STATE SPACE AXIOMS AND STATE DESCRIPTIONS IN CANONICAL FORM

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

İZZET CEM GÖKNAR

1969

This is to certify that the

thesis entitled

State Space Axioms and
State Descriptions in Canonical Form

presented by

Izzet Cem GÖKNAR

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Engineering

James A. Resh
Major professor

Date May 14, 1969

O-169

8

2005

ABSTRACT

STATE SPACE AXIOMS AND STATE DESCRIPTIONS IN CANONICAL FORM

By

İzzet Cem Göknar

Although it dates back to Newton's use of positions and momenta, the concept of "State" has only been given an abstract and rigorous definition in the last decade by Zadeh.

In this thesis, starting with improved versions of Zadeh's "State Axioms," the necessity of another minor modification is shown and the different axiom sets are discussed. With the axioms modified, the important concept of "Equivalence Classes of Inputs" (the major tool of the behavioral approach) is used to investigate the properties of "Reduced and Half Reduced State Descriptions."

Then, the essential properties that "State Descriptions" acquire when the system is "Linear" and/or "Time-Invariant" are examined, and "State-Equations" in canonical form are obtained for a large class of distributed systems. The problem of approximating more general systems, with only minor restrictions on the input space, by systems that possess finite dimensional "State Spaces" is given a solution.

11-14-69
G 58876

STATE SPACE AXIOMS AND STATE DESCRIPTIONS IN CANONICAL FORM

By

İzzet Cem Göknar

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering

1969

ACKNOWLEDGMENTS

I would like to take this opportunity to formally acknowledge some people whose efforts, perhaps not directly related to the thesis, brought me to the state in which I am, influenced my thoughts, and therefore contributed to this work.

Listed in the chronological order I met them, I wish to thank my mother, Vedide Göknar, and my father, Saim Göknar, to whom I owe my existence and my education, my wife, Aytaç Göknar, whose love and support made "those moments" bearable, my Professor Tarık Özker, devoted to his country and the education of his students, who caused an evolution in my thoughts, and finally to my daughter, Elif Göknar, whose addition to the family has been a source of joy and strength.

I am indebted to Professor James A. Resh for his guidance of this thesis and for his unique advising, and to Professor Yilmaz Tokad for his valuable suggestions and discussions.

Finally I extend my gratitude to the institutions of Michigan State University and the Technical University of Istanbul for the fine education and the support they have provided.

"The men where you live," said the little prince, "raise five thousand roses in the same garden-- and they do not find in it what they are looking for."

"They do not find it," I replied.

"And yet what they are looking for could be found in one single rose, or in a little water."

"Yes, that is true," I said.

And the little prince added: "But the eyes are blind. One must look with the heart. . . ."

The Little Prince
Antoine de Saint-Exupéry

TABLE OF CONTENTS

Chapter | Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In order to present the results accomplished in this thesis and to lay down the general background for the subject considered in the thesis the present chapter is divided into four sections. The first section is devoted to the history of the concept of state and the steps toward its abstractization in the framework of modern system theory. We feel that before a meaningful discussion can be given for the findings of the thesis, some general concepts and terminology should be introduced. This is done in section 2. In section 3 the State Axioms proposed by various authors are outlined and some known results are given. Finally, in section 4 the remaining chapters of the thesis are summarized.

If some idea has to be given shortly about the results of the thesis, we can divide our accomplishments into three main groups.

The first group of results is about the State Axioms and what can be said about the State Descriptions in general without any restriction on the system under consideration. An improvement on the State Axioms is given and questions about the size of the State Space and about the nature and system-independent properties of the State Description are answered.

The second group is obtained by placing some restrictions on the nature of the system and then inquiring about the State Description. The basic properties of the State Descriptions of linear, time-invariant systems are investigated and results are obtained by using tools developed in the first group.

In the final third group, we develop analytical formulations of the State Description for some broad classes of systems. These representations can be used in the Theory of Distributed Parameter Systems, or in approximating them by systems with finite dimensional State Descriptions.

Outside the main goals of this thesis, some new Theorems are obtained in the Appendix that center about Orthonormal Series Expansions of Distributions as presented in [ZE2].

I.1--The Modern State Concept

The concept of "state," which dates back to Newton's introduction of positions and momenta as basic mechanical variables, has been used in analytical dynamics, celestial mechanics and quantum mechanics as tied to the concept of stored energy in such physical systems.

The following short discussion, that stems from a treatment on the historical background of the "modern concept of state," appeared in the literature in 1962 [ZA1]. As implied in this reference this modern concept was first used by Turing in his time-discrete machine. Briefly if x_t, u_t, y_t denote, respectively the state, the input and the output at time t, then the machine can be characterized by

$$\begin{aligned} x_{t+1} &= f\,(x_t, u_t) \\ & \qquad\qquad\qquad t = 0,\ 1,\ 2,\ --- \qquad (1) \\ y_t &= g\,(x_t, u_t) \end{aligned}$$

Shannon [SH] in 1948 used equations in the form (1) to characterize probabilistic systems in the sense that x_t and u_t determine the joint probability density function, $p(x_{t+1}, y_t/x_t, u_t)$ instead of x_t and y_t.

Two important notions, namely, equivalent states and equivalent machines were then introduced by Moore [MO] and by Huffman [HU] independently, but in a somewhat restricted form by the latter.

All the above work is in the discrete-state systems context. In the case of differential systems, the equations (1) take the form:

$$\begin{aligned} \frac{d}{dt}\,x(t) &= f\,(x(t),\ u(t)) \\ & \qquad\qquad\qquad\qquad (2) \\ y(t) &= g\,(x(t),\ u(t)) \end{aligned}$$

where x(t), u(t), y(t) are vectors representing the "state," the "input" and the "output." Equations (2) have been used, under different forms, in such fields as ordinary differential equations, analytical dynamics, celestial mechanics, quantum mechanics, etc. Their wide use in the field of automatic control was initiated almost twenty years ago, in Russia, by A. T. Luré, M. A. Aizerman, Ya. Z. Tsypkin, A. A. Fel'dbaum, A. Ya. Lerner, A. M. Letov, N. N. Krasovskii, I. G. Malkin, L. S. Pontryagin and others, and in the United States by Bellman, Kalman, Bertram, LaSalle, Laning, Battin, Friedland and others. General methods of setting up the state equations for RLC networks were later described by Bashkow [BA] and Bryant [BR]. These methods are extended to time varying networks by Kinarawala [KI].

Until recently, the concept of "state" was strongly connected with the specific physical identification of state variables as measurable quantities inside a specific system structure. For example, "the state vector" in an electrical network contains the variables corresponding to the branch capacitor voltages and the chord inductor currents. Thus the "initial state" at the "initial time" is physically the initial charge and the initial flux carried by those elements, and is reflected as the "initial conditions" on the differential equations modeling the network. This notion of state, namely that the state

is a set of internal variables from which everything else about the system can be calculated, is referred to as the "structural approach" to the concept of state [RE5]. In this approach an important property of the state that has to be singled out, is that it ensures a unique output for each given input.

Another approach to the concept of state is in the framework of modern system theory and is introduced in the following.

In Zadeh's view [ZA1], the importance of system theory lies in its abstract generality and in its concern with the mathematical properties of systems and not their physical form. Such an abstraction, however, should be reached from a number of known examples of such systems as physical, socio-economic, biological and others. If the state concept is not to be abandoned during this generalization, one has to be certain whether all the various instances of the state notion that appear in specific systems are sufficiently similar in meaning and usage to be covered by a single abstract definition; if so, what are the essentials of the notion? To elaborate this point further, we consider two examples, one in socio-economic, the other in biological systems.

For the first example, let the community of the Greater Lansing Area be our system, with the price of a certain good, say for example of Nehru jackets, and the advertisement expenditures as inputs and the demand for the same good as output. We shall concern ourselves with an important variable that affects the input-output relationship of the system: "the taste" of the community. It is true that, at different times, for the same price and advertisement expenditures, our community may not have the same demand for Nehru jackets. This is due to a change in "the taste" of the Greater Lansing Area; a taste more in favor of the good will create a larger demand for a given price and advertisement expenditures than a taste less in favor of the good. Thus if the price and the advertisement expenditures were given, as a function of time, one could determine the demand for Nehru jackets, as a function of time, if the taste of the community were known. Equally important is the tie existing between the taste and the past history of the community: the taste will certainly vary depending on the kind and intensity of the advertisement and the past fluctuations of the price. For example: fashion shows, constant T. V. commercials, larger numbers of people wearing Nehru jackets because of low prices, will probably push the taste to be more in favor of the good.

For the second example, we quote from Manning [MA]:

> It is a common observation that the same stimulus given to the same animal at different times does not always evoke the same response. Something inside the animal must have changed and we invoke an "intervening variable." This is something which comes between two things we can measure-- in this case the stimulus we give and the response

> we get out--and affects the relationship between them. . . . Already in this book we have mentioned two factors with different characteristics which alter the relationship between stimulus and response. These were "fatigue" and "maturation." To these we may add two others: "learning" and "motivation" . . .

From these examples we immediately recognize the important property that we noticed in the structural approach to the state, i.e., to make correspond a unique output to a given input, when the input and the state are known. We may therefore conclude that all the various instances of the state notion that appear in specific systems have a very important common property that may lead to a single abstract definition. What can better summarize "the mood of a human being (or an animal)," "the social conditions of a society," "the political conditions of a country" than "the <u>state</u> of the system"?

These examples also bring light to another important aspect of the state notion that was not clearly visible in the structural approach: the strong connection between the <u>history</u> of the system and the state. In fact "the taste of the society," "the fatigue, maturation, . . . of the animal," "the mood of the human being" at a given time, are all results of the past experiences of the system. Even in networks, the flux and the charge at time t_o, are the integral of the voltage and of the current up to time t_o, which certainly bear a relation to the past.

To conclude, the state, in this new context makes a unique output correspond to a given input by at least containing a minimum amount of information that consists in those features of the past experience of the system affecting its future behavior. This experiential aspect is named as "<u>the behavioral approach</u>" to the concept of state [RE5].

In 1962, Zadeh wrote [ZA1]:

> Despite the extensive use of the notion of state in the current literature, one would be hard put to find a satisfactory definition of it in textbooks or papers. A reason for this is that the notion of state is essentially a primitive concept, and as such is not susceptible to exact definition.

However, Zadeh in 1963 [ZA2] and Kalman in 1963-64 [KA, WE] have independently tackled the precise formulation of "<u>state descriptions</u>," and Resh in [RE 1-3] "exposed and eliminated a syndrome of shortcomings in these general formalizations of the state notion" [RE1], and offered two somewhat related though different sets of "State Axioms." Of these syndromes, some important ones were:

> Kalman's formulation, besides being cumbersome (at least to this author) had the obscurity of defining what is to be called a "<u>system</u>" in terms of his state axioms,

in Zadeh's (and Kalman's) formulation,

> The gross properties of the "state space" of a system were not uniquely determined by the system,

The states and the past histories of a system bore no necessary strong relation to one another,

All systems, causal and noncausal alike, possessed state descriptions.

Resh, when modifying Zadeh's axioms, also introduced a powerful tool, "the equivalence classes of pre-t_o inputs" which summarizes the history of the system up to time t_o and which bears strong relations to the states of the system at time t_o.

All the works summarized above being about the gross properties of the state descriptions, some analytical results are also obtained. It has first been pointed out in [ZA2], that for linear, time-invariant systems, the state space is a finite dimensional vector space iff the system can be characterized by ordinary differential equations. The finite dimensional case has then been extended by Balakrishnan, introducing some assumptions on the nature of the state space and the input space in [BA 1-3] and he derived a state description starting from the input-output description with some restrictions in [BA4]. The restrictions are the linearity and time-invariance of the system, except in [BA3] where they were allowed to be time varying. The main tool Balakrishnan used was the analytical theory of semigroups of linear operators as developed by Hille-Phillips and Yesida to obtain results of the form:

$$x(t) = T(t)\cdot x(o) + \int_{o}^{t} T(t-s)\cdot Lu(s)\cdot ds$$

where T(t) is a one-parameter semigroup of linear bounded transformations on the state space, and L a linear bounded transformation on the input space.

Finally, Resh [RE4] and very recently Resh and Göknar [RE5] have given a non-reduced state description of the form:

$$\frac{dx_s(t)}{dt} = s\cdot x_s(t) + u(t) \qquad s\varepsilon\mathbb{C}$$

$$y(t) = \int_{\mathcal{L}^{-1}} C(s)\cdot x_s(t)\cdot ds + \sum_{k=0}^{K} d_k\cdot u^{(k)}(t),$$

where the dimension (sic) of the state space is a two-dimensional continuum.

I.2--Some General Concepts and Terminology

In this and the following section, we specialize in the definitions of systems, objects, existence intervals, uniform objects, etc., and give the different "State Axioms," discuss them more in detail and state some known results. It is our feeling that here is the right place, although it may not be very usual, to do this since we talk of general concepts that underlie our work and present some known theorems for later references.

Webster defines a "SYSTEM" as ". . . an aggregation or assemblage of objects united by some form of interaction or interdependence," which is close to our engineering understanding although still remaining undefined because of the use of the synonymous "object."

The "Mathematics Dictionary" of James and James defines it as:

> (1) A set of quantities having some common property, such as the system of even integers, the system of lines passing through the origin, etc.
> (2) A set of principles concerned with a central objective, as, a coordinate system, a system of notation, etc.

which has no bearing to our concept of system whatsoever.

