is the gradient of  $Gx$ .

Definition 3.4: Critical points of functionals.

A point  $x_0 \in E_1$  at which  $\text{grad } Gx_0 = G^1 x_0 = 0$  is called a critical point of the functional  $Gx$ .

Definition 3.5: Relative extremal points of functionals.

An interior point  $x_0$  of a set  $T \subset E_1$ , such that  $\forall x \in T, Gx_0 \geq Gx$  is called a relative maximum. Similarly if  $Gx_0 \leq Gx \quad \forall x \in T$  then the point

$x_0 \in T$  is called the relative minimum point of  $Gx$  in  $T$ . Relative maximum and minimum points of  $Gx$  are called relative extremal points of  $Gx$ .

Theorem 3.2

A relative extremal point  $x_0 \in T$  of a differentiable functional  $Gx$  is the critical point of  $Gx$ .

Proof:

Let  $g(\mu) = G(x_0 + \mu h)$ ,  $h \in E_1$ ,  $\mu \in \mathbb{R}^1$  then it follows immediately that the real function  $g(\mu)$  has the extremal point  $\mu=0$ .

$$\left( \frac{dg}{d\mu} \right)_{\mu=0} = G'_{x_0} h = 0 \quad \forall h \in E_1$$

Hence  $G'_{x_0} = G'(x_0) = 0$ . This proves the theorem.

For a functional to possess a critical point it is not even necessary that it be weakly continuous (see definition 1.9). It is sufficient if it is only weakly lower or weakly upper semicontinuous. The following definitions will make this fact clear.

Definition: 3.6

A subset  $T$  in a Banach space  $E$  is said to be weakly compact if every infinite sequence of elements of  $T$  has a weakly convergent subsequence, the limit of which may not belong to  $T$ .

It is known that all reflexive Banach spaces [ 5 ] and thus all Hilbert spaces have weakly compact spheres.

Definition 3.7: Weakly lower/upper semicontinuity.

A functional  $Gx$  is said to be weakly lower semicontinuous at  $x_0 \in E$  if  $Gx_0$  does not exceed the limit inferior of the sequence

$Gx_n$ , i.e

$$Gx_0 \leq \bigvee_{n=1}^{\infty} \bigwedge_{k \geq n} Gx_k$$

for all sequences  $\{x_k\}$  that converge weakly to  $x_0$ .

Similarly the functional  $Gx$  is said to be weakly upper semicontinuous if the following condition is satisfied for all sequences  $\{x_k\}$  converging weakly to  $x_0$ :

$$Gx_0 \geq \bigwedge_{n=1}^{\infty} \bigvee_{k \geq n} Gx_k \quad 3.58'$$

A theorem due to Vainberg [12] which will be used subsequently is quoted below:

Theorem 3.3

A weakly lower semicontinuous functional  $Gx$  is bounded below and attains its lower bound on every bounded and weakly compact subset of the Banach space  $E$ .

Since we consider only reflexive Banach space  $E$ , all bounded subsets of  $E$  are weakly compact.

Proof: See Vainberg [12]

Definition 3.8 m-property [12]

A weakly lower semicontinuous functional  $Gx$  is said to have the  $m$ -property on a bounded, and closed subset  $T$  of  $E$  if there exists a point  $x_0 \in T$  such that for all  $x$  on the boundary of  $T$  denoted by  $\partial T$

$$Gx_0 < Gx, \quad \forall x \in \partial T \quad 3.59$$

Theorem 3.4

A Gateaux differentiable functional  $Gx$  with the  $m$ -property has at least one critical point.

Proof:

The proof follows from theorem 3.3 and the definition 3.8. By theorem 3.3  $Gx$  attains its lower bound on the set  $T$  for which 3.59 holds.

If now  $Gx$  is assumed to possess the  $m$  property on  $T$ , then it attains its relative minimum at the point  $x_0 \in T$ .

Q. E. D

Now we can consider the problem (ii) as stated in the introduction to this chapter. To be specific, let us consider the following control system, Fig. 3.2 where,

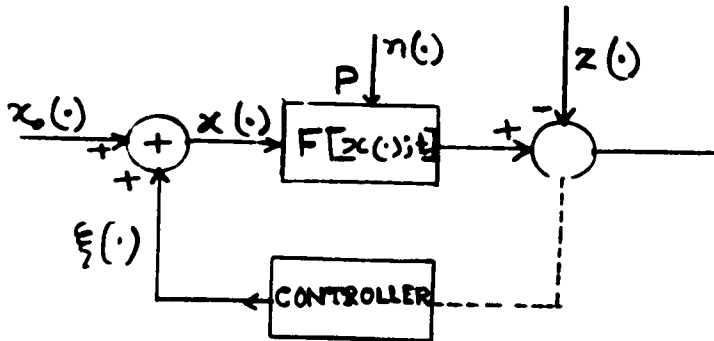


Fig: 3.2: Optimum Synthesis of Control Function.

$P$  is the plant described by the functional operator  $Fx = F[x(\cdot), t]$ ,  $x(t) \in L^2(I)$  is the control signal defined over the interval  $I$ , the period of operation of the system. The set  $I$  may be finite or infinite, depending upon the memory of the system. The function  $z(t) \in L^2(I)$  is the desired output also defined on  $I$ . The actual input to the plant is the sum of a steady input  $x_0(\cdot)$  and the control  $\xi(\cdot)$  determined by the performance of the system. Let  $\lambda > 0$  be the cost per unit of energy delivered to the input of the plant and  $\mu > 0$  the cost per square unit of the norm of error.

Let us denote by  $Gx = G[x(\cdot)]$  the total cost of control of the plant  $P$ .

$$G[x(\cdot)] = \lambda \int_I [x(t)]^2 dt + \mu \int_I [F[x(\cdot), t] - z(t)]^2 dt \quad 3.60$$

Thus the problem of optimizing the control system is reduced to minimizing the functional  $Gx$  with respect to  $x$ . The control input  $\xi(\cdot) = x(\cdot) - x_0(\cdot)$ .

Theorem 3.5

The necessary condition for the optimality of the control system  $P$  described by the functional operator  $F[x(\cdot); t], t \in I$ , and the cost functional  $G[x(\cdot)]$ , is that  $F$  possesses the Gateaux derivative at the point  $x \in E$  and satisfies the functional equation,

$$x(t) = \eta \int_I \wedge(t, \xi | x(\cdot)) [z(\xi) - F[x(\cdot); \xi]] d\xi \quad 3.61$$

$t \in I$

Proof:

Let us compute the first variation of the functional  $G$  about the point  $x \in E$  (here  $L^2(I)$ ). Let  $\beta$  be any real number and  $h$  any element in  $E$ . From equation 3.60 we obtain the equation 3.62.

$$G[x + \beta h] = \lambda \int_I [x(t) + \beta h(t)]^2 dt + \mu \int_I [F[x + \beta h(\cdot), t] - z(t)]^2 dt \quad 3.62$$

$$\frac{dG}{d\beta} = 2\lambda \int_I [x(t) + \beta h(t)] h(t) dt + 2\mu \int_I \frac{dF}{d\beta} [F[x + \beta h(\cdot), t] - z(t)] dt \quad 3.63$$

substituting  $\beta = 0$  in equation 3.63 and equating  $\left(\frac{dG}{d\beta}\right)_{\beta=0} = 0$  we obtain the following condition.

$$0 = \lambda \int_I x(t)h(t)dt + \mu \int_I \int_I \text{grad } F_x (F[x(\cdot), t] - z(t)) h(\xi) dt d\xi \quad 3.64$$

where  $\text{grad } F_x$  is a functional of  $x$  and a function of  $t$  and  $\xi$ , which is denoted by  $\wedge(t, \xi | x)$ .

$$\text{grad } F_x = \wedge(t, \xi | x(\cdot)) \quad 3.65$$

Substituting the expression 3.65 in equation 3.64 we obtain

$$0 = \lambda \int_I x(t)h(t)dt + \mu \int_I \int_I \Lambda(t, \xi | x(\cdot)) [F[x(\cdot), t] - z(t)] h(\xi) dt d\xi, \quad \forall h \in E. \quad 3.66$$

This can be written as,

$$0 = \int_I d\xi h(\xi) \left[ \lambda x(\xi) - \mu \int_I \Lambda(t, \xi | x(\cdot)) (z(t) - F[x(\cdot), t]) dt \right]. \quad 3.67$$

For 3.67 to hold for any  $h \in E$  and otherwise arbitrary it is necessary and sufficient that

$$x(\xi) = \frac{\mu}{\lambda} \int_I \Lambda(t, \xi | x(\cdot)) [z(t) - F[x(\cdot); t]] dt \quad \xi \in I \quad 3.68$$

Letting  $\frac{\mu}{\lambda} = \eta$  and assuming  $\Lambda(t, \xi | x)$  to be symmetric in  $t$  and  $\xi$  we obtain the necessary condition

$$x(t) = \eta \int_I \Lambda(t, \xi | x(\cdot)) [z(\xi) - F[x(\cdot); \xi]] d\xi \quad \forall t \in I \quad 3.69$$

Q. E. D

as stated in the theorem.

**Remark** Equation 3.69 is in general a nonlinear integral equation, which must be satisfied by the control signal  $x(\cdot)$ , so that the cost functional may attain an extremum.

**Theorem 3.6**

Let us assume that the functional  $Gx$  (equation 3.60) is Gateaux-differentiable twice at  $x \in E$ . Then the sufficient condition for the element  $x \in E$  to be the optimal control signal is that  $G$  satisfies the following inequality

$$G''_x \varphi \varphi \geq \beta \|\varphi\|^2 \quad \forall \varphi \in E \quad 3.70$$

where  $g(\beta)$  is continuous and non-negative for  $\beta > 0$  and tends to infinity as  $\beta$  tends to infinity.

Proof:

It can be shown [16] that if  $Hx$  is the gradient, it has the unique potential say  $Gx$  which is given by the following relation

$$Gx = Gx_0 + \int_0^1 (H\{x_0 + \theta(x-x_0)\}, x-x_0) d\theta \quad 3.71a$$

which assumes the value  $Gx_0$  at  $x = x_0$ .

The quantity inside the bracket under the integral sign represents the inner product on  $E$  between  $H$  and  $(x-x_0)$ . In case  $x_0 = 0$ , equation 3.71a reduces to,  $Gx = G(o) + \int_0^1 (H(\theta x), x) d\theta$  . 371b

If  $Gx$  is assumed to be  $G$ -differentiable twice, then it follows from 3.71 that

$$Hx = H(o) + \int_0^1 (H'(\theta x), x) d\theta \quad 3.72$$

ie

$$G^t x = G^t(o) + \int_0^1 (G''(\theta x), x) d\theta \quad 3.73$$

Computing the inner product of  $G^t x$  with  $x$  in  $E$ , we have from 3.73,

$$(G^t x, x) = (G^t(o), x) + \int_0^1 (G''(\theta x) x, x) d\theta \quad 3.74$$

By 3.70 and 3.74 we obtain the following inequality,

$$\begin{aligned} (G^t x, x) &\geq (G^t(o), x) + \int_0^1 \|x\| g(\|x\|) d\theta = \\ &= (G^t(o), x) + \|x\| g(\|x\|) \end{aligned} \quad 3.75$$

which is valid for all  $x \in E$ .

By substituting the expression  $(G^t x, x)$  from 3.75 into equation 3.71b we obtain

$$Gx \geq G(o) + \int_0^1 \{ (G^t(o), \theta x) + \|\theta x\| g(\|\theta x\|) \} \frac{d\theta}{\theta} \quad 3.76$$

$$= G(o) + (G^t(o), x) + \|x\| \int_0^1 g(\theta \|x\|) d\theta \quad 3.77$$

but  $|(G^t(o), x)| \leq \|G^t(o)\| \|x\|$

$$\text{Hence, } (G'(o), x) \geq - \|G'(o)\| \|x\| \quad 3.78$$

Combination of 3.77 and 3.78 yields the following inequality,

$$Gx \geq G(o) + \|x\| \left\{ \int_0^1 g(\theta \|x\|) d\theta - \|G'(o)\| \right\} \quad 3.79$$

This is valid for any  $x$  in the ball  $T_\rho$  of radius  $\rho$  in the space  $E$ .