From an engineering point of view, the "system" definition can be given from two aspects; their main difference being the existence of the concept of "Terminals" in one and not in the other. As an example of the first one, "system" in [NE] is defined by:

$$S \triangleq \{(u,y) : uC_s y\}$$

where $u = [u_j(t)]$, $y = [y_j(t)]$, $j = 1_1 . . .$, k are the admissible pairs at the k terminals of the system with C_s denoting the determining constraints imposed by the system.

As we will be using the second definition of "system" in our context, we will give it in its greater details.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

DEF.I.2.1:

$T \triangleq$ a collection of half open intervals, $(\cdot,\cdot]$, of the real line, i.e.,

$T \triangleq \{(\cdot,\cdot] : (\cdot,\cdot] \subset \mathbb{R}\}$, the intervals in T are called OBSERVATION INTERVALS.

$R_I \triangleq$ a set of ordered pairs of time functions defined on $I \in T$, i.e.,

$R_I \triangleq \{(u,y) : \text{Dom } u = \text{Dom } y = I\}$

$A \triangleq$ is the family of all R_I when $I \in T$, i.e.,

$A \triangleq \{R_I : I \in T\}$.

A CONTINUOUS-TIME SYSTEM (as opposed to discrete-time system) is an ordered pair (T,A) where T and A are defined as above, satisfying:

(C1) If $I \in T$ and $(t_0,t_1] \subset I$ then $(t_0,t_1] \in T$

(C2) If $I \in T$ then $R_I \neq \phi$

(C3) If $I' \in T$ and $I \subset I'$ then $R_I \supseteq R_{I'}/I$

If the first member u of the ordered pair is called an INPUT, and the second one OUTPUT, the SYSTEM is then said to be ORIENTED.

For an oriented system, U_I will denote the set of all inputs whose domain is I and Y_I the set of all outputs whose domain is I, $Y_I(u)$ will be the set of all outputs that can occur as a response to u. $u_0 \circ u$ will mean that two inputs $u_0 \in U_{(t_1,t]}$ and $u \in U_{(t,t_2]}$ are in concatenation.

DEF.I.2.2:

Let a system (T,A) be given; for each observation interval I we define:

$R'_I \triangleq \{(u,y)\in R_I$: u,y are not the restriction to I of pairs in some $R_{I'}$, $\exists\ I \underset{\neq}{\subset} I'\}$

Then: $\hat{T} \triangleq \{I : I\in T$ and $R'_I \neq \Phi\}$.

The intervals in $\hat{T}$ are called the EXISTENCE INTERVALS.

DEF.I.2.3:

An oriented system is UNIFORM iff $\hat{T}$ is a unit set, i.e., contains a unique existence interval.

NOTE I.2.1: Thus for a uniform system it is clear that all pairs $(u,y)\in R_I$, for all $I\in T$ except one are the restrictions to I of some $(\hat{u},\hat{y})\in R_{I'}$, $I\subset I'$.

It has been shown in [RE1]:

> . . . portions of a system (T,A) derived from different existence intervals lead rather independent lives. In fact, one might consider them to be different systems which it has merely been convenient to describe in language suitable for treating them in some unified way.

Thus the loss of generality that entailed by the restriction of our concentration to uniform systems is very little.

FACT I.2.1: The description of a uniform system is completely known when the unique existence interval $\hat{I}\in\hat{T}$ and the input-output list $R_{\hat{I}}$ is given, due to the conditions C1, C2 and C3.

CON.I.2.1: From now on we will talk of OBJECTS and not of systems. No real difference exists between the two things these names describe. However, we will make the following distinction: an object is always a system, but not vice-versa. An object for us will consist of a single $R_{\hat{I}}$, whereas a system may consist of a combination of many objects or systems each given by a different $R_{\hat{I}}$. Briefly we are saying that we do not consider problems arising from the interconnection of systems when we use the name "object."

CON.I.2.2: Def. 1 of a system allows only time functions as inputs and outputs. We think that it would cause no real difficulties, to allow distributions in our input and output spaces, excepting possibly some philosophical arguments that we will try to discuss in the Appendix (see A.1.). Thus we will refer to the elements of the input and output spaces as inputs and outputs meaning distributions or functions, and I-0 will be an abbreviation for "input-output pair."

CON.I.2.3: By an OBJECT we will always understand a "continuous-time, uniform oriented, object. We will denote it by "$\mathcal{O}$," its unique existence interval by $\hat{I}$. will be given by its I-0 list $R_{\hat{I}}$.

We close this section with the following important definition:

DEF.I.2.4:

The object $\mathcal{O}$ is called DETERMINATE iff for each $u \in U_{\hat{I}}$ there is a unique $y \in Y_{\hat{I}}$ such that $(u,y) \in R_{\hat{I}}$. It is said to be NONANTICIPATIVE iff for any $I \in T$ that starts where $\hat{I}$ does and for any u, $u' \in U_{\hat{I}}$ satisfying $u/_I = u'/_I$ there always exists pairs (u,y) and $(u',y') \in R_{\hat{I}}$ such that $y/_I = y'/_I$. Finally $\mathcal{O}$ is said to be CAUSAL iff it is determinate and nonanticipative.

I.3--Previous Work on the Subject

We start with Zadeh's state axioms [ZA 2-3]:

Zadeh's STATE AXIOMS:

The STATE DESCRIPTION of the object $\mathcal{O}$, given by the list R_I of I-O pairs, is the pair $(\Sigma, \bar{A})$ that satisfies the conditions listed below. Here Σ is a set called the STATE SPACE and $\bar{A}$ a relation called the INPUT-OUTPUT-STATE-RELATION (which will be abbreviated as I-O-S-R). More precisely, $\bar{A}$ is a subset of $\{(I,\sigma,u,y): I \subset \hat{I},\ \sigma \in \Sigma,\ (u,y) \in R_{\hat{I}}\}$. The axioms are:

(M1)--For each $I \subset \hat{I}$, $(u,y) \in R_I$ iff $\exists \sigma \in \Sigma \ni (I,\sigma,u,y) \in \bar{A}$.

(S1)--For each $I \subset \hat{I}$, $\sigma \in \Sigma$ and $u \in U_I$ $\exists$exactly one $y \ni (I,\sigma,u,y) \in \bar{A}$.

Denoting by $\bar{A}_I(\sigma,u)$ the unique response guaranteed by (S1), we can define a family of single valued INPUT-OUTPUT-STATE FUNCTIONS $\bar{A}_I : D_I \to Y_I$, for $I \subset \hat{I}$, which completely characterizes the I-O-S-R. The domain of $\bar{A}_I$ is $D_I \triangleq \Sigma \times U_I$.

(S3)--For each $(t_0,t_1] \subset \hat{I}$ and $(\sigma_0,u_0) \in D_{(t_0,t_1]}$, $\exists$ at least one $\sigma_1 \in \Sigma$ with the property that if $(\sigma_o,u_0 0 u_1) \in D_{(t_0,t_2]}$ then $(\sigma_1,u_1) \in D_{(t_1,t_2]}$ and $\bar{A}_{(t_1,t_2]}(\sigma_1,u_1) = \bar{A}_{(t_0,t_2]}(\sigma_0,u_0 0 u_1)/(t_1,t_2]$.

NOTE I.3.1: The MUTUAL CONSISTENCY CONDITION (M1) establishes the relation of the object $\mathcal{O}$ to the state description. The first of the two SELF CONSISTENCY CONDITIONS (S1) and (S3) guarantees the uniqueness of the output for a given input and state, the property that we were after, from the beginning; the second one classifies the states of the description at time t_1.

NOTE I.3.2: To require D_I to be $\Sigma \times U_I$ for each I was shown to be a very important shortcoming by Resh [RE 1-2]. That $D_I \triangleq \Sigma \times U_I$ means no matter what state the system is left in, one can apply any input. Many existing systems, however, do not admit this property. To the examples given by Resh, that extend from the systems of the type homosapiens and certain kinds of inputs termed propaganda, to the very technical one given by Fig. 1, and that include examples such as rocket engines whose fuels can be depleted by the initial input segments, we can add the example of the brain of an animal which became blind as a result of blast (an input). Any form of light, for that matter, any video-input at this state of the system (the animal) are simply not admissible to the brain.

In Fig. 1, the switch closes exactly one second after the applied input u exceeds 1 volt in magnitude and remains closed thereafter. Thus, an input that was admissible before the closing of the switch is not admissible anymore, since only 0 volt can be applied once the switch has been closed. As in the above example the system is left in such a state that all inputs, except 0, are no longer admissible.

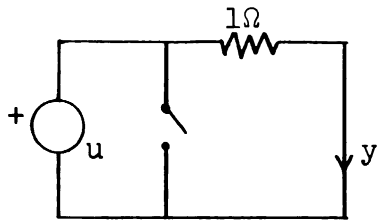


Figure I.3.2

NOTE I.3.3: As it is not desirable to deny a state description to such a large class of objects, since the importance of system theory lies in its abstract generality (page 5), Resh modified axiom (S1), the source of the shortcoming, to read:

(S1')--For each $I \subset \hat{I}$, $\sigma \in \Sigma$ and $u \in U_I$, $\exists$ at most one y $\ni$ $(I, \sigma, u, y) \in \bar{A}$.

This means that the domain D_I of $\bar{A}_I$ is a subset of $\Sigma \times U_I$ consisting of pairs (σ, u) for which there exists a y such that $(I, \sigma, u, y) \in \bar{A}$.

NOTE I.3.4: Unfortunately the replacement of (S1) by (S1'), while eliminating the above shortcoming, introduced some other inconveniences:

All systems had a trivial state description, where the state space Σ' was the set $Y_{\hat{I}}$ of <u>all</u> outputs defined on the existence interval of the object, and the I-O-S-R, $\bar{A}'$ was $\{(I,\sigma,u,\sigma/I): I\subset\hat{I},\ \sigma\in\Sigma'$ and $u = \hat{u}/I$ for some $\hat{u} \ni (\hat{u},\sigma)\in R_{\hat{I}}\}$.

As a result of the state description (Σ',A'), a unit resistor had two reduced (Def. II.2.4) state descriptions: one with a unit state space, the other with a gigantic state space, the set of all outputs of the resistor.

To eliminate the difficulties caused by the change of (S1) to S1') Resh proposed a SECOND MUTUAL CONSISTENCY AXIOM, in two different ways that are not exactly equivalent, to be incorporated in the modified set. Since there are some minor changes in the language of presentation, we present the two axiom sets, proposed in [RE3] and [RE1] respectively:

Let for each t, a set $\Sigma(t)$ be assigned to the object $\mathcal{O}$ as a <u>conjectured state space</u> of $\mathcal{O}$ at time t. Let $\bar{A}$, a subset of $\{(I,\sigma_0,u,y): I = (t_0,t]\subset\hat{I},\sigma_0\in\Sigma(t_0), u\in U_I,y\in Y_I\}$, be the <u>conjectured I-O-S-R</u> of $\mathcal{O}$ meaning: $(I,\sigma_0,u,y)\in\bar{A}$ implies the object in state σ_0 at time t_0 subject to the input u from t_0 to t will respond by producing the output y from t_0 to t. $(\Sigma,\bar{A})$ will be a valid state description iff the following four conditions are satisified:

<u>FIRST AXIOM SET</u> (denoted $\mathcal{A}1$):

(M1)--For each $I = (t_0,t]\subset\hat{I}$, $(u,y)\in R_I \Longleftrightarrow \exists\sigma_0\in\Sigma(t_0)$ $\ni(I,\sigma_0,u,y)\in A$

(M2)--$\Sigma(\hat{t}_0)$ is a unit set, where $\hat{I} = (\hat{t}_0,\hat{t}_1]$ is the existence interval

(S1)--For each $I = (t_0,t]\subset\hat{I}$, $\sigma_0\in\Sigma(t_0)$ and $u\in U_I$, $\exists$ at most one $y\in Y_I\ni(I,\sigma_0,u,y)\in A$.

(S2)--Letting $D_I = \{(\sigma_0,u) : \exists\, y\in Y_1 \ni (I,\sigma_0,u,y)\in A\}$ then defining $\bar{A}_I:D_I\to Y_I$ by $\bar{A}_I(\sigma_0,u) = y$ it is required that: for each $I_0 = (t_0,t_1]$ and (σ_0,u_0) $\in D_{(t_0,t_1]}$ there exists at least one $\sigma_1\in\Sigma(t_1)$ $\ni$:

$$(\sigma_0,u_0 0 u)\in D_{(t_0,t]} \Longrightarrow \begin{cases} (\sigma_1,u)\in D_{(t_1,t]} \\ \qquad\text{and} \\ \bar{A}_{(t_1,t]}(\sigma_1,u) = \\ \\ \bar{A}_{(t_0,t]}(\sigma_0,u_0 0 u)/_{(t_1,t]} \end{cases}$$

<u>SECOND AXIOM SET</u> (denoted $\mathcal{A}2$):

(M1), (S1) and (S2) remain unaltered but (M2) takes the form:

(M2')--For each to $\in(\hat{t}_0,\hat{t}_1] = \hat{I}$ and each $u_0\in U_{(\hat{t}_0,t_0]}$, $\exists\sigma_0\in\Sigma_{(t_0)}\ni$

$$y\in\bar{A}_{(t_0,t]}(\sigma_0,u) \Longleftrightarrow \exists y_0\in Y_{(\hat{t}_0,t_0]}\ni(u_0 0u, y_0 0y)\in R_{(\hat{t}_0,t]}$$

With the introduction of (M2) (or (M2')), the trivial state description (Σ',A') given on page 18 was not a state description anymore [RE2].

NOTE I.3.5: In order to state some more results, we have to introduce the very important concept of equivalence classes of inputs [RE3] which has a strong connection to the past history of the system and which, later, turns out to be a very useful tool in the computation of reduced state descriptions.

DEF.I.3.1:

Let the object $\mathcal{O}$ be given with the list $R_{\hat{I}}$, $\hat{I} = (\hat{t}_0,\hat{t}_1]$. The inputs $u_0\in U_{(\hat{t}_0,t_0]}$ and $u_0'\in U_{(\hat{t}_0,t_0]}$ are said to be EQUIVALENT, denoted $u_0\simeq u_0'$, iff:

(i) $u_0 0w$ is admissible $\Longleftrightarrow u_0' 0w$ is admissible, for $w\in U_{(t_0,\hat{t}_1]}$

(ii) In case $u_0 0w$ and $u_0' 0w$ are admissible, $\exists$ y and y' $\ni$ $(u_0 0w,y)$ and $(u'_0 0w,y')\in R_{\hat{I}}$ then $y/_{(t_0,\hat{t}_1]}$ must equal $y'/_{(t_0,\hat{t}_1]}$

It is trivial to verify that "$\simeq$" is an equivalence relation. Therefore we define: $H_{t_0}[u]\triangleq\{u'\in U_{(\hat{t}_0,t_0]} : u'\simeq u\}$ which are mutually exclusive, collectively inclusive EQUIVALENCE CLASSES OF INPUTS and $\mathcal{H}_{t_0}\triangleq\{H_{t_0}[u] : u\in U_{(\hat{t}_0,t_0]}\}$ as the FAMILY of equivalence classes of inputs.

$$D_{HR(t_0,t]} \triangleq \{(\sigma_0,u)\in D_{(t_0,t]} : \sigma_0\in\Sigma_{HR}(t_0)\}$$ and

$$\bar{A}_{HR(t_0,t]} \triangleq \bar{A}_{(t_0,t]}/D_{HR(t_0,t]}$$ for each

$(t_0,t]$.

DEF.II.2.5:

$(\Sigma_{HR},\bar{A}_{HR})$ will be called a HALF REDUCED STATE DESCRIPTION under $\mathcal{A}i$ iff it satisfies $\mathcal{A}i$, i = 1, 2, 3.

NOTE II.2.3: Under State Axioms $\mathcal{A}1$, there is nothing to guarantee that $(\Sigma_{HR},\bar{A}_{HR})$ or $(\Sigma_R,\bar{A}_R)$ is still a State Description under $\mathcal{A}1$. However under $\mathcal{A}2$ (and $\mathcal{A}3$) this is not the case. M2' in $\mathcal{A}2$ (and S2" in $\mathcal{A}3$) guarantees us the existence of enough non singular states, so that $(\Sigma_R,\bar{A}_R)$ and $(\Sigma_{HR},\bar{A}_{HR})$ are still valid State Descriptions.

NOT. II.2.4: ${}_{t_0}u_{t_1}v_{t_2}z_{t_3}$ will indicate an input (or an output) which consists of segments u defined on $(t_0,t_1]$, v defined on $(t_1,t_2]$ and z defined on $(t_2,t_3]$. $u_{t_1}v$ will mean $t_0 = \hat{t}_0$ and $t_2 = \hat{t}_1$ where the existence interval $\hat{I} = (\hat{t}_0,\hat{t}_1]$. Now we give a State Description for a very simple object, getting payment for the effort; the payments being discussed after the example, the effort is made right now.

DEF.I.3.2:

Here we define a special state description $(\Sigma^*, \bar{A}^*)$ as follows:

$\Sigma^*(t_0)$, for each t_0, is any set of the same cardinality as the family $\mathcal{H}_{t_0}$. $\Sigma^*(\hat{t}_0)$, for $\hat{I} = (\hat{t}_0, \hat{t}_1]$ is any unit set.

To define $\bar{A}^*$, we choose a function C_{t_0} : $\Sigma^*(t_0) \xrightarrow[\text{onto}]{1-1} \mathcal{H}_{t_0}$, that such a function exists is guaranteed by the choice of $\Sigma^*(t_0)$. $(I,\sigma_0,u,y) \in \bar{A}^*$ with $I = (t_0,t]$ iff: for $t_0 > \hat{t}_0$, $\exists (u_0,y_0) \in R_{(\hat{t}_0,t]}$ $\ni u_0/_{(\hat{t}_0,t_0]} \in C_{t_0}(\sigma_0)$ and $(u_0/_I, y_0/_I) = (u,y)$ and for $t_0 = \hat{t}_0$, σ_0 is the single element of $\Sigma^*(\hat{t}_0)$ and (u,y) is arbitrary in R_I.

The results can be summarized in the following two theorems:

THM.I.3.1: If an object has a state description under state axioms $\mathcal{A}2$ then it is causal.

PROOF: [RE1]

THM.I.3.2: The following statements are all equivalent:

(i) $\mathcal{O}$ has a state description under $\mathcal{A}1$

(ii) $\mathcal{O}$ is causal

(iii) (Σ^*, A^*) is a state description under $\mathcal{A}1$.

PROOF: (i)$\Longrightarrow$(ii) [RE2]

(ii)$\Longrightarrow$(iii) [RE3] (in this reference, to show that a causal object always has state description it is proved that (Σ^*, A^*) satisfies $\mathcal{A}1$ for causal objects).

(iii)$\Longrightarrow$(i) is trivial.

<u>NOTE I.3.6</u>: Thus, the state axioms $\mathcal{A}1$ and $\mathcal{A}2$, compared to Zadeh's axioms extending their domain of applicability to such important classes of existing physical systems as in the various examples of pages 16 and 17, have denied state descriptions to non-causal systems, non existing physical systems. However, this is a point much in favor of state axioms $\mathcal{A}1$ and $\mathcal{A}2$ since we can, without hesitation, qualify them as being "more realistic."

<u>CON.I.3.1</u>: Here we make the distinction between A_I and $\bar{A}_I$. As we said earlier, $\bar{A}_I$ denotes the function from D_I into Y_I whereas A_I or $A_I(t)$ denotes the values that the function $\bar{A}_I$ takes on i.e., $y = \bar{A}_I(\sigma,u)$ but $y(t) = A_I(\sigma,u)$.

<u>NOTE I.3.7</u>: When we write $\hat{I} = (\hat{t}_0, \hat{t}_1]$, the case $\hat{I} = (-\infty,\infty)$ is also included.

Finally, State Axioms of Kalman listed merely for completeness close this section.

Kalman's <u>STATE AXIOMS</u>: [KA]

A dynamical system is a mathematical structure defined by the following axioms:

(D1)--There is a given STATE SPACE Σ and a set of values of time Θ at which the behavior of the system is defined: Σ is a topological space and Θ is an ordered topological space which is a subset of the real numbers.

(D2)--There is given a topological space Ω of functions of time defined on Θ, which are the admissible INPUTS to the system.

(D3)--For any initial time $t_0 \in \Theta$, any initial state $\sigma_0 \in \Sigma$ and any input $u \in \Omega$ defined for $t \geqslant t_0$, the future states of the system are determined by the transition function $\phi : \Omega x \Theta x \Theta x \Sigma \rightarrow \Sigma$ which is written as $\phi_u(t;t_0,\sigma_0) = \sigma$. This function is defined only for $t \geqslant t_0$. Moreover, any $t_0 \leqslant t_1 \leqslant t_2$ in Θ, any $\sigma_0 \in \Theta$, and any fixed $u \in \Omega$ defined over $[t_0, t_1] \cap \Theta$ the following relations hold:

(D3-i)--$\phi_u(t_0;t_0,\sigma_0) = \sigma_0$

(D3-ii)--$\phi_u(t_2;t_0,\sigma_0) = \phi_u(t_2;t_1,\phi_u(t_1,t_0,\sigma_0))$

In addition, the system must be NONANTICIPATORY, i.e., if $u,v \in \Omega$ and $u \equiv v$ on $[t_0,t_1] \cap \Theta$ we have

(D3-iii)--$\phi_u(t;t_0,\sigma_0) = \phi_v(t;t_0,\sigma_0)$

(D4)--Every output of the system is a function Ψ : $\Theta \times \Sigma \rightarrow \mathbb{R}$

(D5)--The functions ϕ and Ψ are continuous, with respect to the topologies defined for Σ, Θ and Ω and the induced product topologies.

I.4--A Brief Summary of the Following Chapters

In Chapter II, after introducing some new concepts and modifying some old ones and after presenting an example of deficiency, we conclude that a minor change is necessary in the axiom sets $\mathcal{A}1$ and $\mathcal{A}2$, obtaining axiom set $\mathcal{A}3$ which is stated, for matters of presentation, at the beginning of Sec. II.2. Then, in Sec.II.3 we discuss the interrelations of $\mathcal{A}1$, $\mathcal{A}2$ and $\mathcal{A}3$, and show that $\mathcal{A}3$ is almost equivalent to $\mathcal{A}2$. Finally, Sec. II.4 concentrates on reduced and half reduced state descriptions, yielding important results, for a given object $\mathcal{O}$, such as: the cardinality of any two reduced state space is the same, or any reduced state description is nothing but (Σ^*, A^*) obtained by use of equivalence classes of inputs (DEF.I.3.2), etc.

In Chapter III, we investigate how the properties of the object $\mathcal{O}$--its linearity, time-invariance--are reflected in the properties of its state space. We show that the state space can be constructed to possess corresponding nice properties. An important point about this chapter is that the properties of the system are defined, not in terms of its state description, but rather in terms of its I-O pairs, and then their implications on the state space deduced.

Considering linear, time invariant and continuous objects in Chapter IV, the use of convolutional

representation for such objects is justified, and a state description of the form:

$$\frac{dx_n(t)}{dt} = \sum_{m=1}^{\infty} a_{mn} x_m(t) + b_n u(t) \qquad \frac{dX(t)}{dt} = A\, X(t) + Bu(t)$$

i.e.

$$y(t) = \sum_{n=1}^{\infty} c_n x_n(t) + \sum_{k=0}^{K} d_k u^{(k)}(t) \qquad y(t) = C\, X(t) + \sum_{k=0}^{K} d_k u^{(k)}(t)$$

is given for a large class of distributed systems respectively in sections 2 and 3, where X(t) is a square summable sequence for each t and A an infinite Hilbert Matrix. In the final section of this chapter, the important problem of "approximating a system having a continuum of states with objects having a finite dimensional vector space as their state space" is discussed and solutions offered.

Finally in the Appendix, Chapter V being "the conclusions" chapter, first a justification for using distribution theory, then the "Orthonormal Series Expansions of Distributions," recently developed by Zemanian and others, is given in its general lines. Thirdly some new theorems that are necessary for Chapter IV, such as the convolution of distributions in $\mathfrak{A}'$, the proof that shows certain types of functions are in $\mathfrak{A}$ are presented.

CHAPTER II

AXIOM SETS $\mathcal{A}1$, $\mathcal{A}2$, $\mathcal{A}3$ AND STATE DESCRIPTIONS IN GENERAL

II.1--Introduction

This chapter sets the basic rules, matures the necessary background and develops some very useful tools to be used in Chapters III and IV. Many theorems are proved about State Descriptions in closed form, few of which may be considered as ends by themselves. We consider this chapter of prime importance for the rest of the work and apologize for some long and tedious proofs.

In section 2 we define certain important concepts such as Reachable States, Singular States, Equivalent States, Reduced and Half Reduced State Descriptions, etc., some of which are new, some of which are the modifications of the old ones, in the light of the new Axiom Sets.

An example in the same section shows the insufficiency of the State Axioms $\mathcal{A}1$ and that $\mathcal{A}1$ is not equivalent to $\mathcal{A}2$. To remedy the situation, a modification is introduced to $\mathcal{A}1$ giving rise to $\mathcal{A}3$. The latter, besides being justified physically, deserves attention because of its consequences.

In section 3, we deal mostly with formalities of investigating the interrelations of the Axiom Sets and prove that most former results do still hold under $\mathcal{A}3$. One of these surviving results is the very useful and important State Description of an object, based on the equivalence classes of inputs.

In section 4, we investigate and bring to light the nice properties of Reduced and Half Reduced State Descriptions. We show that Reduced State Descriptions are basically unique and strongly related to equivalence classes of inputs. It is here that we obtain the result "any two Reduced State Spaces for a given object have the same cardinality" which is an end by itself.

Thus briefly section 2 sets the basic rules, section 3 matures the necessary background and section 4 develops the useful tools to be used later.

II.2--An Example of Deficiency and the Axiom Set $\mathcal{A}3$.

Because of the new State Axioms $\mathcal{A}1$, $\mathcal{A}2$ it is necessary to revise the definitions of some important concepts. Some of the subsequent definitions are modifications and some are new. No explicit reference being made with respect to which Axiom Set they are given, they remain the same for $\mathcal{A}1$, $\mathcal{A}2$ and $\mathcal{A}3$ ($\mathcal{A}3$ to be introduced later). Let in the following $(\Sigma,\bar{A})$ be a State Description of $\mathcal{O}$ given by $R_{\hat{I}}$.

DEF.II.2.1:

A state $\sigma_1 \in \Sigma(t_1)$ is said to be REACHABLE FROM A STATE $\sigma_0 \in \Sigma(t_0)$, $t_1 > t_0$ if there exists an input $u_1 \in U_{(t_0,t_1)}$ such that:

(i) $(\sigma_0, u_1) \in D_{(t_0,t_1]}$

(ii) $(\sigma_0, u_1 0 u) \in D_{(t_0,t]} \Longleftrightarrow (\sigma_1, u) \in D_{(t_1,t]}$, for $u \in U_{(t_1,t]}$

(iii) $\bar{A}_{(t_0,t]}(\sigma_0, u_1 0 u)/_{(t_1,t]} = A_{(t_0,t]}(\sigma_1, u)$

Then "the input $u_1 \in U_{(t_0,t_1]}$ takes the state σ_0 into the state σ_1."