It is clear from the inequality 3.79 that for sufficiently large  $\rho$ , the quantity in the braces will be positive for all  $x$  on the boundary of the set  $T_\rho$ .

That is

$$Gx > G(o) \quad \forall x \in \partial T_\rho \quad 3.80$$

Thus the functional  $Gx$  satisfies the  $m$ -property (definition 3.8) since the set  $T_\rho$  is bounded and weakly closed.

Hence  $Gx$  satisfies the hypothesis of theorem 3.4, therefore by theorem 3.4 there exists at least one critical point  $x_0 \in T_\rho$ , which is the relative minimum.

Q. E. D

### Theorem 3.7

For the cost functional  $Gx$  to satisfy the condition as stated in the theorem 3.6, it is sufficient that the system described by the functional operator  $Fx = F[x(\cdot), t]$  satisfies the following condition:

$$(F_x'' \phi \phi, Fx - z) + \|F_x' \phi\|^2 \geq \left( \frac{b-2\lambda}{2\mu} \right) \|\phi\|^2 \quad 3.81$$

for every given  $z \in E$ , and for all  $\phi \in E$  and some real number  $b > 0$ .  $\mu$  and  $\lambda$  are positive real numbers as in theorem 3.5.

Proof:

We can write 3.60 as;

$$G_x = \lambda(x, x) + \mu(Fx-z, Fx-z). \quad 3.82$$

It is clear that for  $G_x$  to be  $G$ -differentiable twice at  $x \in E$  it is necessary that  $F_x$  be also  $G$ -differentiable twice at  $x$ . Since we are concerned with real functionals, the derivative of the functional  $G_x$  taken twice at  $x \in E$  gives,

$$G_x'' = 2\lambda + 2\mu(F_x', F_x') + 2\mu(F_x'', Fx-z) \quad 3.83$$

Thus,

$$G_x'' \phi \phi = 2\lambda \|\phi\|^2 + 2\mu \|F_x' \phi\|^2 + 2\mu (F_x'' \phi \phi, Fx-z) \quad 3.84$$

Substituting the inequality 3.81 in the equation 3.84 we obtain,

$$G_x'' \phi \phi \geq b \|\phi\|^2, \quad \forall \phi \in E. \quad 3.85$$

Theorem 3.6 demands that

$$G_x'' \phi \phi \geq \|\phi\| g(\|\phi\|), \quad \forall \phi \in E.$$

If in 3.85,  $g(\|\phi\|)$  is defined as,

$$g(\|\phi\|) = b \|\phi\|, \quad \forall \phi \in E \quad 3.86$$

with  $b > 0$ , then it clearly satisfies the original hypothesis on the function  $g$ . This completes the proof of the theorem.

Remark:

The consequence of the above theorem is that a function  $g$  satisfying the given hypothesis may not always exist. Much depends on the characteristic of the plant being controlled and on the desired output. That is, it depends on the properties of the functional  $F_x$  and on the function  $z$ . However a function  $g$  satisfying the given hypothesis can always be found if for every  $x$  and  $z \in E$ ,  $(F_x'' \phi \phi, Fx-z)$  is bounded below.

In other words,

$$\lim_{\|\phi\| \rightarrow \infty} (F_x^n \phi \phi, z - Fx) < \infty \quad 3.87$$

$$\forall x, z \in E.$$

Thus an arbitrary combination of system and desired signal may not be admissible for optimal control in the sense of the theorem 3.5 and 3.6.

If certain cost functionals fail to satisfy the hypothesis on the function  $g$  some other admissible cost functional satisfying the hypothesis may be adopted.

The following theorem gives the conditions under which the equation 3.69 will have a unique solution in  $L^2(I)$ .

Theorem 3.8

Let the operator  $F$  and its gradient  $\wedge$  satisfy the Lipschitz condition with constants  $\alpha_F(\rho)$  and  $\alpha_\wedge(\rho)$  respectively in the sphere  $T_\rho$  of radius  $\rho$  around the origin in the space  $E$ .

Further let,

$$\sup_{x \in T_\rho} \|F[x(\cdot), t]\| \leq M_F < \infty \quad 3.88$$

$$\text{and } \sup_{x \in T_\rho} \|\wedge[t, \xi | x(\cdot)]\| \leq M_\wedge < \infty \quad 3.89$$

under these assumptions the functional equation 3.69 has a unique solution in  $E$ , provided, for a given system,

$$\lambda > \mu \{ (\alpha_F(\rho) M_\wedge(\rho) + \alpha_\wedge(\rho) M_F(\rho) ) \}. \quad 3.90$$

and the desired output  $z(\cdot)$  is bounded above in norm by the inequality,

$$\|z\| \leq \frac{1}{\alpha_\wedge(\rho)} \left\{ \frac{\lambda}{\mu} - (\alpha_F(\rho) M_\wedge(\rho) + \alpha_\wedge(\rho) M_F(\rho)) \right\}. \quad 3.91$$

Proof:

Equation 3.69 reads,

$$x(t) = \eta \int_I \wedge(t, \xi | x(\cdot)) (z(\xi) - F[x(\cdot), \xi]) d\xi$$

Let us define a functional  $Hx$  by,

$$Hx \triangleq \eta \int_I \wedge(t, \xi | x(\cdot)) (z(\xi) - F[x(\cdot), \xi]) d\xi \quad 3.92$$

Then equation 3.69 can be written as,

$$x = Hx. \quad 3.93$$

Let  $x_1$  and  $x_2$  be any two elements in  $T_\rho$ , then by 3.92 we have,

$$\begin{aligned} Hx_1 - Hx_2 &= \eta \int_I [\wedge(t, \xi | x_1) - \wedge(t, \xi | x_2)] z(\xi) d\xi \\ &+ \eta \int_I [\wedge(t, \xi | x_2) F[x_2, \xi] - \wedge(t, \xi | x_1) F[x_1, \xi]] d\xi \\ &= \eta \int_I [\wedge(t, \xi | x_1) - \wedge(t, \xi | x_2)] z(\xi) d\xi + \\ &+ \eta \int_I [\wedge(t, \xi | x_2) - \wedge(t, \xi | x_1)] F[x_2, \xi] d\xi \\ &+ \eta \int_I [F[x_2, \xi] - F[x_1, \xi]] \wedge(t, \xi | x_1) d\xi. \end{aligned} \quad 3.94$$

$t \in I$

Therefore,

$$\begin{aligned} \|Hx_1 - Hx_2\| &\leq \eta \left\{ \|\wedge(t, \xi | x_1) - \wedge(t, \xi | x_2)\| (\|z\| + \|F[x_2, \xi]\|) \right. \\ &\left. + \|F[x_1, \xi] - F[x_2, \xi]\| \cdot \|\wedge(t, \xi | x_1)\| \right\}. \end{aligned} \quad 3.95$$

By the hypothesis of the theorem, equation 3.95 becomes,

$$\|Hx_1 - Hx_2\| \leq \eta \left\{ \alpha_{\wedge}(\rho) \|z\| + \alpha_{\wedge}(\rho) M_F(\rho) + \alpha_F(\rho) M_{\wedge}(\rho) \right\} \cdot \|x_1 - x_2\| \quad 3.96$$

Now by the principle of contraction mapping (see chapter 2), we know that the equation  $x = Hx$  has a fixed point in  $T_{\rho} \subset E$  provided  $H$  satisfies the following inclusion relation

$$H(T_{\rho}) \subset T_{\rho}.$$

That is, if

$$\eta \left\{ \alpha_{\wedge}(\rho) \|z\| + \alpha_{\wedge}(\rho) M_F(\rho) + \alpha_F(\rho) M_{\wedge}(\rho) \right\} < 1 \quad 3.97$$

then there exists a unique solution of the equation 3.93 in  $T_{\rho}$  (by theorem 2.1).

Recalling that  $\eta = \frac{\mu}{\lambda}$ , we have, from 3.97

$$\|z\| \leq \frac{1}{\alpha_{\wedge}(\rho)} \left\{ \frac{\lambda}{\mu} - (\alpha_F(\rho) M_{\wedge}(\rho) + \alpha_{\wedge}(\rho) M_F(\rho)) \right\}$$

This proves the inequality 3.91, Inequality 3.90 follows from the non-negativeness of the norm.

The solution of the equation

$x = Hx$  is given by  $x = \lim_n x_n$  where

$$x_n = Hx_{n-1}. \quad 3.98$$

and the limit is taken in the sense of the metric induced by the norm of the space. This completes the proof of the theorem.

Remarks:

The inequalities 3.90 and 3.91 shows that the larger is the cost of control relative to the cost of error, the larger is the domain of control.

### 3.4 DISCUSSIONS:

There are two possible extensions of the work presented in this section. The first is the extension towards multivariable systems. The second is the consideration of optimization on  $L^P$  spaces.

Let us briefly indicate the extension of the first type.

Let  $x_1, \dots, x_s$  be the inputs to the plant and let  $x_i \in L^2(I)$ . Let the functional operator describing the system be denoted by,

$$F_i [ x_1(\cdot), \dots, x_s(\cdot), t ], \quad (i = 1, 2, \dots, r), \quad t \in I$$



Fig. 3.3: Multivariable System.

Then the  $i$ th output of the plant is given by,

$$y_i(t) = F_i [ x_1(\cdot), \dots, x_s(\cdot), t ] . .$$

Let  $\lambda$  be equal to  $\lambda_i$  associated with the input  $x_i$  and  $\mu$  be equal to  $\mu_j$ , associated with the  $j$ th output.

Assuming similar cost functional as in equation 3.60, the one corresponding to the  $j$ th output can be written as,

$$G_j [ x_1, \dots, x_s ] = \sum_{i=1}^s \lambda_i \int_I [ x_i(t) ]^2 dt + \mu_j \int_I [ F_j [ x_1(\cdot), \dots, x_s(\cdot), t ] - z_j(t) ]^2 dt \quad 3.99$$

$$j = 1, 2, \dots, r.$$



In the first problem of this chapter we considered optimization on  $L^2(\Omega_I, B_I, \mu)$  space. If we consider  $L^P(\Omega_I, B_I, \mu)$  space ( $p \geq 1, q \geq 1, \frac{1}{p} + \frac{1}{q} = 1$ ), then it can be shown that the necessary and sufficient condition that  $A$  be the optimum system (Theorem 3.1) is that,

$$\left\langle |Ax-z|^{p/q} \text{sign}(Ax-z), Bx \right\rangle_{\nu} = 0 \quad 3.102$$

for all  $B \in H_N$ . Here  $\nu$  is the Borel measure on the  $L^P$  space.

It is clear from the above expression that for all  $p$  and  $q$  satisfying  $p, q \geq 1$  and  $\frac{1}{p} + \frac{1}{q} = 1$ , the integrand may not be even measurable. However if  $p/q$  is any odd integer, then the equation 3.102 reduces to,

$$\left\langle (Ax-z)^{p/q}, Bx \right\rangle_{\nu} = 0, \quad \forall B \in H_N \quad 3.103$$

which further reduces to equation 3.12 in the case of  $L^2(\Omega_I, B_I, \mu)$ .

In general, the equation 3.103 is quite complicated.

In the second problem considered in this section, we may consider optimization on  $L^P(I)$  space that leads to functional equation,

$$x(t) \left|^{p/q} \text{sign}(x(t)) = - \int_I \wedge(\xi, t | x(\cdot)) |F[x(\cdot); \xi] - z(\xi)|^{p/q} \text{sign}(F[x(\cdot); \xi] - z(\xi)) d\xi \quad 3.104$$

If  $p/q$  is an odd integer, we obtain,

$$(x(t))^{p/q} = \int_I \wedge(\xi, t | x(\cdot)) (z(\xi) - F[x(\cdot), \xi])^{p/q} d\xi \quad 3.105$$

A more interesting case arises, when, in addition to usual energy constraint, a constraint on the amplitude of the control signal  $x(\cdot)$

is also imposed. For example, a constraint of the form,

$$l(t) \leq x(t) \leq u(t), \quad \text{where } l(t) < u(t), \quad \forall t \in I,$$

may be very useful. This constraint may be incorporated in the functional  $G[x(\cdot)]$  by adding the extra term,

$$\int_I [y(t)]^p dt, \quad p \geq 2$$

where,  $y(t) \triangleq \frac{x(t) - a(t)}{b(t)}$  3.106

and,  $a(t) = \frac{1}{2} (u(t) + l(t))$

$b(t) = u(t) - l(t)$  .

The functional  $G$  now becomes,

$$G[x(\cdot)] = \lambda \int_I [x(t)]^2 dt + \mu \int_I [F[x(\cdot), t] - z(t)]^2 dt + \int_I [y(t)]^p dt, \quad 3.107$$

and by similar considerations as before it can be shown that the optimal control signal  $x(\cdot)$  must satisfy the functional equation,

$$x(t) = \frac{1}{2\lambda b^p(t)} (a(t) - x(t))^{p-1} + \eta \int_I \Lambda(\xi, t | x(\cdot)) [z(\xi) - F[x(\cdot), \xi]] d\xi . \quad 3.108$$

$t \in I.$

where  $p$  has been assumed to be any even integer.  $\eta$  and  $\Lambda$  are as defined before. To increase the sharpness of the constraint,  $p$  can be taken as large as desired. The investigation of the technique for solving the functional equation 3.108 is an interesting problem, specially for large  $p$ .

Another interesting case arises when an additive noise is acting at the input of the plant. In the presence of an additive noise  $\{n(t), t \in I\}$  (Fig. 3.4) acting at the input of the plant, the functional to be minimized is the expectation of  $G[x(\cdot), n(\sigma \cdot)]$ , where,

$$G[x(\cdot), n(\sigma \cdot)] = \lambda \int_I (x(t))^2 dt + \mu \int_I (F[x(\cdot) + n(\sigma \cdot), t] - z(t))^2 dt \quad 3.109$$

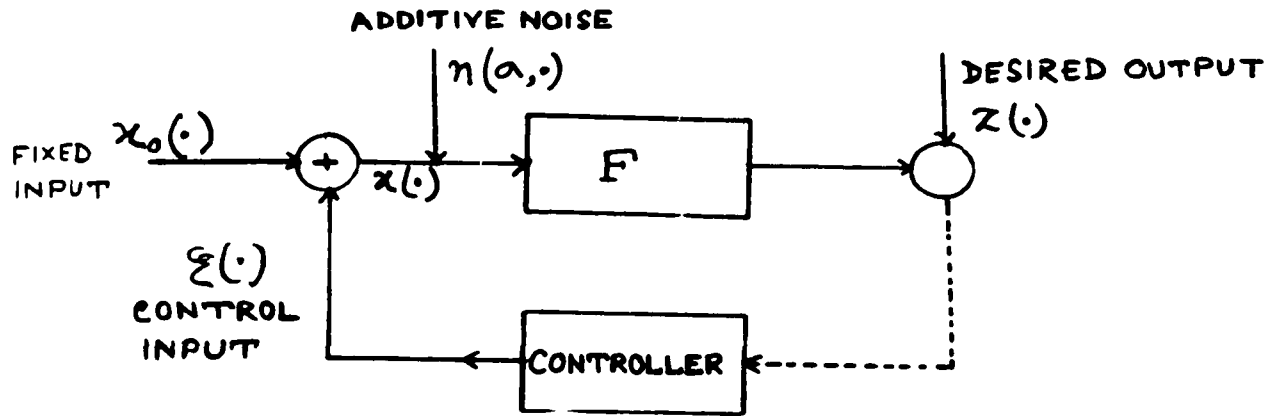


Fig. 3.4. Optimum Synthesis of Control Function in Presence of Noise.

Thus we must minimize,

$$\langle G[x(\cdot), n(\sigma, \cdot)] \rangle_{\mu_B} = \lambda \int_I (x(t))^2 dt + \mu \langle \int_I (F[x(\cdot) + n(\sigma, \cdot), t] z(t))^2 dt \rangle_{\mu_B} \quad 3.110$$

where,  $\mu_B$  is the measure induced by the stochastic process  $\{n(t), t \in I\}$  on the  $\sigma$ -field of measurable sets of the Hilbert space  $H$ . We assume that the functional  $F$  is  $B_I$  measurable in  $\sigma$  and Lebesgue measurable in  $t$  for each  $x \in H$  and that for every fixed  $x \in H$ ,  $F \in L^2(H \times I, B_I \times M_I, \mu_B, m)$ . It is also necessary that  $m(I) < \infty$ . Under these assumptions the equation 3.110 reduces to,

$$\langle G[x(\cdot), n(\sigma, \cdot)] \rangle_{\mu_B} = \lambda \int_I (x(t))^2 dt + \mu \int_I z^2(t) dt + \mu \int_I E_2[x(\cdot), t] dt - 2\mu \int_I E_1[x(\cdot), t] z(t) dt, \quad 3.111$$

where,

$$E_2[x(\cdot), t] = \int_H F^2[x(\cdot) + n(\sigma, \cdot), t] d\mu_B \geq 0$$

and 
$$E_1[x(\cdot), t] = \int_H F[x(\cdot) + n(\sigma, \cdot), t] d\mu_B$$

If the Functional  $F$  is continuous and bounded then both  $E_1$  and  $E_2$  are continuous and bounded. It can be shown, similarly as in theorem 3.5, that the necessary condition that  $x(\cdot)$  be the optimal control signal is that it satisfies the following functional equation,

$$x(t) = \frac{\mu}{\lambda} \int_I E_1^{(1)}[x(\cdot), \xi, t] z(\xi) d\xi + \frac{\mu}{2\lambda} \int_I E_2^{(1)}[x(\cdot), \xi, t] d\xi. \quad 3.112$$

where  $E_1^{(1)}$  and  $E_2^{(1)}$  are the first derivatives of the functionals  $E_1$  and  $E_2$  in the sense of Gateaux at the point  $x \in H$ .

For  $x(\cdot)$  to be the optimal control signal it is sufficient that both  $E_1^{(1)}$  and  $E_2^{(1)}$  possess Gateaux derivative at  $x(\cdot) \in H$  and that there exists an  $\epsilon > 0$  such that for the given  $z(\cdot) \in H$  the following inequality holds:

$$(E_1^{(n)} \varphi \varphi, z) \leq \left( \frac{2\lambda - \epsilon}{2\mu} \right) \|\varphi\|^2 + \frac{1}{2} (E_2^{(n)} \varphi \varphi, 1)$$

$$\forall \varphi \in H,$$

This follows exactly as in theorem 3.7.

**CHAPTER IV**  
**THE PROBLEM OF CANONICAL REPRESENTATION OF**  
**NONLINEAR SYSTEMS BY ORTHOGONAL**  
**FUNCTIONALS.**

#### 4.1 INTRODUCTION:

In the preceding chapters we considered that a physical system is represented by either of the operators of type  $A_V$ ,  $A_H$  or  $A_U$  and discussed several problems arising in nonlinear systems. In engineering practice it is often desirable to represent the system in a form suitable for experimental evaluation of the parameters determining the system. Just as it is convenient to have a complete orthonormal set of functions for the Fourier expansion of a given function belonging to some class, so also is it convenient to have a complete orthonormal set of functionals that can be used to develop any functional belonging to some class. In the case of Fourier expansion of a function, the domain of the function is either an interval on the real line or some subset of Euclidian  $n$ -space ( $n < \infty$ ). In the case of Fourier expansion of a functional, the domain is invariably an infinite dimensional function space for example any of the Banach spaces. Let us assume that  $\mu$  is a countably additive measure defined on some Banach space  $B$  and let  $I$  be an interval on which the elements of  $B$  are defined; Let  $\{F_i [x(\cdot), t]\}$  be a set of complete orthonormal functionals defined on  $B$  with  $x(\cdot) \in B$  and  $t \in I$ . If  $F[x(\cdot), t]$  is an element of the corresponding measure space  $L^2(B, M, \mu)$  then  $F$  could be represented as,

$$F [ x(\cdot), t ] = \sum_i a_i(t) F_i [ x(\cdot), t ] \quad \text{[ measure ]}^{\text{in}}$$

and the Fourier co-efficients would be given by,

$$a_k(t) = \int_B F [ x(\cdot), t ] F_k [ x(\cdot), t ] d\mu(x(\cdot)).$$

The problem is not quite that simple. In fact there exists no countably additive measure on an arbitrary Banach space, and the theory of integration on such spaces is just beginning to develop. Several examples of nonexistence of countably additive measure on Hilbert space are presented by Gross, Loewner [ 13, 14 ] .

To avoid this delicate question, we will concentrate on the so-called tame subsets of the Hilbert space  $H$ .

A tame subset of a real Hilbert space  $H$  is a subset  $A$  defined by  $A = \{x \in H : Px \in B\}$  where,  $P$  is a finite dimensional projection and  $B$  is a Borel set in the range of  $P$  to be denoted by  $R(P)$ . Then,  $A$  is said to be based on  $R(P)$ . For such a set  $A$ , one may define its measure by,

$$\mu(A) = \left(\frac{\lambda}{\pi}\right)^{n/2} \int_B \exp - \lambda \|x\|^2 dm(x) \quad 4.3$$

where  $dm(x)$  denotes Lebesgue measure on  $R(P)$  and  $n$  is the dimension of  $R(P)$ , and  $\lambda > 0$  is a parameter. Here  $P$  may be thought of as an orthogonal projection of dimension  $n$  of  $H$  onto  $\mathcal{L}^2$ . Let  $P'$  be another projection of dimension  $m > n$  such that  $R(P') \supset R(P)$ , then  $A$  is also based on  $R(P')$  ie.  $A$  is based on all subspaces of dimension larger than that of  $R(P)$ .

A tame set is usually known as a cylinder set. The cylinder sets form a ring  $\Sigma$  and  $\mu$  is additive on  $\Sigma$  since any two cylinder sets are based on a common finite dimensional subspace. However it can be proved that  $\mu$  as defined by equation 4.3 is not countably additive on  $\Sigma$ , [ 13a ] .

However, when  $\mu$  is restricted to the  $\sigma$  - ring  $\delta_{R(P)}$  of cylinder sets based on a fixed finite dimensional subspace  $R(P)$  then clearly  $\mu$ ,

as defined by equation 4.3, is countably additive.

A functional  $F[x(\cdot)]$  defined on  $H$  is called a tame function on  $H$  if  $F[x(\cdot)] = F[Px(\cdot)]$  for some finite dimensional projection  $P$  and if  $F$  is measurable with respect to  $\Sigma$ . In that case  $F$  is said to be based on  $P$ .

#### 4.1-1 STOCHASTIC PROCESSES:

A stochastic process  $\{x(\tau), \tau \in I\}$  is a family of real or complex valued functions defined on an arbitrary index set  $I$  such that for each  $\tau \in I$ ,  $x(\tau)$  is a random variable taking values on the real line  $R_\tau = (-\infty, +\infty)$ .

In order to satisfy the countable additivity property demanded by probability measures it is necessary that the class of intervals on  $R_\tau$  form a  $\sigma$ -ring. This is because all classes of subsets of  $R_\tau$  are not measurable [2]: The minimal  $\sigma$ -field over the class of all subsets of the real line  $R_\tau$  satisfying the above property is the Borel field  $B_\tau$ . The elements of  $B_\tau$  are Borek sets in  $R_\tau$ , and the corresponding measurable space  $(R_\tau, B_\tau)$  is the Borel line.

Let  $\Omega_I$  be a linear space of functions defined on  $I$  and  $P_n \triangleq P_{\tau_1 \dots \tau_n}$  be a projection operator which assigns to each element in  $\Omega_I$  a set of  $n$ -tuples.

$$P_{\tau_1 \dots \tau_n} [x(\tau)] = (x(\tau_1), \dots, x(\tau_n)) \quad 4.4$$

with  $\tau_1, \tau_2, \dots, \tau_n \in I$ .