DEF.II.2.2:

A state $\sigma_1 \in \Sigma(t_1)$ is SINGULAR iff it is not reachable from a state $\sigma_0 \in \Sigma(\hat{t}_0)$. Or equivalently: a state $\sigma_1 \in \Sigma(t_1)$ is NON-SINGULAR iff it is reachable from a state $\sigma_0 \in \Sigma(\hat{t}_0)$, where $\hat{I} = (\hat{t}_0, \hat{t}_1]$ is the existence interval.

DEF.II.2.3:

A state $\sigma_0' \in \Sigma(t_0)$ is SUBSUMED by the state $\sigma_0'' \in \Sigma(t_0)$ iff:

(i) $(\sigma_0', u) \in D_{(t_0,t]} \Longrightarrow (\sigma_0', u) \in D_{(t_0,t]}$, for all $t > t_0$ in $\hat{I}$

(ii) $\bar{A}_{(t_0,t]}(\sigma_0', u) = \bar{A}_{(t_0,t]}(\sigma_0'', u)$ holds for all u described in (i).

Two states $\sigma_0'\in\Sigma(t_0)$ and $\sigma_0''\in\Sigma(t_0)$ are EQUIVALENT iff σ_0' is subsumed by σ_0'' and σ_0'' is subsumed by σ_0', i.e.,

(i) $(\sigma_0',u)\in D_{(t_0,t]} \Longleftrightarrow (\sigma_0'',u)\in D_{(t_0,t]}$

and $t\in\hat{I}$, and

(ii) $\bar{A}_{(t_0,t]}(\sigma_0',u) = \bar{A}_{(t_0,t]}(\sigma_0'',u)$ is true for all u described in (i).

We believe that these definitions are self explanatory and need no further justification or physical interpretation. We now prove a simple fact.

<u>FACT II.2.1</u>: Let $\sigma_2\in\Sigma(t_2)$ be reachable from $\sigma_1\in\Sigma(t_1)$ and $\sigma_1\in\Sigma(t_1)$ be reachable from $\sigma_0\in\Sigma(t_0)$. Then $\sigma_2\in\Sigma(t_2)$ is reachable from $\sigma_0\in\Sigma(t_0)$.

<u>PROOF</u>: σ_1 is reachable from σ_0 and σ_2 is reachable from σ_1 implies respectively that there is an input $u_1\in U_{(t_0,t_1]}$ that takes σ_0 into σ_1 and another input $u_2\in U_{(t_1,t_2]}$ that takes σ_1 into σ_2. Now we claim that:

$u_1 0 u_2$ is admissible. Since $(\sigma_1,u_2)\in D_{(t_1,t_2]} \Longleftrightarrow (\sigma_0,u_1 0 u_2)\in D_{(t_0,t_2]}$, which can happen only if $u_1 0 u_2\in U_{(t_0,t_2]}$.

$u_1 0 u_2$ is an input that takes σ_0 into σ_2:

(i) $(\sigma_0,u_1 0 u_2)\in D_{(t_0,t_2]}$ from above,

(ii) $(\sigma_0,u_1 0 u_2 0 u)\in D_{(t_0,t]} \Longleftrightarrow (\sigma_1,u_2 0 u)\in D_{(t_1,t]}$

$\Longleftrightarrow (\sigma_2,u)\in D_{(t_2,t]}$ for $u\in U_{(t_2,t]}$

(iii) $\bar{A}_{(t_0,t]}(\sigma_0,u_1 0 u_2 0 u)/_{(t_2,t]} =$

$\{\bar{A}_{(t_0,t]}(\sigma_o,u_1 0 u_2 0 u)/_{(t_1,t]}\}/_{(t_2,t]} =$

$\bar{A}_{(t_1,t]}(\sigma_1,u_2 0 u)/_{(t_2,t]} =$

$\bar{A}_{(t_2,t]}(\sigma_2,u)$ proving the fact.

Another simple fact that can be proved easily is the following one:

FACT II.2.2: The equivalence of the states defined in Def.II.2.3 is an equivalence relation and thus partitions the state space $\Sigma(t)$ into equivalence classes of states for each $t \in \hat{I}$.

Now we proceed to the "CONSTRUCTION OF REDUCED STATE DESCRIPTION."

NOT.II.2.1: Let $(\Sigma_2,\bar{A}_2)$ be a State Description of $\mathcal{O}$ under the Axiom Set $\mathcal{A}2$. In this case the Reduced State Description has to be obtained in two steps, as compared to a State Description under $\mathcal{A}1$ (or $\mathcal{A}3$), because of (M2') that allows more than one state at the creation instant. We obtain a new State Description $(\Sigma,\bar{A})$ from $(\Sigma_2,\bar{A}_2)$ proceeding as follows:

$\Sigma(t) \triangleq \Sigma_2(t) \quad \forall t > \hat{t}_0$, $\hat{t}_0$ the creation instant.

$\Sigma(\hat{t}_0) \triangleq$ any unit set.

$D_{(t_0,t]} \triangleq D_{2(t_0,t]} \quad \forall t_0 > \hat{t}_0$

$D_{(\hat{t}_0,t]} \triangleq \{(\sigma_{\hat{t}_0},u) : \sigma_{\hat{t}_0} \in \Sigma(\hat{t}_0)$ and $u \in U_{(\hat{t}_0,t]}\}$

$\bar{A}_{(t_0,t]} \triangleq \bar{A}_{2(t_0,t]}$ $\forall t_0 > \hat{t}_0$. $\bar{A}_{(\hat{t}_0,t]} \triangleq$
$\{((\hat{t}_0,t],\ \sigma_{\hat{t}_0},u,y) : \sigma_{\hat{t}_0}\in\Sigma(\hat{t}_0),\ (u,y)\in R_{(\hat{t}_0,t]}\}$

<u>NOTE II.2.1</u>: The definition of $\bar{A}_I$, makes sense for $I = (\hat{t}_0,t]$. For, to each $u\in U_{(\hat{t}_0,t]}$, there corresponds a unique $y\in Y_{(\hat{t}_0,t]}$, since $\mathcal{O}$ has a State Description under $\mathcal{A}2$ and must therefore be causal by Thm.I.3.1.

<u>NOTE II.2.2</u>: The pair $(\Sigma,\bar{A})$ obtained from $(\Sigma_2,\bar{A}_2)$, as explained in Not.II.2.1, is a State Description of $\mathcal{O}$, under $\mathcal{A}2$.

(M1), (M2'), (S1) are trivially true since $(\Sigma_2,\bar{A}_2)$ satisfies $\mathcal{A}2$, and by Note II.2.1, (S2) is also true for $I_0 = (t_0,t_1]$, $t_0 > \hat{t}_0$, by the same reasons. For $I_0 = (\hat{t}_0,t_1]$ and $(\sigma_{\hat{t}_0},u_0)\in D_{(\hat{t}_0,t_1]}$ let $\sigma_1\in\Sigma(t_1)$, as required by (S2'), to be the state $\sigma_1\in\Sigma(t_1)$ guaranteed by (M2'). Then:

$$(\sigma_{\hat{t}_0},u_0 0 u)\in D_{(\hat{t}_0,t_2]} \iff u_0 0 u\in U_{(-\infty,t_2]}$$

$$\iff \exists\, y_0 0 y\in Y_{(\hat{t}_0,t_2]} \ni (u_0 0 u, y_0 0 y)\in R_{(\hat{t}_0,t_2]}$$

$$\iff y = \bar{A}_{2(t_1,t_2]}(\sigma_1,u)$$

proving (S2).

NOT.II.2.2: Let $(\Sigma, \bar{A})$ be a State Description of $\mathcal{O}$ under $\mathcal{A}1$, $\mathcal{A}2$ (or $\mathcal{A}3$), (if under $\mathcal{A}2$ the description contains more than one state at $\hat{t}_0$, we apply the procedure of Not.II.2.1. to obtain one that possesses a unique state). Then for each t, we form:

$\Sigma_s(t) \triangleq \{\sigma \in \Sigma(t) : \sigma \text{ is singular}\}$

$\Sigma_\sigma(t) \triangleq \{\sigma' \in \Sigma(t) : \sigma' \simeq \sigma\} \cap \Sigma_s^c(t)$, the complement being with respect to $\Sigma(t)$. Note that the classes $\Sigma_\sigma(t)$ are mutually exclusive.

$\Sigma_R(t) \triangleq$ a subset of $\Sigma(t)$ obtained by taking only one element from each class $\Sigma_\sigma(t)$.

$D_{R(t_0,t]} \triangleq \{(\sigma_0,u) \in D_{(t_0,t]} : \sigma_0 \in \Sigma_R(t_0)\}$ for each interval $(t_0,t]$

$\bar{A}_{R(t_0,t]} \triangleq \bar{A}_{(t_0,t]}/D_{R(t_0,t]}$ for each interval $(t_0,t]$.

DEF.II.2.4:

$(\Sigma_R, \bar{A}_R)$ will be called a REDUCED STATE DESCRIPTION UNDER $\mathcal{A}i$, i = 1, 2, 3 iff it satisfies the Axiom Set $\mathcal{A}i$, i = 1, 2, 3.

NOT.II.2.3: Let $(\Sigma, \bar{A})$ be a State Description of $\mathcal{O}$ under $\mathcal{A}i$, i = 1, 2, 3. We obtain $\Sigma_s(t)$ as in Not.II.2.2, then we define:

$\Sigma_{HR}(t) \triangleq \Sigma_s^c(t)$ for each t, i.e., we only keep non-singular states for each t in our state space.

AN EXAMPLE OF DEFICIENCY

Let the object $\mathcal{O}$ be given by the list $R_{\hat{I}}$, $\hat{I} = (-\infty,\infty)$, shown in Fig.II.2.1. $R_{\hat{I}} = \{(0_0 0, 0_0 0), (0_0 1, 0_0 0), (1_0 0, 1_0 1), (1_0 0, 1_0 1)\}$.

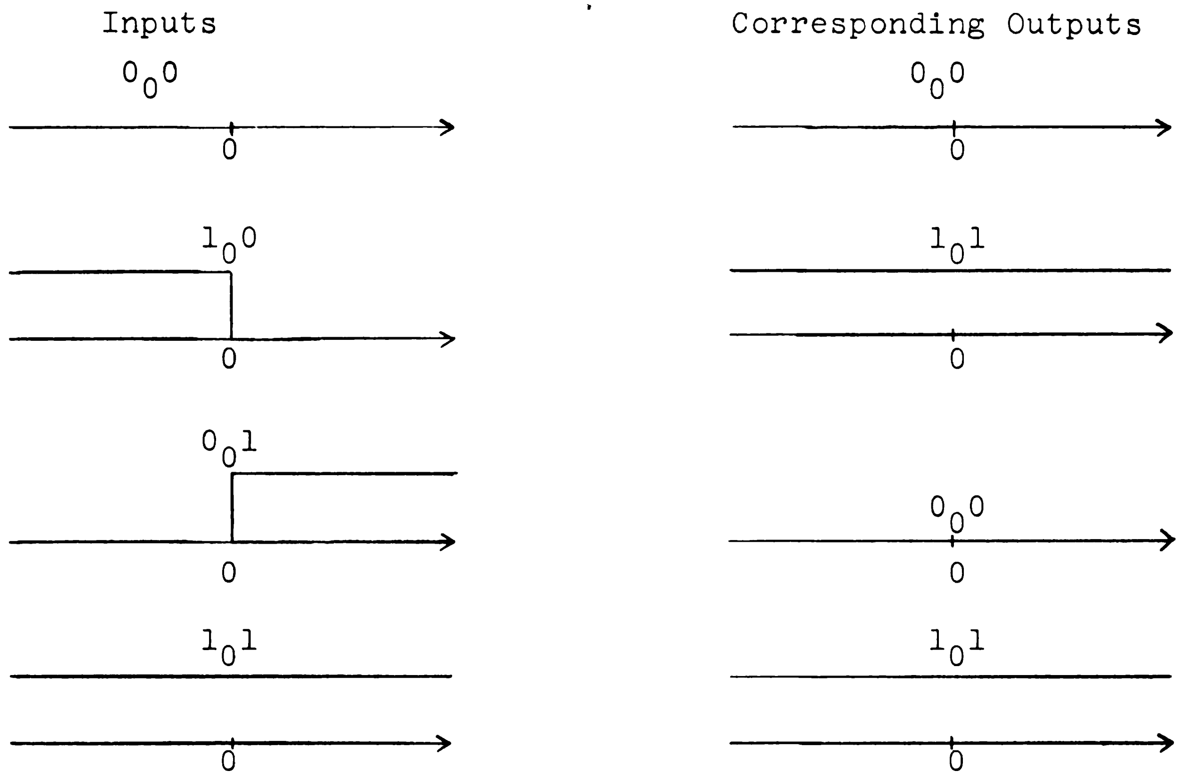


Figure II.2.1.

Let our <u>conjectured state space</u> be:

$$\Sigma(-\infty) = \{\sigma\}$$
$$\Sigma(t) = \{\sigma_1, \sigma_2\} \qquad -\infty < t \leq 0$$
$$\Sigma(t) = \{\alpha, \beta, \gamma, \delta\} \qquad t > 0$$

together with the <u>conjectured I-O-S-R</u> which is defined as follows for $I = (t_0, t]$:

For

$$
\begin{array}{lll}
t_0 = -\infty & \cdot t \leqslant 0, & \bar{A} = \{(I,\sigma,0,0),\ (I,\sigma,1,1)\} \\
 & \cdot t > 0, & \bar{A} = \{(I,\sigma,0_0 0,0_0 0),\ (I,\sigma,0_0 1,0_0 0), \\
 & & \qquad (I,\sigma,1_0 0,1_0 1),\ (I,\sigma,1_0 1,1_0 1)\} \\
t_0 \leqslant 0 & \cdot t \leqslant 0, & \bar{A} = \{(I,\sigma_1,0,0),\ (I,\sigma_2,1,1), \\
 & & \qquad \underline{(I,\sigma_1,1,1)}\} \\
 & \cdot t > 0, & \bar{A} = \{(I,\sigma_1,0_0 0,0_0 0),\ (I,\sigma_1,0_0 1,0_0 0), \\
 & & \qquad (I,\sigma_2,1_0 0,1_0 1),\ (I,\sigma_2,1_0 1,1_0 1)\} \\
t_0 > 0 & , & \bar{A} = \{(I,\alpha,0,0),\ (I,\beta,1,0),\ (I,\gamma,0,1), \\
 & & \qquad (I,\delta,1,1)\}
\end{array}
$$

We first bring to attention that the need for parametrization (i.e., the need for a State Description) shows up, as discussed in Chapter I, when $I = (t_0,t]$ with $t_0 > 0$. If we were given, e.g., the input 0 on $I = (1, 3]$ we would not know what output to make correspond to it. But if we are given the input 0, with the state α, we now can say that the corresponding output is 0.

A--<u>The conjectured State Description $(\Sigma,\bar{A})$ of $\mathcal{O}$ satisfies $\mathcal{A}1$</u>:

(M1) Let $I = (t_0,t]$, then I must be in one of the five categories, i.e., either ($t_0 = -\infty$, $t \leqslant 0$ or $t > 0$) or ($t_0 \leqslant 0$, $t \leqslant 0$ or $t > 0$) or ($t_0 > 0$). In all the five cases: $(u,y) \in R_I \Longleftrightarrow \exists$ a state (either σ, or one of σ_1,σ_2 or one of $\alpha,\beta,\gamma,\delta$) $\sigma_0 \ni (I,\sigma_0,u,y) \in \bar{A}$.

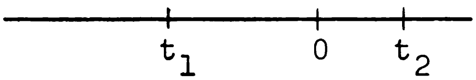
(M2) is satisfied by definition of the state space.

(S1) By checking the list defining $\bar{A}$ one can see that for each $\sigma_0 \in \Sigma(t_0)$, each $u \in U_I$ $\exists$ at most one output $y \ni (I,\sigma_0,u,y) \in \bar{A}$.

(S2) is the most tedious one to check. We must go through all possibilities.

Let $I_0 = (t_0,t_1]$ with its position indicated at the beginning of each case.

$t_0 = -\infty$, $t_1 \leqslant 0$



Consider $(\sigma,0) \in D_{I_0}$, then σ_1 is the state required by (S2). In fact:

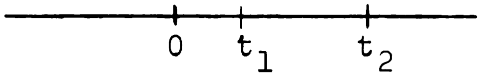
For $t_2 \leqslant 0$, $(\sigma, {}_{t_1}0u_{t_2}) \in D_{(-\infty,t_2]} \Longrightarrow u = {}_{t_1}0_{t_2} \Longrightarrow$ $(\sigma_1, {}_{t_1}0_{t_2}) \in D_{(t_1,t_2]}$ and $\bar{A}_{(t_1,t_2]}(\sigma_1,0) =$ $\bar{A}_{(-\infty,t_2]}(\sigma, {}_{t_1}0_{t_1}0_{t_2})/_{(t_1,t_2]} = 0.$

For $t_2 > 0$ $(\sigma, {}_{t_1}0u_{t_2}) \in D_{(-\infty,t_2]} \Longrightarrow$ either $u =$ ${}_{t_1}0_0 0_{t_2}$ or $u = {}_{t_1}0_0 1_{t_2}$.

In either case $(\sigma_1, {}_{t_1}u_{t_2}) \in D_{(t_1,t_2]}$ and again in either case $\bar{A}_{(t_1,t_2]}(\sigma_1,u) = A_{(-\infty,t_2]}(\sigma, {}_{t_1}0u_{t_2})/_{(t_1,t_2]} = 0.$

For the same I_0, but for $(\sigma,1) \in D_{I_0}$ we can go through the same arguments by replacing "0 inputs" with "1 inputs" and vice versa in the above discussion, the state required by (S2) being σ_2 this time.

$t_0 = -\infty,\ t_1 > 0$

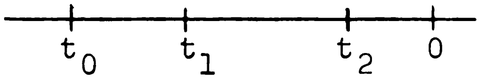


Consider $(\sigma, 0_0 0_{t_1}) \in D_{I_0}$. The state (S2) requires, is $\alpha \in \Sigma(t_1)$. In fact:

$(\sigma, 0_{t_1} u_{t_2}) \in D_{(-\infty, t_2]} \Longrightarrow u = {}_{t_1}0_{t_2}$ since that is the only possible concatenation. Then $(\alpha, {}_{t_1}0_{t_2}) \in D_{(t_1, t_2]}$ and $\bar{A}_{(t_1,t_2]}(\alpha, 0) = \bar{A}_{(-\infty,t_2]}(\sigma, 0_{t_1} 0_{t_2})/(t_1,t_2] = 0$.

For the three possibilities $(\sigma, 0_0 1_{t_1})$, $(\sigma, 1_0 0_{t_1})$ and $(\sigma, 1_0 1_{t_1}) \in D_{I_0}$ the states required by (S2) are respectively β, γ and δ, the proof remaining the same.

$t_0 \leqslant 0,\ t_1 \leqslant 0$



Let us take $(\sigma_1, {}_{t_0}0_{t_1}) \in D_{I_0}$. The state required by (S2) is $\sigma_1 \in \Sigma(t_1)$. Since:

For $t_2 < 0$, $(\sigma_1, {}_{t_0}0_{t_1} u_{t_2}) \in D_{(t_0,t_2]} \Longrightarrow u = {}_{t_1}0_{t_2} \Longrightarrow$

$(\sigma_1, {}_{t_1}0_{t_2}) \in D_{(t_1,t_2]}$ and $\bar{A}_{(t_1,t_2]}(\sigma_1, 0) =$

$\bar{A}_{(t_0,t_2]}(\sigma_1, {}_{t_0}0_{t_1} 0_{t_2})/(t_1,t_2] = 0.$

For $t_2 > 0$, $(\sigma_1, {}_{t_0}0_{t_1} u_{t_2}) \in D_{(t_0,t_2]} \Longrightarrow$ either

$u = {}_{t_1}0_0 0_{t_2}$ or $u = {}_{t_1}0_0 1_{t_2}$.

In either case: $(\sigma_1, u) \in D_{(t_1,t_2]}$ and $\bar{A}_{(t_1,t_2]}(\sigma_1, u) =$

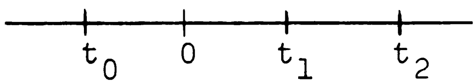
$A_{(t_0,t_2]}(\sigma_1, {}_{t_0}0_{t_1} u_{t_2})/(t_1,t_2] = 0.$

For the same I_0, but either for $(\sigma_1, 1)$ or for $(\sigma_2, 1) \in D_{I_0}$, $\sigma_2 \in \Sigma(t_1)$ this time is the state required by (S2). In fact for $I_0 = (t_0,t_1]$ and $(\sigma_2,1) \in D_{I_0}$, both σ_1 and σ_2 work.

For $t_2 \leqslant 0$, $(\sigma_1, {}_{t_0}1_{t_1}u_{t_2}) \in D_{(t_0,t_2]} \Rightarrow u = {}_{t_1}1_{t_2}$, is the only choice $\Rightarrow$ both $(\sigma_1, {}_{t_1}1_{t_2}) \in D_{(t_1,t_2]}$ and $(\sigma_2, {}_{t_1}1_{t_2}) \in D_{(t_1,t_2]}$ and for both $A_{(t_1,t_2]}(\sigma_1,1) = \bar{A}_{(t_1,t_2]}(\sigma_2,1) = A_{(t_0,t_2]}(\sigma_1, {}_{t_0}1_{t_1}1_{t_2})/(t_1,t_2] = 1$.

For $t_2 > 0$, $\nexists\, u \in U_{(t_1,t_2]} \ni (\sigma_1, {}_{t_0}1_{t_1}u_{t_2}) \in D_{(t_0,t_2]}$, so there is nothing to be checked and (S2) is automatically satisfied.

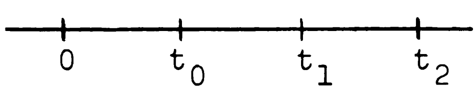
$t_0 \leqslant 0$, $t_1 > 0$



Consider $(\sigma_1, {}_{t_0}0_0 0_{t_1}) \in D_{I_0}$. Then $\alpha \in \Sigma(t_1)$ satisfies the requirements of (S2). For: $(\sigma_1, {}_{t_0}0_0 0_{t_1} 0_{t_2}) \in D_{(t_0,t_2]} \Rightarrow u = {}_{t_1}0_{t_2}$ is the only possible sequel to ${}_{t_0}0_0 0_{t_1}$. Then $(\alpha,0) \in D_{(t_1,t_2]}$ and $\bar{A}_{(t_1,t_2]}(\alpha,0) = A_{(t_0,t_2]}(\sigma_1, {}_{t_0}0_{t_1}0_{t_2})/(t_1,t_2] = 0$.

For the remaining $(\sigma_1, {}_{t_0}0_0 1_{t_1})$, $(\sigma_2, {}_{t_0}1_0 0_{t_1})$, $(\sigma_2, {}_{t_0}1_0 1_{t_1}) \in D_{I_0}$ the states required by (S2) are respectively: β, γ, δ, the proof being the same as above.

$t_0 > 0$



Finally for $(\alpha, {}_{t_0}0_{t_1}) \in D_{I_0}$, α is the state that satisfies (S2). In fact:

$(\alpha, {}_{t_0}0_{t_1}u_{t_2}) \in D_{(t_0,t_2]} \Rightarrow u = {}_{t_1}0_{t_2}$ and thus $(\alpha, {}_{t_1}0_{t_2}) \in D_{(t_1,t_2]}$ with $\bar{A}_{(t_1,t_2]}(\alpha,0) = \bar{A}_{(t_0,t_2]}(\alpha, {}_{t_0}0_{t_1}0_{t_2})/(t_1,t_2] = 0.$ Similarly the states required by (S2) for the pairs $(\beta,1)$, $(\gamma,0)$ and $(\delta,1)$ in D_{I_0} are β, γ, $\delta \in \Sigma(t_1)$ respectively.

B--<u>$(\Sigma,\bar{A})$ is not a State Description of $\mathcal{O}$ under $\mathcal{A}2$</u>:

Since (M1), (S1) and (S2) are the same in $\mathcal{A}1$ and $\mathcal{A}2$, the only possible contradiction can be to (M2'). To show this we consider the case: $t_0 < 0$ and $u_0 = 0 \in U_{(-\infty,t_0]}$, then what must be proved is:

"$\nexists \sigma_0 \in \Sigma(t_0) \ni \{y = \bar{A}_{(t_0,t]}(\sigma_0,u) \Leftrightarrow \exists y_0 \in Y_{(-\infty,t_0]} \ni (u_0 0 u, y_0 0 y) \in R_{(-\infty,t]}\}$."

As $\sigma_0 = \sigma_1$ and $\sigma_0 = \sigma_2$ are the only two possible states at time $t_0 < 0$:

Let first $\sigma_0 = \sigma_1 \in \Sigma(t_0)$. Note that $(\sigma_1,1) \in D_{(t_0,t]}$ for $t < 0$ which implies $y = 1 = \bar{A}_{(t_0,t]}(\sigma_1,1)$. But since t_0 is negative it is clear that $u_0 u = 0_{t_0} 1_t \notin U_{(-\infty,t]}$ and hence there cannot be any $y_0 \ni (u_0 0 u, y_0 0 y) \in R_{(-\infty,t]}$. We conclude that σ_1 is not the state required by (M2').

Let now $\sigma_0 = \sigma_2 \in \Sigma(t_0)$. The only possible input u such that $(\sigma_0,u) \in D_{(t_0,t]}$ is $u = {}_{t_0}1_t$ for $t < 0$.

But again $0_{t_0}1_t$ is not admissible $\Longrightarrow$ σ_2 is not the state required by (M2').

The following Notes discuss the State Axioms $\mathcal{A}i$, $i = 1, 2, 3$ in the light of the above example, providing the promised payments.

<u>NOTE II.2.4</u>: It is clear, since there are no singular, and equivalent states, that the State Description of the object $\mathcal{O}$ in the above example is a reduced State Description. Besides that, our example serves two main purposes:

It shows that the two sets of axioms $\mathcal{A}1$ and $\mathcal{A}2$ are not equivalent.

It proves that the next Axiom Set we are going to define cannot be obtained from $\mathcal{A}1$.

<u>NOTE II.2.5</u>: One could also introduce into the I-O-S-R, $\bar{A}$, many other quadruples, as we did for the case: $t_0 \leqslant 0$, $t \leqslant 0$ and $I = (t_0,t]$ by introducing $(I,\sigma_1,1,1)$, that are not really necessary to get a State Description. For example, for the case, $t_0 \leqslant 0$, $t > 0$ we could include in $\bar{A}$ the quadruples $(I,\sigma_1,1_0 0,1_0 1)$, $(I,\sigma_1,1_0 1,1_0 1)$ and still get a State Description which is valid under $\mathcal{A}1$. This turns out to be a deficiency of $\mathcal{A}1$, which does not occur under State Axioms $\mathcal{A}2$, as we have seen and will see. We call this a deficiency of $\mathcal{A}1$ for two reasons:

Even after including such "superfluous" quadruples in $\bar{A}$, we still have a reduced State Description and can have many different ones by adding such quadruples at will. There is no way of getting rid of these quadruples by throwing out some states.