Let  $R_{\tau_i}$  be the family of real lines ( $i = 1, 2, \dots, n$ ) and  $B_{\tau_i}$  be the corresponding family of Borel sets. Then an  $n$ -dimensional Borel measurable space  $(\Omega_{I_n}, B_{I_n})$  is defined by the cartesian product of the sequence  $\{(R_{\tau_i}^n, B_{\tau_i}^n)\}$  of Borel lines, and is given by

$$(\Omega_{I_n}, B_{I_n}) = \prod_{i=1}^n (R_{\tau_i}, B_{\tau_i}) \quad 4.5$$

It is clear that  $\tau$  plays only the role of a parameter and  $P_{\tau_1 \dots \tau_n}(x(\cdot))$  can be written more conveniently as,

$$P_{\tau_1 \dots \tau_n}(x(\cdot)) = (x_{\tau_1}; \dots; x_{\tau_n})$$

Since, the value of  $x_{\tau_i}$  depends on the element  $x(\cdot) \in \Omega_{I_n}$ , it is further convenient to introduce a parameter to explicitly indicate this dependence. Thus we may write

$$P_{\tau_1 \dots \tau_n}(x(\cdot)) = (x_{\tau_1}(\sigma); \dots; x_{\tau_n}(\sigma))$$

where  $\sigma$  is an element in  $B_{I_n}$ .  $B_{I_n}$  may be generated by a class  $C_{I_n}$  of sets of the form

$$\{C_{I_n}^k\} = \{(a_1^k, \dots, a_n^k; b_1^k, \dots, b_n^k)\} \text{ where,}$$

$$A_k \triangleq (a_1^k, \dots, a_n^k; b_1^k, \dots, b_n^k) = \{\sigma \in \Omega_{I_n} : a_i^k < \sigma_i \leq b_i^k, i = 1, 2, \dots, n\}.$$

In this case  $B_{I_n}$  is said to be generated by the class  $C_{I_n}$  of  $n$ -dimensional cells in  $\Omega_{I_n}$ .

If the linear space  $\Omega_{I_n}$  is known to possess a complete orthonormal basis, the projection  $P_n$  may be taken as the usual orthogonal projection of  $\Omega_{I_n}$  into its  $n$ -dimensional subspace. In this case,

$$P_n x(\cdot) = ((x, \varphi_1)(\sigma), \dots, (x, \varphi_n)(\sigma)) \quad 4.6$$

Let  $A_{I_n}^k$  be an element of  $B_{I_n}$  (the class of Borel sets in  $\Omega_{I_n}$ ) then a Borel cylinder in  $\Omega_I$  is defined by,  $\{A_{I_n}^k \times \Omega_{I_n}^c\}_k$  for any finite  $n$  with  $A_{I_n}^k$  as the base. For a fixed  $n$ , the cylinders so defined form a  $\sigma$ -ring  $\delta_{\Omega_{I_n}}$ . Kolmogorav's consistency relation demands that

$$\mu_{B_I}(A_{I_n}^k \times \Omega_{I_n}^c) = \mu_{B_{I_n}}(A_{I_n}^k). \quad 4.7$$

It is known that the measure  $\mu_{B_I}$  defined by

$$d\mu_{B_{I_n}}(\sigma) = \left(\frac{\lambda}{\pi}\right)^{n/2} \frac{1}{\sqrt{|A_n|}} \exp - \lambda(X_n^T A_n X_n) dx_1 \dots dx_n \quad 4.8$$

where;  $A_n$  is a  $n \times n$  positive definite moment matrix,

$$\sigma \triangleq X_n = (x_1 \dots x_n) \in \Omega_{I_n}, \text{ and } \lambda \in (0, \infty), \text{ is a}$$

parameter; provides an important example of countably additive measure function for any finite  $n$ . Here the sequence  $\{x_n\}$  may be either equal to  $\{(x, \varphi_n)\}$  or  $\{x_\tau\}$ . It is clear that all finite dimensional measure functions  $\mu_{B_{I_n}}$  obtained from some suitable family of distribution functions define the measure of all finite dimensional Borel cylinders in  $\Omega_I$  through the consistency relation of Kolmogorov.

The measurable space corresponding to the function space  $\Omega_I$  may be given either of the following two forms, a and b

$$(a) \quad (\Omega_I, B_I) = \pi_{t \in I} (\Omega_t, B_t) \quad 4.9$$

$$(b) \quad (\Omega_I, B_I) = \pi_{t \in I_c} (\Omega_t, B_t) \quad 4.10$$

where  $I_c$  is a dense countable subset of the set  $I$  such that  $\overline{I_c} = I$ .

Sets characterized by,  $\{\sigma : \sup_{t \in I} x_t(\sigma) \leq a\}$  corresponding to the former class do not even belong to any  $\sigma$ -ring and they are not Borel measurable. However, if the sets  $\{\sigma : \sup_{t \in I} x_t(\sigma) \leq a\}$  and  $\{\sigma : \sup_{t \in I} x_t(\sigma) \leq a'\}$  differ at most by a set of measure zero then (a)  $\{t \in I_c\}$  will form a measurable space. Stochastic processes satisfying this condition are called separable [15]. We will concern ourselves with separable processes and define the corresponding measure space by the three tuples  $(\Omega_I, B_I, \mu_{B_I})$  where  $\mu_{B_I}$  is the Borel measure induced by the stochastic process  $(x(t, \sigma) \mid t \in I, \sigma \in B_I)$ . In summary, stochastic processes may be considered as elements of a measure space with  $\mu_{B_I}(\Omega_I) = 1$ .

#### 4.1-2 MEASURABLE FUNCTIONALS:

A functional  $F[x(\cdot)]$  defined over  $\Omega_I$  is called a cylinder or tame function if there exists at least one  $\Omega_I^n$ , a finite dimensional subspace of  $\Omega_I$ , such that  $F[P_n x(\cdot)] = F[x(\cdot)]$  for some finite  $n$ . Here  $R(P_n) = \Omega_{I_n}$ .

Let  $g_n(u_1 \dots u_n)$  be a continuous function defined on  $\Omega_{I_n}$  and let  $u_i$  be identified with  $x(\tau_i, \sigma)$  then  $g_n(x(\tau_1, \sigma), \dots, x(\tau_n, \sigma))$  is a Baire function on  $\Omega_{I_n}$ , with  $\sigma \in B_{I_n}$ .

We define a Borel measurable functional of order  $n$  by the  $n$ -fold Lebesgue integral on  $I^n \triangleq \underbrace{I \times I \dots \times I}_n$

$$B_n[x(\sigma, \cdot), \sigma] = \int_{I^n} \dots \int K_n(\tau_1 \dots \tau_n) g_n(x(\sigma, \tau_1), \dots, x(\sigma, \tau_n)) d\tau_1 \dots d\tau_n, \quad 4.11$$

for every  $\sigma \in B_{I_n}$  and for all positive integers  $n$ , where  $\{K_n\}$  are Lebesgue measurable and Lebesgue square integrable Kernels on  $I^n$ . Since in this section we will be mainly concerned with Volterra-Frechet operators [16], we will define  $g_n(x(\sigma, \tau_1); \dots x(\sigma, \tau_n)) = x(\sigma, \tau_1) \cdot x(\sigma, \tau_2) \dots x(\sigma, \tau_n)$ . The corresponding sequence of functionals measurable with respect to the  $\sigma$ -ring are the volterra homogeneous functionals  $F_n [x(\sigma, \cdot), \sigma]$ : We will consider that the sequence  $F_n [x(\sigma, \tau), \sigma]$  is  $B_I$  measurable and Lebesgue square integrable on  $\Omega_I$  with respect to the measure  $\mu_B$ . That is,

$$\int_{\Omega_I} |F_n [x(\sigma, \cdot), \sigma]|^2 d_{\mu_B}(\sigma) < \infty, \forall n \in J^+ \quad 4.12$$

where,  $\mu_B = \mu_{B_I}$  is as defined by expressions 4.7 and 4.8.

According to measure theory, the volterra regular functionals of any finite degree defined by,

$$R_n(\sigma) = \sum_{s=0}^n F_s [x(\sigma, \cdot), \sigma], \forall n \in J^+ \quad 4.13$$

are  $B_I$  measurable. By repeated application of Minkowski's inequality, it is easily proved that  $R_n(\sigma)$  is also square integrable on  $\Omega_I$  with respect to the measure  $\mu_B$ . It is now clear by 4.12 and 4.13, that the class of  $B_I$  measurable and volterra homogeneous functionals form a linear vector space. If condition 4.12 is satisfied it is exactly the Hilbert space  $L^2(\Omega_I, B_I, \mu_B)$  since it is easily verified that  $V$  satisfies all the postulates of Hilbert space.

The following relations are well defined,

$$\forall F, G, H \in L^2(\Omega_I, B_I, \mu_B) \text{ and } \forall a \in F(R/C)$$

$$(i) \quad \langle a F, H \rangle_{\mu_B} = a \langle F, H \rangle_{\mu_B} \triangleq a \int_{\Omega_I} F(\sigma) \overline{H(\sigma)} d\mu_B(\sigma).$$

$$(ii) \quad \langle F + G, H \rangle_{\mu_B} = \langle F, H \rangle_{\mu_B} + \langle G, H \rangle_{\mu_B}$$

$$(iii) \quad \langle F, H \rangle_{\mu_B} = \overline{\langle H, F \rangle_{\mu_B}}$$

and

$$(iv) \quad \langle F, F \rangle_{\mu_B} \geq 0$$

where the bar indicates the complex conjugate. The equality in (iv) holds if and only if  $F = \theta \in L^2(\Omega_I, B_I, \mu_B)$  almost everywhere except on sets of  $\mu_B$ -measure zero.  $F(R/C)$  is the field of real or complex numbers.

### Multiple Correlation.

In the construction of orthogonal functionals on the measure space  $(\Omega_I, B_I, \mu_B)$  using the sequence of measurable functions  $\{R_n(\mathcal{G})\}$  we need the following quantity,

$$q(\tau_1 \dots \tau_n) = \int_{\Omega_I} x(\sigma, \tau_1) \dots x(\sigma, \tau_n) d\mu_B.$$

By use of equation 4.8 it can be shown that,

$$\begin{aligned} q(\tau_1 \dots \tau_n) &= 0 \quad \text{for } n\text{-odd.} & (a) \\ &= \frac{1}{(2\lambda)^p} \sum_{\substack{n=2p \\ * \\ i \neq j}} \prod a_{ij}^{-1} \quad \text{for } n = 2p & (b) \end{aligned} \left. \vphantom{\begin{aligned} q(\tau_1 \dots \tau_n) \\ = \frac{1}{(2\lambda)^p} \sum_{\substack{n=2p \\ * \\ i \neq j}} \prod a_{ij}^{-1} \end{aligned}} \right\} 4.14$$

$i, j = 1, 2, \dots, 2p.$

where  $p$  is any non-negative integer,  $a_{ij}^{-1}$  are the elements of the variance co-variance matrix  $A_n^{-1}$  and the  $*$  indicates the sum over all distinct ways of pairing  $2p$  objects among themselves.

The product is taken over all distinct pairs of indices. For convenience in orthogonalization we will assume that  $a_{ij}^{-1} = \delta(\tau_i - \tau_j)$  for each  $i$  and  $j$ . In that case the parameter  $\lambda$  may be interpreted as the reciprocal of the power density of the process and the process may be referred to as the white Gaussian process. In fact this process is stationary both in wide and strict sense [15]. Many of these restrictions will be removed in the sequel.

#### 4.2. ORTHOGONAL FUNCTIONALS AND THEIR PROPERTIES:

##### 4.2-1 Construction of Orthogonal Functionals on the Measure Space $(\Omega_I, B_I, \mu_B)$ .

The procedure of orthogonalization is very similar to that of Hilbert schmidt and can be described as follows. The zero degree functional  $R_0$  is normalized. The first degree regular functional  $R_1$  is orthogonalized with respect to all zero degree functionals and the inner product of the resulting functional with itself may be evaluated. The second degree regular functional  $R_2$  may be orthogonalized with respect to all  $B_I$  measurable zero degree and first degree homogenous functionals, and the inner product of the resulting  $B_I$  measurable orthogonal functional with itself may be evaluated. Similarly the regular functional  $R_n$  may be orthogonalized with respect to all  $B_I$  measurable homogeneous functionals of degree  $m < n$  and the inner product of the resulting orthogonal functional with itself is to be evaluated.

By the term "all  $B_I$  measurable functionals" we mean any arbitrary sequence  $\{H_n(\sigma)\}$  of  $B_I$  measurable homogeneous functionals associated with a corresponding set  $\{L_n\}$  of Lebesgue

measurable and Lebesgue square integrable Kernels defined on  $\Gamma^n \triangleq (I \times I \times \dots \times I)$ , which are otherwise arbitrary. That is,

$H_n(\sigma)$  is expressed as,

$$H_n(\sigma) = \int \dots \int_{\Gamma^n} L_n(\tau_1 \dots \tau_n) x(\sigma, \tau_1) \dots x(\sigma, \tau_n) d\tau_1 \dots d\tau_n \quad 4.15$$

for all  $\sigma \in B_I$  and all  $L_n \in L^2(\Gamma^n)$ , and  $R_n(\sigma)$  is expressed by,

$$R_n(\sigma) = \sum_{s=0}^n \int \dots \int_{I^s} K_s(\tau_1 \dots \tau_s) x(\sigma, \tau_1) \dots x(\sigma, \tau_s) d\tau_1 \dots d\tau_s \quad 4.16$$

for  $\forall \sigma \in B_I$  and  $\forall K_s \in L^2(I^s)$ . Since we are mainly interested in real valued stochastic processes we will consider  $L^2(\Omega_I, B_I, \mu_B)$  to be the real Hilbert space of all  $B_I$  measurable functions. The inner product of any two elements  $f(\sigma)$  and  $g(\sigma) \in L^2(\Omega_I, B_I, \mu_B)$  will be denoted by  $\langle f, g \rangle_{\mu_B} \triangleq \int_{\Omega_I} f(\sigma) g(\sigma) d\mu_B(\sigma)$ . The set of

Kernels  $\{K_s\}$  and  $\{L_s\}$  will be assumed to be symmetric in all the variables. If they are not symmetric they can be symmetrized by writing,  $K_n(\tau_1 \dots \tau_n) = \sum_n \frac{P_n K_n^1(\tau_1 \dots \tau_n)}{n!}$

where  $P_n$  effects the permutation of all the variables  $(\tau_1 \dots \tau_n)$  of  $K_n^1$  in all possible ways.

We illustrate the procedure of orthogonalization by constructing few orthogonal functionals. We normalize  $R_0(\sigma)$  by setting

$$\langle R_0(\sigma), R_0(\sigma) \rangle_{\mu_B} = 1 = |K_0|^2$$

Taking  $K_0 = 1$  we have the normalized zeroth degree  $B_I$  measurable orthogonal functional

$$G_0 [K_0, x(\sigma), \sigma] = 1.$$

we orthogonalize  $R_1(\sigma)$  with any zeroth degree functional, and since

a zero degree functional is any constant, it is equivalent to orthogonalizing  $R_1(\sigma)$  with unity, thus,

$$\langle R_1(\sigma), 1 \rangle_{\mu_B} = \langle F_0, 1 \rangle_{\mu_B} + \langle F_1, 1 \rangle_{\mu_B} = 0$$

By using 4.14 and the fact that  $a_{ij}^{-1} = \delta(\tau_i - \tau_j)$  for all  $i$  and  $j$  we have,  $K_0 = 0$ , Hence the corresponding orthogonal functional to be denoted by  $G_1 [K_1, x(\sigma, \tau), \sigma]$  is,

$$G_1 [K_1, x(\sigma, \tau), \sigma] = \int_I K_1(\tau) x(\sigma, \tau) d\tau \quad 4.17$$

and,

$$\langle G_1, G_1 \rangle_{\mu_B} = \frac{1}{2\lambda} \int_I |K_1(\tau)|^2 d\tau. \quad 4.18$$

Next, we consider  $R_2(\sigma)$ , orthogonalize it with zero degree functional, then with first degree functional and finally evaluate the inner product of the corresponding orthogonal functional  $G_2$  with itself.

$$\begin{aligned} \langle R_2(\sigma), 1 \rangle_{\mu_B} &= \langle F_0, 1 \rangle_{\mu_B} + \langle F_1, 1 \rangle_{\mu_B} + \langle F_2, 1 \rangle_{\mu_B} = 0 \\ &= K_0 + \frac{1}{2\lambda} \int_I K_2(\tau, \tau) d\tau = 0 \end{aligned}$$

$$\text{Therefore, } K_0 = -\frac{1}{2\lambda} \int_I K_2(\tau, \tau) d\tau, \quad 4.19$$

assuming that  $K_2(\tau, \tau) \in L^1(I)$ .

$$\langle R_2(\sigma), H_1(\sigma) \rangle_{\mu_B} = \langle F_0, H_1 \rangle_{\mu_B} + \langle F_1, H_1 \rangle_{\mu_B} + \langle F_2, H_1 \rangle_{\mu_B} = 0$$

$$\text{by 4.14 " } = \frac{1}{2\lambda} \int_I K_1(\tau) L_1(\tau) d\tau = 0, \quad 4.20$$

Since  $L_1(\tau) \in L^2(I)$  and otherwise arbitrary we can rewrite equation 4.20,

$$\frac{1}{2\lambda} \int_I |K_1(\tau)|^2 d\tau = 0$$

Hence  $K_1(\tau) = 0$  a.e except on sets of Lebesgue measure zero.

Therefore,  $G_2$  is given by,

$$G_2 [K_2(\tau_1, \tau_2), x(\sigma, \tau), \sigma] = \int_{I^2} \int K_2(\tau_1, \tau_2) x(\sigma, \tau_1) x(\sigma, \tau_2) d\tau_1 d\tau_2 - \frac{1}{2\lambda} \int_I K_2(\tau, \tau) d\tau \quad 4.21$$

and

$$\langle G_2, G_2 \rangle_{\mu_B} = \frac{2!}{(2\lambda)^2} \int_{I^2} \int |K_2(\tau_1, \tau_2)|^2 d\tau_1 d\tau_2 \quad 4.22$$

Similarly we consider  $R_3(\sigma)$ , orthogonalize it with any zero-degree, first degree, and second degree  $B_I$  measurable homogeneous functionals and finally obtain the inner product of the resulting orthogonal functional  $G_3$  with itself :

$$\langle R_3(\sigma), 1 \rangle_{\mu_B} = \langle F_0, 1 \rangle_{\mu_B} + \langle F_1, 1 \rangle_{\mu_B} + \langle F_2, 1 \rangle_{\mu_B} + \langle F_3, 1 \rangle_{\mu_B} = 0$$

by 4.14a,  $\langle F_1, 1 \rangle_{\mu_B} = \langle F_3, 1 \rangle_{\mu_B} = 0$  and by 4.14 b

$$\langle F_0, 1 \rangle_{\mu_B} = K_0 \text{ and } \langle F_2, 1 \rangle_{\mu_B} = \frac{1}{2\lambda} \int_I K_2(\tau, \tau) d\tau$$

$$\text{Hence, } K_0 = - \frac{1}{2\lambda} \int_I K_2(\tau, \tau) d\tau \quad 4.23$$

$$\langle R_3(\sigma), H_1(\sigma) \rangle_{\mu_B} = \langle F_0, H_1 \rangle_{\mu_B} + \langle F_1, H_1 \rangle_{\mu_B} + \langle F_2, H_1 \rangle_{\mu_B} + \langle F_3, H_1 \rangle_{\mu_B} = 0$$

By 4.14a,  $\langle F_0, H_1 \rangle_{\mu_B} = \langle F_2, H_1 \rangle_{\mu_B} = 0$  and by 4.14b, we have,

$$\langle F_1, H_1 \rangle_{\mu_B} = \frac{1}{2\lambda} \int_I K_1(\tau) L_1(\tau) d\tau \text{ and}$$

$$\langle F_3, H_1 \rangle_{\mu_B} = \frac{3}{(2\lambda)^2} \int_{I^2} \int K_3(\tau_1, \tau_2, \tau_2) L_1(\tau_1) d\tau_1 d\tau_2$$

Hence,

$$\langle R_3, H_1 \rangle_{\mu_B} = \int_I L_1(\tau_1) \left[ \frac{1}{2\lambda} K_1(\tau_1) + \frac{3}{(2\lambda)^2} \int_I K_3(\tau_1, \tau_2, \tau_2) d\tau_2 \right] d\tau_1 = 0$$

Since, this must be true  $\forall L_1 \in L^2(I)$  it is necessary and sufficient that,

$$K_1(\tau_1) = - \frac{3}{2\lambda} \int_I K_3(\tau_1, \tau_2, \tau_2) d\tau_2 \text{ a.e. on } I. \quad 4.24$$

Again,

$$\langle R_3, H_2 \rangle_{\mu_B} = \langle F_0, H_2 \rangle_{\mu_B} + \langle F_1, H_2 \rangle_{\mu_B} + \langle F_2, H_2 \rangle_{\mu_B} + \langle F_3, H_2 \rangle_{\mu_B} = 0$$

By 4.14a,  $\langle F_1, H_2 \rangle_{\mu_B} = \langle F_3, H_2 \rangle_{\mu_B} = 0$  and by 4.14b,

$$\langle F_0, H_2 \rangle_{\mu_B} = \frac{K}{2\lambda} \int_I L_2(\tau, \tau) d\tau \quad 4.25$$

and

$$\begin{aligned} \langle F_2, H_2 \rangle_{\mu_B} &= \frac{2}{(2\lambda)^2} \int_{I^2} \int K_2(\tau_1, \tau_2) L_2(\tau_1, \tau_2) d\tau_1 d\tau_2 \\ &+ \frac{1}{(2\lambda)^2} \int_{I^2} \int K_2(\tau_1, \tau_1) L_2(\tau_2, \tau_2) d\tau_1 d\tau_2 \end{aligned} \quad 4.26$$

Hence, by 4.23, 4.24 and 4.25, we have,

$$\langle R_3, H_2 \rangle_{\mu_B} = \frac{1}{(2\lambda)^2} \int_{I^2} \int K_2(\tau_1, \tau_2) L_2(\tau_1, \tau_2) d\tau_1 d\tau_2 = 0 \quad 4.27$$

for all  $L_2 \in L^2(I^2)$ . Therefore,  $K_2(\tau_1, \tau_2) = 0$

almost everywhere on  $I^2$ , which in addition implies that  $K_0 = 0$ , (equation 4.23). Thus, we obtain the corresponding orthogonal functional  $G_3$  of degree three,

$$G_3 [ K_3(\tau_1, \tau_2, \tau_3), x(\sigma, \tau), \sigma ] \\ = \int_{I^3} K_3(\tau_1, \tau_2, \tau_3) x(\sigma, \tau_1) x(\sigma, \tau_2) x(\sigma, \tau_3) d\tau_1 d\tau_2 d\tau_3 \\ - \frac{3}{2\lambda} \int_{I^2} K_3(\tau_1, \tau_2, \tau_2) x(\sigma, \tau_1) d\tau_1 d\tau_2 \quad 4.28$$

and

$$\langle G_3, G_3 \rangle_{\mu_B} = \frac{3!}{(2\lambda)^3} \int_{I^3} |K_3(\tau_1, \tau_2, \tau_3)|^2 d\tau_1 d\tau_2 d\tau_3 \quad 4.29$$

In general an  $n$ th degree orthogonal functional  $G_n$  is given by,

$$G_n [ k_n(\tau_1 \dots \tau_n), x(\sigma, \tau), \sigma ] \\ = \int_{I^n} K_n(\tau_1 \dots \tau_n) \sum_{p=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^p \frac{n!}{(4\lambda)^p (n-2p)! p!} S_p(x(\sigma, \tau_1), \dots, x(\sigma, \tau_n)) d\tau_1 \dots d\tau_n \quad 4.30$$

where,

$$S_p(x(\sigma, \tau_1), \dots, x(\sigma, \tau_n)) = \prod_{i=1}^n x(\sigma, \tau_i) \text{ for } p = 0, \quad 4.30a$$

and for  $p \geq 1$

$$S_p(x(\sigma, \tau_1), \dots, x(\sigma, \tau_n)) = \left( \prod_{i \neq j}^{2p} \frac{-1}{a_{ij}} \right) \cdot \prod_{i=2p+1}^n x(\sigma, \tau_i) \quad 4.30b$$

$$\text{where, } \prod_{i \neq j}^{2p} \frac{-1}{a_{ij}} = \delta(\tau_1 - \tau_2) \dots \delta(\tau_{2p-1} - \tau_{2p}). \quad 4.30c$$

$$\langle G_n, G_n \rangle_{\mu_B} = \frac{n!}{(2\lambda)^n} \int_{I^n} K_n(\tau_1 \dots \tau_n)^2 d\tau_1 \dots d\tau_n. \quad 4.31$$