Whenever we tried to prove the key results of this chapter (such as Thms. II.4.2, II.4.3, II.4.5) we were always stopped by the presence of such superfluous quadruples.

All this trouble owes its presence to (S2) of $\mathcal{A}1$, and was not present in $\mathcal{A}2$ due to (M2') which requires more than (M2). We will give a new Axiom Set $\mathcal{A}3$ which will be justified by its physical interpretation and by its ends.

<u>NOTE II.3.2</u>: Let us suppose an observer wants to experiment at time t_1, t_1 not a creation instant, on the object $\mathcal{O}$ which is left at state $\sigma_1 \in \Sigma(t_1)$. Let us further assume the object $\mathcal{O}$ came to the state σ_1, from a state $\sigma_0 \in \Sigma(t_0)$, $t_0 < t_1$ by an input u_0, i.e., σ_1 is one of the states required by (S2) for the pair $(\sigma_0, u_0) \in D_{(t_0,t_1]}$. Now if the observer can apply the input $u \in U_{(t_1,t_2]}$, i.e., if $(\sigma_1, u) \in D_{(t_1,t_2]}$, that means u can follow u_0, that also means the concatenation $u_0 0 u$, besides being admissible, can be taken as an input pairable with the state $\sigma_0 \in \Sigma(t_0)$. What we are trying to say is that $(\sigma_1, u) \in D_{(t_1,t_2]}$ should imply $(\sigma_0, u_0 0 u) \in D_{(t_0,t_2]}$. This property is not

reflected in (S2) of $\mathcal{A}1$ and constitutes the only change in $\mathcal{A}3$.

We list all of $\mathcal{A}3$ for ease of reference.

THIRD AXIOM SET $\mathcal{A}3$:

(M1)-For each $I = (t_0,t]\subset\hat{I}$, $(u,y)\in R_I$ iff $\exists\,\sigma_0\in\Sigma(t_0)$ $\ni(I,\sigma_0,u,y)\in\bar{A}$.

(M2)-$\Sigma(\hat{t}_0)$ is a unit set, $\hat{I} = (\hat{t}_0,\hat{t}_1]$

(S1)-For each $I = (t_0,t]$, for each $\sigma_0\in\Sigma(t_0)$ and for each $u\in U_I$, there exists at most one output y such that $(I,\sigma_0,u,y)\in\bar{A}$.

(S2')-For each $I_0 = (t_0,t_1]$ and $(\sigma_0,u_0)\in D_{(t_0,t_1]}$ there exists at least one $\sigma_1\in\Sigma(t_1)$.

(i) $(\sigma_0,u_0 0 u)\in D_{(t_0,t]} \Leftrightarrow (\sigma_1,u)\in D_{(t_1,t]}$

(ii) $\bar{A}_{(t_1,t]}(\sigma_1,u) = \bar{A}_{(t_0,t]}(\sigma_0,u_0 0 u)/_{(t_1,t]}$,

$\forall$u that satisfy (i) where D_I and $\bar{A}_I$ are as defined for $\mathcal{A}1$.

II.3--Interrelations of the Axiom Sets $\mathcal{A}1$, $\mathcal{A}2$, $\mathcal{A}3$

NOTE II.3.1: ABOUT $\mathcal{A}1$ and $\mathcal{A}3$. It is obvious that $\mathcal{A}3$ is a restriction of $\mathcal{A}1$, in the sense that any State Description of Θ that satisfies $\mathcal{A}3$, satisfies $\mathcal{A}1$. That the converse is not true is easily established by the Example of II.2:

PROOF of $\mathcal{A}1 \not\Rightarrow \mathcal{A}3$: We take the object Θ of p.34 with the conjectured state description $(\Sigma,\bar{A})$ of p.35. We show this $(\Sigma,\bar{A})$, already a State Description under $\mathcal{A}1$, does not satisfy $\mathcal{A}3$, the contradiction being to (S2'). Consider, $I_0 = (-\infty,t_1]$, $t_1 < 0$ and $(\sigma,0) \in D_{I_0}$

$\sigma_0 = \sigma_1 \in \Sigma(t_1)$ cannot be the state required by (S2') since: $(\sigma_1,1) \in D_{(t_1,t_2]}$ say for $t_1 < t_2 \leqslant 0$, but $(\sigma,0_{t_1}1_{t_2}) \notin D_{(-\infty,t_2]}$ as $0_{t_1}1_{t_2}$ is not an admissible input.

$\sigma_0 = \sigma_2 \in \Sigma(t_1)$ is also not the state required by (S2'), since again: $(\sigma_2,1) \in D_{(t_1,t_2]}$ for $t_1 < t_2 \leqslant 0$ but $0_{t_1}1_{t_2}$ is not admissible.

Thus we conclude, there exists no state at t_1 that satisfies (S2') for the pair $(\sigma,0) \in D_{I_0}$ and therefore $\mathcal{A}1 \not\Rightarrow \mathcal{A}3$.

Finally we would like to note that the same pair $(\Sigma,\bar{A})$, only without the quadruple $(I,\sigma_1,1,1)$ for $t_0 \leqslant 0$, $t \leqslant 0$, $I = (t_0,t]$, is a State Description under $\mathcal{A}1$, $\mathcal{A}2$ and $\mathcal{A}3$ for Θ.

Now before anything else we have to proceed through the formalities of proving the previous results, Thm. I.3.2, under our new Axiom Set $\mathcal{A}3$. In order to do this we demonstrate two simple facts, already mentioned in [RE3] for $\mathcal{A}1$, about the State Description $(\Sigma^*,\bar{A}^*)$ obtained by use of equivalence classes of inputs (Def.I.3.1).

Let in the following the object $\mathcal{O}$ be given by $R_{\hat{I}}$, $\hat{I} = (\hat{t}_0, \hat{t}_1]$ the existence interval.

<u>FACT II.3.1</u>: Two inputs $v \in U_{(\hat{t}_0, t_0]}$ and $u \in U_{(\hat{t}_0, t_0]}$ are equivalent iff:

(i) $\forall t > t_0$ and $\forall w \in U_{(t_0, t]}$, $u_0 w$ is an admissible input iff $v_0 w$ is.

(ii) If for $w \in U_{(t_0, t]}$, $u_0 v$ and $v_0 w$ are admissible and $\hat{y}, \hat{z} \in Y_{(\hat{t}_0, t]}$ are the corresponding outputs then $\hat{y}/_{(t_0, t]} = \hat{z}/_{(t_0, t]}$.

<u>PROOF</u>:

$\Longrightarrow$ Let $v \simeq u$ in $U_{(\hat{t}_0, t]}$

(i) Let $u_0 w$ be admissible for $w \in U_{(t_0, t]}$ $\Longrightarrow$ $\exists \hat{u} \in U_{\hat{I}}$ $\ni \hat{u} = u_0 w_0 \hat{u}/_{(t, \hat{t}_1]}$ $\Longrightarrow$ $v_0 w_0 \hat{u}/_{(t, \hat{t}_1]}$ is admissible (since $u \simeq v$) $\Longrightarrow$ $v_0 w$ is admissible. And vice-versa.

(ii) Using the input $\hat{u}$ of (i), $u_0 w_0 \hat{u}/_{(t, \hat{t}_1]}$ and $v_0 w_0 \hat{u}/_{(t, \hat{t}_1]}$ are admissible. $\Longrightarrow$ by equivalence of u and v, $\hat{y}/_{(t_0, \hat{t}_1]} = \hat{z}/_{(t_0, \hat{t}_1]}$ where $\hat{y}$ and $\hat{z}$ are the corresponding outputs to $u_0 w_0 \hat{u}/_{(t, \hat{t}_1]}$ and $v_0 w_0 \hat{u}/_{(t, \hat{t}_1]}$. Hence clearly $\hat{y}/_{(t_0, t]} = \hat{z}/_{(t_0, t]}$.

$\Leftarrow$ Now let (i) and (ii) be true for u, $v \in U_{(\hat{t}_0, t_0]}$. Then for $t = \hat{t}_0$ we obtain the definition of $u \simeq v$.

<u>FACT II.3.2</u>: Let D_I^* be the domain of $\bar{A}_I^*$. Then $(\sigma_0, u) \in D^*_{(t_0,t]}$ iff $\sigma_0 \in \Sigma^*(t_0)$ and $u_0 0 u \in U_{(\hat{t}_0, t]}$ for any $u_0 \in C_{t_0}(\sigma_0)$ and $t_0 > \hat{t}_0$.

<u>PROOF</u>: $(\sigma_0, u) \in D^*_{(t_0,t]} \Leftrightarrow \exists y \in Y_{(t_0,t]} \ni y = \bar{A}^*_{(t_0,t]}(\sigma_0, u)$, (by def. of D_I) $\Longleftrightarrow \exists (\hat{u}, \hat{y}) \in R_{(\hat{t}_0, t]} \ni \hat{u}/_{(t_0,t]} \in C_{t_0}(\sigma_0)$ and $(\hat{u}, \hat{y})/_{(t_0,t]} = (u, y)$, (by def. of A^*_I).

$\Longrightarrow$ Let $(\sigma_0, u) \in D^*_{(t_0,t]} \Longrightarrow \sigma_0 \in \Sigma^*(t_0)$ and $u_0 0 u \in U_{(\hat{t}_0, t]}$ for any $u_0 \in C_{t_0}(\sigma_0)$, by Fact II.3.1, since any two inputs in $C_{t_0}(\sigma_0)$ are equivalent.

$\Longleftarrow$ Assume $\sigma_0 \in \Sigma^*(t_0)$ with $u_0 0 u \in U_{(\hat{t}_0, t]}$ for any $u_0 \in C_{t_0}(\sigma_0)$. $u_0 0 u \in U_{(\hat{t}_0, t]} \Longrightarrow \exists \bar{y} \in Y_{(\hat{t}_0, t]} \ni (u_0 0 u, \bar{y}) \in R_{(\hat{t}_0, t]}$. Note that $(\sigma, u) \in D^*_I$ iff $\exists y \ni (I, \sigma, u, y) \in \bar{A}^*$. So by def. of $\bar{A}^*$, $((t_0, t], \sigma_0, u, \bar{y}/_{(t_0,t]}) \in \bar{A}^*$ and hence $(\sigma_0, u) \in D^*_{(t_0,t]}$.

The following Thm.II.3.1 is the counterpart of Thm.I.3.2 under $\measuredangle$3.

THM.II.3.1: Under the State Axioms $\mathcal{A}3$ the following are equivalent:

(i) $\mathcal{O}$ has a State Description under $\mathcal{A}3$.

(ii) $\mathcal{O}$ is causal.

(iii) $(\Sigma^*,\bar{A}^*)$ is a State Description under $\mathcal{A}3$.

PROOF: (i) $\Longrightarrow$ (ii) $\mathcal{O}$ has a State Description under $\mathcal{A}3 \Longrightarrow \mathcal{O}$ has a State Description under $\mathcal{A}1$ (Note II.3.1) $\Longrightarrow \mathcal{O}$ is causal (Thm.I.3.2)

(ii) $\Longrightarrow$ (iii) To show $(\Sigma^*,\bar{A}^*)$ is a State Description under $\mathcal{A}3$, the only axiom that needs verification is (S2'), the others being verified in (Thm.I.3.2).

Let $I_0 = (t_0,t_1]$ and $(\sigma_0,u_0)\in D^*_{(t_0,t_1]}$ be given.

Then $(\sigma_0,u_0)\in D^*_{(t_0,t_1]} \Longleftrightarrow \sigma_0\in\Sigma^*(t_0)$ and $u'_0 0 u_0\in U_{(\hat{t}_0,t_1]}$ for any $u'_0\in C_{t_0}(\sigma_0)$ by Fact II.3.2. $\sigma_1 = C^{-1}_{t_1}[H_{t_1}(u'_0 0 u_0)]$ is the state required by (S2'). In fact: $(\sigma_0,u_0 0 u)\in D^*_{(t_0,t_2]} \Longrightarrow u'_0 0 u_0 0 u\in U_{(\hat{t}_0,t_2]}$ for any $u'_0\in C_{t_0}(\sigma_0)$ (by Fact II.3.2).

$\Longrightarrow u' 0 u\in U_{(\hat{t}_0,t_2]}$ for any $u'\in H_{t_1}(u'_0 0 u_0)$ (by definition of $H_{t_1}(u'_0 0 u_0)$) $\Longrightarrow$ $(\sigma_1,u)\in D^*_{(t_1,t_2]}$ (by Fact II.3.2).

$(\sigma_1,u)\in D^*_{(t_1,t_2]} \Longrightarrow u'ou\in U_{(\hat{t}_0,t_2]}$ for any $u'\in C_{t_1}(\sigma_1)$,

i.e., for any $u'\in H_{t_1}(u'_0 0 u_0)$

$\Longrightarrow u'_0 0 u_0 0 u\in U_{(\hat{t}_0,t_2]}$ and hence

$\bar{u}_0 0 u_0 0 u\in U_{(\hat{t}_0,t_2]}$ for any $\bar{u}_0\in H_{t_0}(u'_0)$

$\Longrightarrow (\sigma_0,u_0 0 u)\in D^*_{(t_0,t_2]}$ (by Fact II.3.2).