$\forall n \in J^+$

4.2-2: PROPERTIES OF THE ORTHOGONAL FUNCTIONALS:

The properties of the set of  $B_I$  measurable orthogonal functionals  $\{G_n\}$  can be summarized as follows.

$$\forall K_n \in L^2(I^n) \text{ and } L_m \in L^2(I^m) \text{ and } m \neq n,$$

$$\langle G_n[K_n(\tau_1, \dots, \tau_n), x(\sigma, \tau), \sigma], G_m[L_m(\theta_1, \dots, \theta_m), x(\sigma, \theta), \sigma] \rangle_{\mu_B} = 0 \quad 4.32$$

$$\langle G_n[K_n(\tau_1, \dots, \tau_n), x(\sigma, \tau), \sigma], G_n[L_n(\theta_1, \dots, \theta_n), x(\sigma, \theta), \sigma] \rangle_{\mu_B}$$

$$= \frac{n!}{(2\lambda)^n} \int_{I^n} \dots \int K_n(\tau_1, \dots, \tau_n) L_n^*(\tau_1, \dots, \tau_n) d\tau_1 \dots d\tau_n \quad 4.33$$

In case  $K_n = L_n$  a.e. in  $I^n$ , this reduces to the equation 4.31.

The natural question that arises in connection with any orthogonal system of functions is, can a given function defined on the same basic set as the given orthogonal system, be represented by their suitable combination? This question can be answered relatively easily if we consider the representation in the mean square sense. That is, the sequence  $f_n(\sigma)$  defined as,  $f_n(\sigma) = \sum_{s=0}^n G_s[K_s(\tau_1, \dots, \tau_s), x(\sigma, \tau), \sigma]$  is said to converge in the limit in the mean of order two, to a function  $f(\sigma) \in L^2(\Omega_I, B_I, \mu_B)$  if to every  $\epsilon > 0 \exists$  a number  $n(\epsilon)$  such that

$$\int_{\Omega_I} |f(\sigma) - f_n(\sigma)|^2 d\mu_B < \epsilon, \quad \forall n \geq n(\epsilon).$$

Convergence in the mean is usually assured if the function  $f(\sigma) \in L^2(\Omega_I, B_I, \mu_B)$  and the sequence  $\{G_n\}$  is complete in the class  $L^2(\Omega_I, B_I, \mu_B)$ . The following theorems which are analogous to those arising in the theory of Fourier series of functions defined on finite dimensional spaces will clarify this fact.

**Theorem 4.1**

Let  $\{G_s [K_s(\tau_1 \dots \tau_s), x(\sigma), \sigma]\}$  be a denumerable sequence of orthogonal functions defined on the Borel measurable sets of the function space  $\Omega_I$  with each of the Kernels  $K_s$  belonging to  $L^2(I^s)$ . Let, for every  $f(\sigma) \in L^2(\Omega_I, B_I, \mu_B)$ ,

$$\begin{aligned} & \langle f(\sigma), G_s [K_s(\tau_1 \dots \tau_s), x(\sigma), \sigma] \rangle_{\mu_B} \\ &= \frac{s!}{(2\lambda)^s} \int \dots \int_{I^s} L_s(\tau_1 \dots \tau_s) K_s^*(\tau_1 \dots \tau_s) d\tau_1 \dots d\tau_s. \end{aligned} \quad 4.34$$

for some Kernel  $L_s \in L^2(I^s)$  and for every  $s \in J^+$ .

Then,

$$\|f(\sigma)\|^2 \geq \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int \dots \int_{I^s} |L_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s. \quad 4.35$$

This inequality is analogous to the Bessel's inequality arising in the theory of Fourier series.

Proof:

$$\text{Let } J_n \triangleq \int_{\Omega_I} \left| f(\sigma) - \sum_{s=0}^n G_s [K_s(\tau_1 \dots \tau_s), x(\sigma), \sigma] \right|^2 d\mu_B. \quad 4.36$$

$$\begin{aligned} J_n &= \langle f(\sigma) - \sum_s G_s [K_s, x, \sigma], f(\sigma) - \sum_m G_m [K_m, x, \sigma] \rangle_{\mu_B} \\ &= \langle f(\sigma), f(\sigma) \rangle_{\mu_B} - \langle \sum_s G_s [K_s, x, \sigma], f(\sigma) \rangle_{\mu_B} \\ &\quad - \langle f(\sigma), \sum_m G_m [K_m, x, \sigma] \rangle_{\mu_B} + \langle \sum_s G_s [K_s, x, \sigma], \sum_m G_m [K_m, x, \sigma] \rangle_{\mu_B} \\ &= \|f(\sigma)\|^2 - \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int \dots \int_{I^s} K_s(\tau_1 \dots \tau_s) L_s^*(\tau_1 \dots \tau_s) d\tau_1 \dots d\tau_s \\ &\quad - \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int \dots \int_{I^s} L_s(\tau_1 \dots \tau_s) K_s^*(\tau_1 \dots \tau_s) d\tau_1 \dots d\tau_s \\ &\quad + \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int \dots \int_{I^s} |K_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s \geq 0 \end{aligned} \quad 4.37$$

$$\begin{aligned}
 J_n = \|f(\sigma)\|^2 + \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int_{I^s} \dots \int |L_s(\tau_1 \dots \tau_s) - K_s(\tau_1, \dots, \tau_s)|^2 d\tau_1 \dots d\tau_s \\
 - \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int_{I^s} \dots \int |L_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s \geq 0 \quad 4.38
 \end{aligned}$$

Since the first and the third terms are independent of the choice of the Kernels of the orthogonal system  $\{G_s[K_s, x, \sigma]\}$  it is clear that the non-negative integral  $J_n$  attains its minimal value if and only if for each  $s$ ,  $K_s = L_s$  a.e on  $I^s$ . That is, the Kernels  $K_s$  in the expansion of  $f$  with respect to the orthogonal system  $\{G_s\}$  must be chosen as those arising from taking the inner-product (equation 4.34). Let the corresponding minimal value of  $J_n$  be denoted by  $J'_n$ . Then,

$$\begin{aligned}
 J'_n = \|f(\sigma)\|^2 - \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int_{I^s} \dots \int |L_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s \geq 0 \quad 4.39 \\
 \forall n \in J^+
 \end{aligned}$$

Thus for any  $f(\sigma) \in L^2(\Omega_I, B_I, \mu_B)$  and any positive integer  $n$ ,

$$\|f(\sigma)\|^2 \geq \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int_{I^s} \dots \int |L_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s \quad 4.40$$

This is exactly equivalent to the Bessel's inequality encountered in the theory of Fourier series. This inequality shows that the infinite series, is always, convergent,

$$\sum_{s=0}^{\infty} \frac{s!}{(2\lambda)^s} \int_{I^s} \dots \int |L_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s < \infty \quad 4.41$$

and  $J'_n$  can be put into the form,

$$\begin{aligned}
 J'_n = (\|f(\sigma)\|^2 - \sum_{s=0}^{\infty} \frac{s!}{(2\lambda)^s} \int_{I^s} \dots \int |L_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s) \\
 + \sum_{s=n+1}^{\infty} \frac{s!}{(2\lambda)^s} \int_{I^s} \dots \int |L_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s \quad 4.42
 \end{aligned}$$

$J'_n = \eta + \epsilon_n$ , where the first term inside the bracket is replaced by  $\eta$  and the second term by  $\epsilon_n$ .

It is clear that  $\eta \geq 0$  and by the inequality 4.40,  $\lim \epsilon_n = 0$ . Hence we see that a necessary and sufficient condition for a function  $f(\sigma) \in L^2(\Omega_I, B_I, \mu_B)$  to be approximated in the mean by a suitable choice of  $L^2$  Kernels  $\{K_s\}$  of the orthogonal system  $\{G_s\}$  is that  $\eta = 0$ . This immediately leads to an equality analogous to Parseval's equality in Fourier theory,

$$\|f(\sigma)\|^2 = \sum_{s=0}^{\infty} \frac{s!}{(2\lambda)^s} \int_{I^s} |L_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s \quad 4.43$$

Q.E.D

Remark:

It is clear that if the Kernels  $\{K_s\}$  in the system of orthogonal functionals  $\{G_s(K_s, x, \sigma)\}$  were chosen such that,

$$\frac{s!}{(2\lambda)^s} \int_{I^s} |K_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s = 1 \quad 4.44$$

$\forall s \in J^+$ , then  $\langle G_s, G_s \rangle = 1$  and the set  $\{G_s\}$  becomes an orthonormal set. In that case a function  $f(\sigma) \in L^2(\Omega_I, B_I, \mu_B)$  can be represented as,

$$f(\sigma) = \text{l.i.m.}_n \sum_{s=0}^n a_s G_s [K_s, x, \sigma].$$

In this case the Bessel's inequality 4.40 and the Parseval's equality 4.43 become respectively,

$$\|f(\sigma)\|^2 \geq \sum_{s=0}^n |a_s|^2 \quad 4.40'$$

and 
$$\|f(\sigma)\|^2 = \sum_{s=0}^n |a_s|^2 \quad 4.43'$$

where  $a_s = \langle f(\sigma), G_s \rangle_{\mu_B}$ . These have exactly the same form

as those arising in the theory of ordinary Fourier series. This is very advantageous from the point of view of representation of a system provided for a given  $I$  and a given power spectral density  $\lambda$  of the process, a complete set of Kernels  $\{K_s\}$  with the property 4.44 can be found. This is evidently a difficult job for an arbitrary choice of  $I$ .

It can be shown that if the set  $\{G_s [K_s, x, \sigma]\}$  is complete in the class of  $L^2(\Omega_I, B_I, \mu_B)$  functions then any function  $f(\sigma) \in L^2(\Omega_I, B_I, \mu_B)$  can be uniquely represented in terms of the set  $\{G_s\}$ .

For let

$$f(\sigma) = \text{l.i.m.}_n \left( \sum_{s=0}^n G_s [K_s, x, \sigma] = f_n \right) \quad 4.45$$

and

$$g(\sigma) = \text{l.i.m.}_n \left( \sum_{s=0}^n G_s [L_s, x, \sigma] = g_n \right) \quad 4.46$$

where  $K_s, L_s \in L^2(I^s)$  for each  $s$ .

If  $f(\sigma) = g(\sigma)$  a.e on  $\Omega_I$  except on sets of  $B_I$  measure zero

Then,  $\text{l.i.m.}_n \sum_{s=0}^n G_s [K_s - L_s, x, \sigma] = \phi$  the null function.

That is,

$$\lim_n \left\langle \sum_{s=0}^n G_s [K_s - L_s, x, \sigma], \sum_{s=0}^n G_s [K_s - L_s, x, \sigma] \right\rangle_{\mu_B} = 0$$

$$\lim_n \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int_{I^s} \int |K_s(\tau_1 \dots \tau_s) - L_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s = 0 \quad 4.47$$

The above equality holds if, for each  $s \in J^+$ ,  $K_s = L_s$  a.e on  $I^s$ . Conversely, if  $K_s = L_s$  (a.e) on  $I^s$  for each  $s \in J^+$  then  $f(\sigma) = g(\sigma)$  a.e in  $\Omega_I$  except on sets of  $B_I$  measure zero.

For,

$$\|f-g\| = \|f-f_n + g_n - g + f_n - g_n\| \leq \|f-f_n\| + \|g-g_n\| + \|g_n-f_n\|$$

Since the left hand side is independent of  $n$ ,

$$\|f-g\| \leq \lim_n \|g_n-f_n\| = \lim_n \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int_{\mathbf{I}^s} \left| L_s(\tau_1, \dots, \tau_s) - K_s(\tau_1, \dots, \tau_s) \right|^2 d\tau_1 \dots d\tau_s = 0$$

Hence  $f = g$  a. e.