Let u be such that $(\sigma_0,u_0 0 u)\in D^*_{(t_0,t_2]}$ and $(\sigma_1,u)\in D^*_{(t_1,t_2]}$. As $\sigma_0 = C^{-1}_{t_0}[H_{t_0}(u'_0)]$ and $\sigma_1 = C^{-1}_{t_1}[H_{t_1}(u'_0 0 u_0)]$ we have $y = \bar{A}^*_{(t_0,t_2]}(\sigma_0,u_0 0 u)$ and $y' = \bar{A}^*_{(t_1,t_2]}(\sigma_1,u)$ are such that: $\exists\, \bar{y} \ni (\bar{y},u'_0 0 u_0 0 u)\in R_{(\hat{t}_0,t_2]}$ and $\bar{y}/_{(t_0,t_2]} = y$ and $\exists \bar{y}' \ni (\bar{y}',\hat{u}_0 0 u)\in R_{(\hat{t}_0,t_2]}$ and $\bar{y}'/_{(t_1,t_2]} = y'$ with $\hat{u}_0 \simeq u'_0 0 u_0$. By definition of $\hat{u}_0 \simeq u'_0 0 u_0$, the $(t_1,t_2]$ portions of the outputs corresponding to the inputs $\hat{u}_0 0 u$ and $u'_0 0 u_0 0 u$ are equal.

$\Longrightarrow y/_{(t_1,t_2]} = y'$. Hence (S2') is satisfied.

(iii) $\Longrightarrow$ (i) is a triviality.

The next theorem is one that was promised in Note II.2.3. It shows that the reduced State Description of Def.II.2.4 satisfies $\mathcal{A}3$.

THM.II.3.2: Let $(\Sigma,\bar{A})$ be a State Description of $\mathcal{O}$ under $\mathcal{A}3$. Then the reduced State Description $(\Sigma_R,\bar{A}_R)$ obtained from $(\Sigma,\bar{A})$ always satisfies $\mathcal{A}3$.

PROOF: We verify each axiom to get the proof.

(M1) For each $I = (t_0,t]\subset\hat{I}$

$\exists\sigma_0\in\Sigma_R(t_0) \ni (I,\sigma_0,u,y)\in\bar{A}_R \Longrightarrow (u,y)\in R_I$ since $\sigma_0\in\Sigma(t_0)$ and $(\Sigma,\bar{A})$ satisfy A3.

$(u,y)\in R_I \Rightarrow \exists(\bar{u},\bar{y})\in R_{(\hat{t}_0,t]} \ni (\bar{u},\bar{y})/_I = (u,y)$.

Applying (S2') to $(\sigma_{\hat{t}_0},\bar{u}/_{(\hat{t}_0,t_0]})\in D_{(\hat{t}_0,t]}$ we can say $\exists\sigma_0\in\Sigma(t_0) \ni$

(i) $(\sigma_{\hat{t}_0},\bar{u}/_{(\hat{t}_0,t_0]}\circ u')\in D_{(\hat{t}_0,t]} \Longleftrightarrow (\sigma_0,u')\in D_{(t_0,t]}$

and

(ii) $\bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},\bar{u}/_{(\hat{t}_0,t]}\circ u')/_{(t_0,t]} = A_{(t_0,t]}(\sigma_0,u')$

for any u' that satisfies (i).

Since $\bar{u}/_{(\hat{t}_0,t_0]}\circ u$ is admissible,

$(\sigma_{\hat{t}_0},\bar{u}/_{(\hat{t}_0,t_0]}\circ u)\in D_{(\hat{t}_0,t]} \Longrightarrow (\sigma_0,u)\in D_{(t_0,t]}$ and $[\bar{y} = \bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},\bar{u}/_{(\hat{t}_0,t_0]}\circ u)]/_{(t_0,t]} = \bar{A}_{(t_0,t]}(\sigma_0,u) = y$.

Now $\sigma_0\in\Sigma_S(t_0)$ (i.e., σ_0 is not a singular state) since it is reachable from $\sigma_{\hat{t}_0}\in\Sigma(\hat{t}_0)$ by $\bar{u}/_{(\hat{t}_0,t_0]}$, due to (i) and (ii) above. Then the only way for $\sigma_0\notin\Sigma_R(t_0)$ to occur is when we throw all states in $\Sigma_{\sigma_0}(t_0)$ but one. If ever σ_0 is thrown out there must exist a state $\sigma_0'\in\Sigma_R(t_0) \ni \sigma_0 \simeq \sigma_0'$. In this case $y = \bar{A}_{(t_0,t]}(\sigma_0',u)$ by definition of equivalence of states. So $\exists\,\sigma_0'\in\Sigma_R(t_0) \ni (\sigma_0',u)\in D_{R(t_0,t]}$ and $y = \bar{A}_{R(t_0,t]}(\sigma_0',u)$.

(M2) and (S1) are trivially true since $(\Sigma,\bar{A})$ satisfied them.

(S2') Let $I = (t_0,t_1]$ and $(\sigma_0,u_0) \in D_{R(t_0,t_1]}$, then $(\sigma_0,u_0) \in D_{(t_0,t_1]}$ and $\sigma_0 \in \Sigma(t_0)$.

Since $(\Sigma,\bar{A})$ satisfies $\mathcal{A}3$, for $I = (t_0,t_1]$ and $(\sigma_0,u_0) \in D_I$ there exists at least one $\sigma_1 \in \Sigma(t_1)$ such that:

$$\text{(i)}\ (\sigma_0,u_0 0 u) \in D_{(t_0,t]} \Longleftrightarrow (\sigma_1,u) \in D_{(t_1,t]}$$

$$\text{(ii)}\ \bar{A}_{(t_1,t]}\ (\sigma_1,u) = \bar{A}_{(t_0,t]}(\sigma_0,u_0 0 u)/_{(t_1,t]}$$

$\forall u$ satisfying (i).

From (i) and (ii) it is obvious that σ_1 is reachable from σ_0 by u_0. As $\sigma_0 \in \Sigma_R(t_0)$ it is non-singular and thus reachable from σ_{t_0}, say by u'_0. Then by Fact II.2.1 $\sigma_1 \in \Sigma(t_1)$ is reachable from $\sigma_{\hat{t}_0}$. Thus the only way for $\sigma_1 \notin \Sigma_R(t_1)$ is that it be thrown out when we keep a single state from $\Sigma_{\sigma_1}(t_1)$. In that case $\exists\, \sigma_1' \in \Sigma_R(t_1) \ni \sigma_1' \simeq \sigma_1$ and σ_1' is the state required by (S2') for $(\Sigma_R,\bar{A}_R)$. In fact:

$$\begin{aligned}
\text{(i)}\ (\sigma_0,u_0 0 u) \in D_{R(t_0,t]} &\Longleftrightarrow (\sigma_0,u_0 0 u) \in D_{(t_0,t]} \\
&\qquad \text{since } \sigma_0 \in \Sigma_R(t_0) \\
&\Longleftrightarrow (\sigma_1,u) \in D_{(t_1,t]} \text{ by} \\
&\qquad \text{(S2') for } (\Sigma,\bar{A}) \\
&\Longleftrightarrow (\sigma_1',u) \in D_{(t_1,t]} \text{ since} \\
&\qquad \sigma_1 \simeq \sigma_1'
\end{aligned}$$

$\Longrightarrow (\sigma_1', u) \in D_{R(t_1,t]}$ since $\sigma_1' \in \Sigma_R(t_1)$ and by def. of $D_{R(t_1,t]}$

(ii) $\bar{A}_{R(t_0,t]}(\sigma_0, u_0 0 u) = \bar{A}_{(t_0,t]}(\sigma_0, u_0 0 u)$ by def. of A_R. Hence

$\bar{A}_{R(t_0,t]}(\sigma_0, u_0 0 u)/(t_1,t] = A_{(t_1,t]}(\sigma_1, u)$ by by (S2') for $(\Sigma, \bar{A})$

$= \bar{A}_{(t_1,t]}(\sigma_1', u)$ since $\sigma_1' \simeq \sigma_1$

$= \bar{A}_{R(t_1,t]}(\sigma_1', u)$ since $(\sigma_1', u) \in D_{R(t_1,t]}$

proving (S2').

<u>NOTE II.3.3</u>: ABOUT $\mathcal{A}2$ and $\mathcal{A}3$. The example in II.2 has already shown us that the two Axiom Sets are not obtainable one from another. So now we concentrate our attention on the relation between $\mathcal{A}2$ and $\mathcal{A}3$ and show that those two sets are almost equivalent, with a minor restriction on the State Descriptions satisfying $\mathcal{A}2$.

<u>THM.II.3.3</u>:

(i) Let $(\Sigma, \bar{A})$ be a State Description of $\mathcal{O}$ under $\mathcal{A}3$. Then $(\Sigma, \bar{A})$ is a State Description of $\mathcal{O}$ under $\mathcal{A}2$.

(ii) Let $(\Sigma,\bar{A})$ be a State Description of $\mathcal{O}$ under $\mathcal{A}2$. Then $(\Sigma_{HR},\bar{A}_{HR})$ obtained from $(\Sigma,\bar{A})$ as defined in Def.II.2.5, is a State Description of $\mathcal{O}$ under $\mathcal{A}3$.

PROOF:

(i) (M1) and (S1) are the same for both $\mathcal{A}2$ and $\mathcal{A}3$, and (S2) is trivially implied by (S2').

(M2') is (S2') applied to the interval $I_0 = (\hat{t}_0,t_0]$. In fact: $u_0 \in U_{(\hat{t}_0,t_0]} \Longleftrightarrow (\sigma_{\hat{t}_0},u_0) \in D_{(\hat{t}_0,t_0]}$; then (S2') requires there be a state $\sigma_0 \in \Sigma(t_0)$ such that: (i) $(\sigma_{\hat{t}_0},u_0 0 u) \in D_{(\hat{t}_0,t]} \Longleftrightarrow (\sigma_0,u) \in D_{(t_0,t]}$

(ii) $\bar{A}_{(t_0,t]}(\sigma_0,u) = [\bar{A}_{(\hat{t}_0,t]}(\sigma_0,u_0 0 u) = \bar{y}]/_{(t_0,t]}$

Naturally $\sigma_0 \in \Sigma(t_0)$ is the state required by (M2').

Verifying: $y = \bar{A}_{(t_0,t]}(\sigma_0,u) \Longrightarrow (\sigma_0,u) \in D_{(t_0,t]} \Longrightarrow (\sigma_{\hat{t}_0},u_0 0 u) \in D_{(\hat{t}_0,t]} \Longrightarrow \exists \bar{y} \in Y_{(\hat{t}_0,t]} \ni [\bar{y} = \bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},u_0 0 u)]/_{(t_0,t]} = y \Longrightarrow \exists y_0$, namely $y_0 = \bar{y}/_{(\hat{t}_0,t_0]} \ni (y_0 0 y,u_0 0 u) \in R_{(\hat{t}_0,t]}$.

$\exists y_0 \in Y_{(\hat{t}_0,t]} \ni (u_0 0 u,y_0 0 y) \in R_{(\hat{t}_0,t]} \Longrightarrow y_0 0 y = \bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},u_0 0 u) \Longrightarrow (\sigma_{\hat{t}_0},u_0 0 u) \in D_{(\hat{t}_0,t]} \Longrightarrow (\sigma_0,u) \in D_{(t_0,t]} \Longrightarrow [y_0 0 y = \bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},u_0 0 u)]/_{(t_0,t]} = \bar{A}_{(t_0,t]}(\sigma_0,u) = y$.

(ii) We have shown in Note II.2.2, on our way to the construction of reduced State Descriptions, that a State Description $(\Sigma,\bar{A})$ was obtainable from a given one, under $\mathcal{A}2$, such that $\Sigma(\hat{t}_0)$ was a unit set. Very little modification was necessary and indicated. Here then, we assume that this modification is already made and $\Sigma(\hat{t}_0)$ is a unit set.

(M2') is satisfied by the above comment and (S1) is automatically satisfied since $(\Sigma,\bar{A})$ was already a State Description.

(M1) needs verification since there may not be enough states left after obtaining Σ_{HR}. Let $I = (t_0,t]\subset\hat{I}$.

$\exists\,\sigma_0\in\Sigma_{HR}(t_0)\ \ni(I,\sigma_0,u,y)\in\bar{A}_{HR} \Longrightarrow (u,y)\in R_I$ clearly.

Let $(u,y)\in R_I \Longrightarrow \exists(\bar{u},\bar{y})\in R_{(\hat{t}_0,t]}\ \ni(\bar{u},\bar{y})/_I = (u,y)$.

By (M2'), as $\Sigma(\hat{t}_0)$ is a unit set, $\exists\sigma_0\in\Sigma(t_0)$ for $u_0 = \bar{u}/_{(\hat{t}_0,t_0]}\ \ni$

(1) $(\sigma_{\hat{t}_0},u_0 0 u')\in D_{(t_0,t']} \Longleftrightarrow (\sigma_0,u')\in D_{(t_0,t']}$

(2) $[\bar{y} = \bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},u_0 0 u')]/_{(t_0,t']} =$

$\bar{A}_{(t_0,t']}(\sigma_0,u') = y'$

From (1) and (2), σ_0 is reachable from $\sigma_{\hat{t}_0}$; hence $\sigma_0\in\Sigma_{HR}(t_0)$ and $(\sigma_{\hat{t}_0},u_0 0 u')\in D_{HR(\hat{t}_0,t']} \Longleftrightarrow$ $(\sigma_{\hat{t}_0},u_0 0 u')\in D_{(\hat{t}_0,t']} \Longleftrightarrow (\sigma_0,u')\in D_{(t_0,t']} \Longleftrightarrow$ $(\sigma_0,u')\in D_{HR(t_0,t']}$ and $\bar{A}_{HR(t_0,t']}(\sigma_0,u') =$

$\bar{A}_{(t_0,t']}(\sigma_0,u') = [\bar{A}_{(\hat{t}_0,t']}(\sigma_{\hat{t}_0},u_0 0 u') = \bar{y} =$
$\bar{A}_{HR(\hat{t}_0,t']}(\sigma_{\hat{t}_0},u_0 0 u')]/_{(t_0,t']}$. Hence for $u' = u$ we get $y = \bar{A}_{HR(t_0,t]}(\sigma_0,u) \Longrightarrow \exists\, \sigma_0 \in \Sigma_{HR}(t_0) \ni (I,\sigma_0,u,y) \in \bar{A}_{HR}$.