This proves the uniqueness of representation of an  $L^2(\Omega_I, B_I, \mu_B)$  function with respect to the orthogonal set  $\{G_s\}$ .

An important theorem analogous to the celebrated theorem of Riesz-Fischer in the theory of Fourier series can also be proved for the orthogonal set  $\{G_s\}$ .

Theorem 4.2

Given the orthogonal set  $\{G_s [K_s, x, \sigma]\}$  and the sequence of  $L^2$  Kernels  $\{K_s\}$  associated with the set, then a necessary and sufficient condition that there exists an  $L^2(\Omega_I, B_I, \mu_B)$  function with  $\{K_s\}$  as its Kernels is that the series,

$$\sum_{s=0}^{\infty} \frac{s!}{(2\lambda)^s} \int_{\mathbf{I}^s} \left| K_s(\tau_1, \dots, \tau_s) \right|^2 d\tau_1 \dots d\tau_s \text{ converges.}$$

Proof:

If  $f(\sigma) \in L^2(\Omega_I, B_I, \mu_B)$  and

$$\langle f(\sigma), G_s \rangle_{\mu_B} = \frac{s!}{(2\lambda)^s} \int_{\mathbf{I}^s} \left| K_s(\tau_1, \dots, \tau_s) \right|^2 d\tau_1 \dots d\tau_s \tag{4.48}$$

then the necessary condition follows from the Bessel's inequality, 4.40.

For the proof of the sufficient condition let  $f_n(\sigma)$  be defined as,

$$f_n(\sigma) = \sum_{s=0}^n G_s [K_s, x, \sigma]. \tag{4.49}$$

Let  $m = n+p$  where  $p \gg 1$ , then,

$$f_m - f_n = \sum_{s=n+1}^m G_s [K_s, x, \sigma]$$

and

$$\langle f_m(\sigma) - f_n(\sigma), f_m(\sigma) - f_n(\sigma) \rangle_{\mu_B} = \sum_{s=n+1}^m \frac{s!}{(2\lambda)^s} \int_{I^s} |K_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s$$

ie for all  $p \gg 1$ ,

$$\|f_{n+p} - f_n\|^2 = \sum_{s=n+1}^{n+p} \frac{s!}{(2\lambda)^s} \int_{I^s} |K_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s \quad 4.50$$

By hypothesis the series on the right is convergent; so for every  $\epsilon > 0 \exists$  an integer  $n(\epsilon) \Rightarrow$  for every  $p \gg 1$

$$\|f_{n+p} - f_n\| \leq \epsilon \quad \forall n > n(\epsilon) \quad 4.51$$

This proves that the sequence  $\{f_n\}$  as defined by equation 4.49 is a cauchy sequence in the  $L^2(\Omega_I, B_I, \mu_B)$  space. Since an  $L^2$  space is complete it implies that, the sequence  $f_n$  converges in the limit in the mean to a function  $f$  also belonging to  $L^2(\Omega_I, B_I, \mu_B)$ . This function  $f$  has the given Kernels  $\{K_s\}$  as the Fourier Kernels in the expansion with respect to the set  $\{G_s\}$ . We can prove this by establishing contradiction.

Let  $\{L_s\}$  be the corresponding Kernels instead of the given set  $\{K_s\}$ .

Then,

$$\begin{aligned} \langle f - f_n, f - f_n \rangle_{\mu_B} &= \langle f, f \rangle_{\mu_B} + \langle f_n, f_n \rangle_{\mu_B} - \langle f, f_n \rangle_{\mu_B} - \langle f_n, f \rangle_{\mu_B} \\ &= \|f\|^2 + \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int_{I^s} |K_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s \\ &\quad - \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int_{I^s} L_s(\tau_1 \dots \tau_s) K_s^*(\tau_1 \dots \tau_s) d\tau_1 \dots d\tau_s - \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int_{I^s} K_s(\tau_1 \dots \tau_s) \times \\ &\quad \quad \quad L_s^*(\tau_1 \dots \tau_s) d\tau_1 \dots d\tau_s \end{aligned}$$

$$\begin{aligned}
 &= (\|f\|^2 - \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int_{I^s} |L_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s) \\
 &\quad + \sum_{s=0}^n \frac{s!}{(2\lambda)^s} \int_{I^s} |L_s(\tau_1 \dots \tau_s) - K_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s \quad 4.52
 \end{aligned}$$

By Bessel's inequality the first term is equal or greater than zero for any  $n$  and hence,

$$\lim_n \|f - f_n\|^2 \geq \sum_{s=0}^{\infty} \frac{s!}{(2\lambda)^s} \int_{I^s} |L_s(\tau_1 \dots \tau_s) - K_s(\tau_1 \dots \tau_s)|^2 d\tau_1 \dots d\tau_s .$$

Hence even if one of the Kernels say  $L_m \neq K_m$

$$\lim_n \|f - f_n\|^2 \geq \frac{m!}{(2\lambda)^m} \int_{I^m} |L_m(\tau_1 \dots \tau_m) - K_m(\tau_1 \dots \tau_m)|^2 d\tau_1 \dots d\tau_m . \quad 4.53$$

This leads to the contradiction that

$$\lim_n f_n = f .$$

Hence for each  $s$ ,  $L_s$  must be equal a.e on  $I^s$  to the given Fourier Kernels  $K_s$ . Q.E.D

Some additional properties of the orthogonal set are:

- (a)  $G_n$  is an  $n$ th degree  $B_I$  measurable homogenous functional of  $x$ , for each  $n \in J^+$ .
- (b)  $G_n$  is a linear functional of  $K_n \in L^2(I^n)$  for each  $n \in J^+$ .

For, each  $K_n$  and  $L_n \in L^2(I^n)$  and  $a, b \in F$

$$G_n [aK_n + bL_n; x, \sigma] = a G[K_n, x, \sigma] + b G[L_n, x, \sigma] .$$

- (c) Since  $L^2(\Omega_I, B_I, \mu_B)$  is a metric space, the distance between two systems  $s_1$  and  $s_2$  is given by

$$\|s_1 - s_2\| = \|f_{s_1}(\sigma) - f_{s_2}(\sigma)\| = \sqrt{\sum_{n=0}^{\infty} \frac{n!}{(2\lambda)^n} \int_{I^n} |K_n(\tau_1 \dots \tau_n) - L_n(\tau_1 \dots \tau_n)|^2 d\tau_1 \dots d\tau_n}$$

where  $\{K_n\} \in L^2(I^n)$  are the Kernels corresponding to the system  $s_1$

and  $\{L_n\} \in L^2(I^n)$  are those corresponding to the system  $s_2$ . Therefore two systems are equivalent iff  $K_n = L_n$  a.e on  $I^n$  for each  $n \in J^+$ .

(d) A system  $S$  is unstable in  $L^2(\Omega_I, B_I, \mu_B)$  sense if and only if,

$\mu_B \{ \sigma \in B_I : |f_s(\sigma)| > a \} = 1$ , for every  $a \in \mathbb{R}^1$ , ie if the system is unstable with probability one.

A system which is unstable in the present sense is also unstable in any other sense, and vice versa.

### REMARKS:

R1: The orthogonal functionals  $\{G_n\}$  constructed above are different from those of Wiener [17] only in the factor  $\lambda > 0$ . In the case of Wiener's  $G$  functionals this factor is taken equal to  $\frac{1}{2}$ . The factor  $\lambda$  is inversely proportional to the power density of the process i.e  $p(\omega) \propto \frac{1}{\lambda}$ . The advantage of using this factor is that if the power content in the signal is low the sequence  $f_n(\sigma)$  will converge rapidly to the output  $f(\sigma)$ . thus leading to the practical advantage of using fewer number of terms in approximating the system output. However, both these classes of orthogonal functionals suffer from few principal limitations both with respect to the signal and system.

(a) Signal:

- (1) The input is Gaussian.
- (2) The input is white.
- (3) The input is stationary

(b) System:

- (4) The system is time-invariant.
- (5) The system is of Volterra-Frechet type.

We will reserve comments on the limitation (1) until the end of this section. The second limitation can be removed by simply replacing the equation 4.30c by 4.30c'

$$\prod_{i \neq j}^{2p} a_{ij}^{-1} = R(\tau_1 - \tau_2) \dots R(\tau_{2p-1} - \tau_{2p}) \quad 4.30c'$$

where,

$$R(\tau_i - \tau_j) = \int_{\Omega_I} x(\sigma, \tau_i) x(\sigma, \tau_j) d\mu_B = R(\tau_j - \tau_i)$$

for  $i$  and  $j = 1, 2, \dots, 2p$ . It is only necessary to assume that the autocorrelation function is positive definite.

The third and fourth limitations can also be removed by introducing the following assumptions.

The input process,  $z(\sigma, t) \in L^2(\Omega_I \times I, B_I \times M, \mu_B \times m)$  and for almost every  $t \in I$ ,  $z$  is  $B_I$  measurable and that

$$\sigma_z^2(t) = \int_{\Omega_I} z^2(\sigma, t) d\mu_B < \infty \quad \text{with } t \in I: [0, T]$$

and  $m_2(t) = \int_{\Omega_I} z(\sigma, t) d\mu_B$  is measurable on  $I$  and finite almost everywhere on  $I$ . Let us define a random process  $x(\sigma, t)$  by,

$x(\sigma, t) = z(\sigma, t) - m(t)$  and let the measure induced by the stochastic process  $x(\sigma, t)$  be Gaussian as before with the elements of the matrix  $A_n^{-1} = \| a_{ij}^{-1} \|$  redefined as,

$$a_{ij}^{-1} = R(t_i, t_j) = \int_{\Omega_I} z(\sigma, t_i) z(\sigma, t_j) d\mu_B - m(t_i)m(t_j) .$$

The function  $R(t, \tau)$  is assumed to be symmetric and positive definite as before. Equations 4.15 and 4.16 are replaced by,

$$H_n(\sigma, t) = \int_{I^n} \dots \int_{I^n} L_n(m(\cdot); t, \tau_1 \dots \tau_n) x(\sigma, t - \tau_1) \dots x(\sigma, t - \tau_n) d\tau_1 \dots d\tau_n \quad 4.15'$$

and

$$R_n(\sigma, t) = \sum_{s=0}^n \int_{I^s} \dots \int K_s(m(\cdot); t, \tau_1 \dots \tau_s) x(\sigma, t-\tau_1) \dots x(\sigma, t-\tau_s) d\tau_1 \dots d\tau_s \quad 4.16'$$

and are assumed to belong to  $L^2(\Omega_I \times I, B_I \times M, \mu_B \times m)$  for each  $n$  and the Kernels are symmetric in  $\tau$  variables.

With these assumptions a similar set of orthogonal functionals can be constructed. In this case the equation 4.30 representing a general member of the sequence of orthogonal functionals becomes,

$$G_n [ K_n(m(\cdot); t, \tau_1 \dots \tau_n) \dots x(\sigma, \cdot), \sigma, t ] \\ = \int_{I^n} \dots \int K_n(m(\cdot); t; \tau_1, \dots, \tau_n) \sum_{p=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^p \frac{n!}{(4\lambda)^p (n-2p)! p!} S_p(x(\sigma, t-\tau_1) \dots x(\sigma, t-\tau_n)) \\ d\tau_1 \dots d\tau_n \quad 4.30'$$

where

$$S_0(x(\sigma, t-\tau_1), \dots, x(\sigma, t-\tau_n)) = \prod_{i=1}^n x(\sigma, t-\tau_i), \text{ for } p = 0$$

and for  $p \geq 1$

$$S_p(x(\sigma, t-\tau_1) \dots x(\sigma, t-\tau_n)) = \prod_{i \neq j}^{2p} a_{ij}^{-1} \prod_{i=2p+1}^n x(\sigma, t-\tau_i)$$

where,

$$\prod_{i \neq j}^{2p} a_{ij}^{-1} = R(t-\tau_1, t-\tau_2) \dots R(t-\tau_{2p-1}, t-\tau_{2p})$$

The properties of these orthogonal functionals are similar to those corresponding to the previous ones.

So far as the question of construction of orthogonal functionals is concerned, challenging theoretical problems are presented by the first and the fifth limitations. The fifth restriction excludes discontinuous nonlinear systems and even nonlinear systems with Kernels

containing products of impulse functions. To the knowledge of the writer there seems to be no immediate prospect of lifting this restriction. A theory of distribution for functionals, analogous to what has already led to the concept of generalized functions, may prove to be useful in this direction. Even if this problem is overcome in future and a theory of generalized functionals is developed it is quite conceivable that measurable functions may be transformed into non-measurable ones under these generalized transformation.

An attempt to lift the first limitation brings in several questions: Any other suitable measure defined on the function space say  $\Omega_I$  may be another special measure just as the one induced by Gaussian processes. If the input process is other than Gaussian then Wiener's orthogonal functionals cannot be used since these are only complete in the class  $L^2(\Omega_I, B_I, \mu_B)$  where  $\mu_B$  is the Gaussian measure. It is therefore, desirable to construct a suitable and useful family of measure spaces on a given function space say  $L^2(I)$  and the corresponding family of orthogonal functionals. The author made an attempt towards this direction by using a family of multi-dimensional probability functions discovered by P. W. Cooper [18]. These probability functions are defined as,

$$P_{m,n}(\underline{x}) = \frac{m \Gamma(n/2) \sqrt{|A|}}{2 \Gamma(n/m) \pi^{n/2}} \exp -(\underline{x}-\underline{\mu})^\dagger A(\underline{x}-\underline{\mu}) \frac{m}{2} \quad 4.54$$

$\forall m \in (0, \infty)$

where  $\underline{x}$  is n vector,  $A$ ;  $n \times n$  moment matrix and  $\underline{\mu}$  is the mean of the n-vector  $\underline{x}$ . For  $m=2$ ,  $P_{m,n}(\underline{x})$  reduces to the Gaussian probability function. The writer faced few difficulties while using these probability functions in the construction of a similar set of orthogonal functionals as those of Wiener. We could evaluate the multiple correlation of this process and it was found

for the case  $\underline{\mu} = \underline{0}$  that,

$$\left. \begin{aligned} q(\tau_1 \dots \tau_n) &= 0, \text{ for } n \text{ odd} \\ &= a(p, m) \sum_{\substack{n=2p \\ * \\ i \neq j}} \pi_{ij}^{-1} a_{ij}, \text{ for } n \text{ even} \end{aligned} \right\} 4.14'$$

This expression has the same form as that of equation 4.14 except that  $(\frac{1}{2\lambda})^p$  is replaced by  $a(p, m)$

$$a(p, m) = 2^{2p} \left( \frac{2}{m} - \frac{3}{2} \right) \frac{\Gamma\left(\frac{2p}{m} + \frac{1}{2}\right)}{\Gamma\left(p + \frac{1}{2}\right)},$$

for all  $p \in J^+$  and  $m \in (0, \infty)$ . For,  $m=2$  this reduces to the multiple correlation of the corresponding Gaussian process.

A typical problem that was faced in the process of orthogonalization of the sequence  $\{R_n(\sigma)\}$  with respect to the measure corresponding to the probability functions 4.54 is the following;

Let us consider orthogonalization of  $R_3(\sigma)$  with the zeroth, the first and the second degree homogeneous functionals as in section 4.2-1 by use of 4.14'.

Similarly as before, it can be shown that

$$K_0 = -a(1, m) \int_I K_2(\tau, \tau) d\tau \quad 4.23'$$

$$K_1(\tau_1) = \frac{3a(2, m)}{a(1, m)} \int_I K_3(\tau_1, \tau_2, \tau_2) d\tau_2 \quad 4.24'$$

and  $\langle R_3, H_2 \rangle$  leads to the following relation for an arbitrary  $L_2(\tau_1, \tau_2) \in L^2(I \times I)$ .

$$\begin{aligned} 2 a(2, m) \int_I \int_I L_2(\tau_1, \tau_2) K_2(\tau_1, \tau_2) d\tau_1 d\tau_2 \\ + (a(2, m) - a^2(1, m)) \int_I \int_I L_2(\tau_1, \tau_1) K_2(\tau_2, \tau_2) d\tau_1 d\tau_2 = 0 \end{aligned} \quad 4.27'$$

For an arbitrary  $L_2(\tau_1, \tau_2) \in L^2(I^2)$ , equation 4.27 is satisfied if and only if,  $K_2 = 0$  a. e, where as  $K_2 = 0$  a. e, is only a sufficient condition for 4.27' to be satisfied.

Similar situations are encountered in the process of orthogonalization of other members of the sequence  $\{R_n(\sigma)\}$ . Prof. Wonham expressed (private communication), his doubt about the consistency of the family of probability functions defined by equation 4.54, which escaped the notice of this author.

It is believed that this is the reason why problems like the one described above were encountered.

R2: Synthesis and analysis (including identification) of nonlinear systems subject to Gaussian noise are substantially simplified if the Wiener orthogonal functionals are used, [23, 24, 25]. In many engineering problems Gaussian process is almost a fiction. This leads to the question of construction on Banach spaces measures other than Gaussian. The answer to this problem is partly available in the famous Radon-Nikodym theorem [26] which asserts the existence of a continuous transformation between two absolutely continuous measures defined on the same  $\sigma$ -field of Borel measurable sets of the given Banach space. This shows that, if for a given non Gaussian process we can find a continuous transformation which when applied to the given non Gaussian process results in the Gaussian process, we can still use Wiener's results. In this case we can replace  $x$  by  $g(x)$  (equation 4.30) and obtain a set of orthogonal functionals for the non-Gaussian process. But if there, exists no such continuous transformation then we have to construct the orthogonal functionals again from basic considerations.

CONCLUSION:

In this thesis the main emphasis was on the functional approach to analysis and synthesis of nonlinear control problems, rather than differential equation approach. Functional representation includes differential systems, integral systems and also integro-differential systems. For example a representation of the form

$$y(t) = F \left[ x_1 \left( \begin{smallmatrix} \cdot \\ a_1 \end{smallmatrix} \right), \dots, x_n \left( \begin{smallmatrix} \cdot \\ a_n \end{smallmatrix} \right); t \right]$$

includes differential systems when  $F$  depends only on the values of the functions  $x_1(\cdot), \dots, x_n(\cdot)$  taken at the point  $t$ , i.e. when  $a_1 = a_2 = \dots = a_n = t$ . In certain problems of synthesis (such as the first problem of chapter III) such an abstract representation is not possible since it is the functional  $F$  that must be determined from the knowledge of  $y(\cdot)$  and  $x(\cdot)$ . In this dissertation such abstract representation was attempted at wherever it was found convenient.

Limitations and possible extensions of the work presented here have been discussed as far as possible at the end of each chapter.

The author feels that in future functional approach is going to be the principal tool in system engineering.

REFERENCE

- (1) Munroe, M.E.,  
"Measure and Integration"  
Addison-Wesley Publishing Co. Inc., 1959.
- (2) Halmos, P.R.,  
"Measure Theory"  
D. Van Nostrand Company, Inc., 1950.
- (3) Krasnosel'skii, M.A.,  
"Topological Methods in the Theory of Nonlinear  
Integral Equations"  
The MacMillan Company, New York, 1964.
- (4) Krasnosel'skii, M.A.,  
"Some Problems in Nonlinear Analysis",  
American Mathematical Society Translation  
Series 2, Vol. 10. 1958
- (5) Dunford, N. and Schwartz, J.T.,  
"Linear Operators", Part I.  
John Wiley, Second Edition, 1964.
- (6) Nemyckii, V.V.  
"The Fixed Point Method in Analysis",  
American Mat. Soc. Translation, Vol, 34, 1963.
- (7) Tricomi, F.G.,  
"Integral Equations",  
Interscience Publishers, Inc., N.Y., 1957.

- (8) Lovit, W. V.  
"Linear Integral Equations",  
Dover Publications Inc., N. Y., 1950.
- (9) Greub, W. and Rheinboldt, W.  
"On a generalization of an inequality of L. V. Kantorovich"  
American Math. Soc. Proceedings, Vol 10, 1959.
- (10) Kantorovich, L. V.  
"Functional analysis and applied Mathematics",  
Uspehi, Mat, Nauk, vol. 3 (1948), English translation -  
National Bureau of Standard, Report No. 1509 March 7, 1952.
- (11) Balakrishnan, A. V.  
"A general theory of nonlinear estimation problems in  
control systems"  
Journal of Mathematical analysis and applications  
Vol 8, No 1., Feb. 1964.
- (12) Vainberg, M. M.  
"Variational Methods for the Investigation of Nonlinear  
Operators"  
Gostekhizdat, Moscow, 1956.
- (13) Gross, Leonard,  
(a) "Classical Analysis on Hilbert Space"  
Proceeding of a Conference on the Theory and  
Applications of Analysis in Function Space,  
Endicot House, Massachusetts, June 9-13, 1963.  
(b) "Measurable Functions on Hilbert Space",  
Trans. Amer. Math. Soc. 105, 372-390, 1962.

- (14) Loewner, Karl,  
"Grundzuge Einer Inhaltslehre Im Hilbert Schen  
Raume" Ann. Math., 40, 816-833 (1939).
- (15) Doob, J.L.  
"Stochastic Processes",  
John Wiley and Sons Inc., New York, 1953.
- (16) Volterra, Vito  
"Theory of Functionals and of Integral and Integro  
Differential Equations",  
Dover Publications Inc., New York.
- (17) Wiener, N.  
"Nonlinear Problems in Random Theory"  
M.I. T. Press, 1958.
- (18) Cooper, P.W.  
"Multivariate Extension of Univariate Distribution",  
I.E.E.E Transactions on Electronic Computers,  
Oct. 1963.
- (19) Cronin, J.  
"Fixed Points and Topological Degree in Nonlinear  
Analysis",  
American Mathematical Society, Mathematical Surveys,  
No. 11, 1964.
- (20) Orlicz, W.  
"On Spaces of  $\phi$ -Integrable Functions"  
Proceedings of the International Symposium on Linear  
Spaces, A Publication of the Israel Academy of Sciences  
and Humanities, July 5-12, 1960.

- (21) Krasnosel'skii, M. A., Rutickii, Ya. B  
"Convex Functions and Orlicz Spaces"  
Noordhoff, Groningen, 1961.
- (22) Dolph, C. L. and Minty G. J.  
"On Nonlinear Integral Equations of Hammerstein Type",  
Nonlinear Integral Equations,  
Proceedings of an Advanced Seminar Conducted by  
the Mathematics Research Centre, U. S. Army,  
University of Wisconsin, Madison, April 21-24, 1963.
- (23) Lee, Y. W. Schetzen,  
"Measurement of the Kernels of a Nonlinear System by  
Cross correlation" Q. P. R. No 60, Research Lab.  
Electronics M. I. T., Jan 15, 1961.
- (24) George, D. A.  
"Continuous Nonlinear Systems",  
TR 355, Research Lab of Electronics, M. I. T., Jul. 24, 1959.
- (25) Glinski, G. S., Ahmed, N. U.  
"Representation of Nonlinear Systems by Orthogonal  
Functionals" T. R. 64-12, July, 1964, Ottawa;  
(presented at the Symposium on Signal Processing,  
Queens' University, Kingston, Canada, June 17-19, 1964.
- (26) Loève, M.  
"Probability Theory"  
D. Van Nostrand Co. Inc., Third Edition, 1963.

VITA

NAME Nasir Uddin Ahmed

BORN Pakistan, Jan. 1934

EDUCATED

SECONDARY St. Joseph's High School  
Khulna, Pakistan

COLLEGE Daulatpur College  
Khulna, East Pakistan

UNIVERSITY Ahsanullah Engineering College  
Dacca University, Pakistan

Harwell Reactor School  
A.E.R.E., England, U.K.

COURSES

Electrical Engineering

Reactor Engineering

DEGREES

B. Sc. (Engg.) (Dacca)

M. Sc. (Engg.) (Ottawa)