(S2') We let $I_0 = (t_0,t_1]$ and $(\sigma_0,u_0) \in D_{IHR}$.
$(\sigma_0,u_0) \in D_{HR(t_0,t]} \Longleftrightarrow \sigma_0$ is non singular and $(\sigma_0,u_0) \in D_I$ by def. of D_{HR}.

$$\Longleftrightarrow \exists\, \bar{u}_0 \in U_{(\hat{t}_0,t_0]} \ni \begin{cases} (3)\ (\sigma_{\hat{t}_0},\bar{u}_0 0 u') \in D_{(t_0,t']} \Longleftrightarrow \\ \qquad (\sigma_0,u') \in D_{(t_0,t']} \\ (4)\ \bar{A}_{(t_0,t']}(\sigma_0,u') = \\ \qquad \bar{A}_{(\hat{t}_0,t']}(\sigma_{\hat{t}_0},u_0 0 u')/_{(t_0,t']} \end{cases}$$

Applying (M2') to $(\sigma_{\hat{t}_0},u_0 0 u) \in D_{(\hat{t}_0,t_1]}$, $\exists\, \sigma_1 \in \Sigma_2(t_1) \ni$

(5) $(\sigma_{\hat{t}_0},\bar{u}_0 0 u_0 0 u) \in D_{HR(\hat{t}_0,t]} \Longleftrightarrow (\sigma_{\hat{t}_0},\bar{u}_0 0 u_0 0 u) \in D_{(\hat{t}_0,t]}$
$\Longleftrightarrow (\sigma_1,u) \in D_{(t_1,t]} \Longleftrightarrow (\sigma_1,u) \in D_{HR(t_1,t]}$, since clearly σ_1 is non singular and

(6) $[\bar{A}_{HR(\hat{t}_0,t]}(\sigma_{\hat{t}_0},\bar{u}_0 0 u_0 0 u) =$
$\bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},\bar{u}_0 0 u_0 0 u)]/_{(t_1,t]} = A_{(t_1,t]}(\sigma_1,u) =$
$A_{HR(t_1,t]}(\sigma_1,u)$.

Thus $(\sigma_{\hat{t}_0},\bar{u}_0 0 u_0 0 u) \in D_{HR(\hat{t}_0,t]} \Longleftrightarrow (\sigma_0,u_0 0 u) \in D_{HR(t_0,t]}$ by (3) and since σ_0 is non singular. Combining this with (5) we obtain:

(7) $(\sigma_0,u_0 0 u) \in D_{HR(t_0,t]} \Longleftrightarrow (\sigma_1,u) \in D_{HR(t_1,t]}$

And (4) put together with (6) yields:

$$(8)\ \{\bar{A}_{HR(t_0,t]}(\sigma_0,u_0 0u) = \bar{A}_{(t_0,t]}(\sigma_0,u_0 0u) =$$
$$[\bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},\bar{u}_0 0u_0 0u) =$$
$$\bar{A}_{HR(\hat{t}_0,t]}(\sigma_{\hat{t}_0},\bar{u}_0 0u_0 0u)]/_{(t_0,t]}\}/_{(t_1,t]} =$$
$$\bar{A}_{HR(t_1,t]}(\sigma_1,u),\ \text{i.e. } \bar{A}_{HR(t_0,t]}(\sigma_0,u_0 0u)/_{(t_1,t]} =$$
$$\bar{A}_{HR(t_1,t]}(\sigma_1,u).$$

(7) and (8) prove (S2') and complete the theorem.

<u>COR.II.3.1</u>: The half reduced State Description $(\Sigma_{HR},\bar{A}_{HR})$ obtained from the pair $(\Sigma,\bar{A})$ satisfying $\mathcal{A}2$ is a State Description under $\mathcal{A}2$.

<u>PROOF</u>: $(\Sigma,\bar{A})$ satisfies $\mathcal{A}2 \implies (\Sigma_{HR},\bar{A}_{HR})$ is a State Description under $\mathcal{A}3 \implies (\Sigma_{HR},\bar{A}_{HR})$ is a State Description under $\mathcal{A}2$, the implications following from parts (ii) and (i) respectively.

In order to emphasize more the equivalence of $\mathcal{A}2$ and $\mathcal{A}3$ we can restate Thm.II.3.3 as follows:

<u>THM.II.3.3</u>': A conjectured State Description $(\Sigma,\bar{A})$ for an object Θ with $\Sigma(\hat{t}_0)$ a unit set and $\Sigma(t)$ containing no singular states for each t, is a State Description under $\mathcal{A}2$ iff it is a State Description under $\mathcal{A}3$.

<u>NOTE II.3.4</u>: Nothing stronger than that has been obtained about the equivalence of the two Sets of Axioms $\mathcal{A}2$ and $\mathcal{A}3$. However as we always will be dealing with

State Descriptions that are at least half-reduced, this much is what is needed.

<u>NOTE II.3.5</u>: The next two simple corollaries of Thm.II.3.3, together with Thm.II.3.2, Thm.II.3.3, constitute an answer to the problem posed in Note II.2.3.

<u>COR.II.3.2</u>: Let $(\Sigma,\bar{A})$ be a State Description of Θ under $\mathcal{A}2$. Then the reduced State Description $(\Sigma_R,\bar{A}_R)$ of Θ satisfies $\mathcal{A}2$.

<u>PROOF</u>: By Thm.II.3.3 $(\Sigma_{HR},\bar{A}_{HR})$, obtained from $(\Sigma,\bar{A})$ is a State Description of Θ under $\mathcal{A}3$. Then by Thm.II.3.2 the reduced State Description $(\Sigma_R,\bar{A}_R)$ obtained from $(\Sigma_{HR},\bar{A}_{HR})$ satisfies $\mathcal{A}3$. Hence, again by Thm.II.3.3 $(\Sigma_R,\bar{A}_R)$ is a State Description of Θ under $\mathcal{A}2$.

<u>COR.II.3.3</u>: A conjectured, reduced State Description $(\Sigma_R,\bar{A}_R)$ satisfies $\mathcal{A}2$ iff it satisfies $\mathcal{A}3$.

<u>PROOF</u>: This is a direct result of Thm.II.3.3 and Cor.II.3.2.

<u>NOTE II.3.6</u>: From now on, only Axiom Set $\mathcal{A}3$ will be used as the basic one. We will briefly say "let $(\Sigma,\bar{A})$ be a State Description" or "let $\Sigma(t)$ be the State Space and $\bar{A}$ the I-O-S-R." These will mean "satisfying $\mathcal{A}3$." However for reduced or half-reduced State Descriptions the set $\mathcal{A}2$ can be referred to as a Theorem.

<u>II.4--About Reduced and Half-Reduced State Descriptions</u>

<u>NOTE II.4.1</u>: With the following theorems we harvest the fruits of our efforts in the previous two sections. That these fruits are very nutritious will be appreciated as we proceed into the next chapters. One must have remarked in the Example of II.2 that to test whether a conjectured State Description is one that satisfies $\mathcal{A}3$ may be a very difficult task for some objects. The following theorems give us some algorithms to ease this task.

<u>THM.II.4.1</u>: The State Description $(\Sigma^*,\bar{A}^*)$, obtained by use of equivalence classes of inputs, for an object $\mathcal{O}$, is reduced.

<u>PROOF</u>:

(i) First we show that $\Sigma^*(t)$, for each t, contains only non-singular states. Let $\sigma_0 \in \Sigma^*(t_0)$, then by definition of $\Sigma^*(t_0)$, $\exists\, u_0 \in U_{(\hat{t}_0,t_0]} \ni C_{t_0}(\sigma_0) = H_{t_0}[u_0]$. This input u_0, or any $u'_0 \in H_{t_0}[u_0]$ is the input that takes $\sigma_{\hat{t}_0}$ into σ_0. For:

$u_0 0 u \in U_{(\hat{t}_0,t]} \Longrightarrow \exists$ a unique (unique since $\mathcal{O}$ is causal by Thm.II.3.1) output $\bar{y}_{t_0} y_t \in Y_{(\hat{t}_0,t]}$ $\ni (u_0 0 u, \bar{y}_{t_0} y_t) \in R_{(\hat{t}_0,t]}$. $\Longrightarrow ((t_0,t],\sigma_0,u,y) \in \bar{A}^*$. Since by definition of $\bar{A}^*$ $(I',\sigma'_0,u',y')\in\bar{A}^* \Longleftrightarrow$ $\exists(u'_0,y'_0)\in R_{(\hat{t}_0,t']} \ni u'_0/_{(\hat{t}_0,t'_0]}\in C_{t'_0}(\sigma'_0)$ and $(u'_0,y'_0)/_I = (u',y')$ which is clearly the case here. Hence $u_0 0 u\in U_{(\hat{t}_0,t]} \Longrightarrow (\sigma_0,u)\in D^*_{(t_0,t]}$.

$(\sigma_0,u)\in D^*_{(t_0,t]} \Longrightarrow \exists y\in Y_{(t_0,t]} \ni ((t_0,t],\sigma_0,u,y)\in\bar{A}^*$ $\Longrightarrow u_0 0 u\in U_{(\hat{t}_0,t]}$ again by definition of $\bar{A}^*$. Thus we have proved: $(\sigma_0,u)\in D^*_{(t_0,t]} \Longleftrightarrow u_0 0 u\in U_{(\hat{t}_0,t]}$ $\Longleftrightarrow (\sigma_{\hat{t}_0},u_0 0 u)\in D^*_{(\hat{t}_0,t]}$ since $\sigma_{\hat{t}_0}$ is unique.

$[\bar{y}_{t_0} y_t = \bar{A}^*_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},u_0 0 u)]/_{(t_0,t]} = y =$ $\bar{A}^*_{(t_0,t]}(\sigma_0,u)$ from above.

(ii) Second we prove that $\Sigma^*(t)$ does not contain equivalent states for each t. We do this by showing: $\sigma'_0 \simeq \sigma''_0 \Longrightarrow \sigma'_0 = \sigma''_0$ for σ'_0 and $\sigma''_0\in\Sigma^*(t_0)$, t_0 arbitrary.

$$\sigma'_0 \simeq \sigma''_0 \Longleftrightarrow \begin{cases} (\sigma'_0,u)\in D^*_I \Longleftrightarrow (\sigma''_0,u)\in D^*_I \text{ and} \\ \qquad\qquad\qquad \text{where } I = (t_0,t]. \\ y = \bar{A}^*_I(\sigma'_0,u) = \bar{A}^*_I(\sigma''_0,u)\ \forall u \text{ in (1).} \end{cases}$$

$y = \bar{A}^*_I(\sigma'_0,u) \Longrightarrow \exists(u'_0,y'_0)\in R_{(\hat{t}_0,t]} \ni u'_0/_{(\hat{t}_0,t_0]} =$ $C_{t_0}(\sigma'_0)$ and $(u'_0,y'_0)/_I = (u,y)$.

$y = \bar{A}^*_I(\sigma''_0,u) \Longrightarrow \exists(u''_0,y''_0)\ R_{(\hat{t}_0,t]} \ni u''_0/_{(\hat{t}_0,t_0]} =$ $C_{t_0}(\sigma''_0)$ and $(u''_0,y''_0)/_I = (u,y)$.

We will be done if we can show $u'_0/(\hat{t}_0,t_0] \simeq u''_0/(\hat{t}_0,t_0]$ which would imply the equality of the equivalence classes $C_{t_0}(\sigma'_0)$ and $C_{t_0}(\sigma''_0)$, since the equivalence classes have to be mutually exclusive. Then, as C_{t_0} is one-to-one by definition $C_{t_0}(\sigma'_0) = C_{t_0}(\sigma''_0)$ will yield $\sigma'_0 = \sigma''_0$.

To prove $u'_0/(\hat{t}_0,t_0] \simeq u''_0/(\hat{t}_0,t_0]$

(a) Consider any $w \in U_{(t_0,\hat{t}_1]} \ni u'_0/(\hat{t}_0,t_0]0w$ is admissible. Then: $\exists\, \hat{y} \in Y_{\hat{I}} \ni (u'_0/(\hat{t}_0,t_0]0w, \hat{y}) \in R_{\hat{I}}$. But, as $\Sigma(\hat{t}_0)$ is a unit set $\hat{y} = \bar{A}^*_{\hat{I}}(\sigma_{\hat{t}_0}, u'_0\, /(\hat{t}_0,t_0]0w)$. By (S2') $\exists\, \sigma_0 \in \Sigma(t_0)$ and by [RE3], $\sigma_0 = C^{-1}_{t_0}(H_{t_0}[u'_0/(\hat{t}_0,t_0]])$, which is σ'_0 itself, $\ni$ (1) $\hat{y}/_{(t_0,\hat{t}_1]} = \bar{A}^*_{(t_0,\hat{t}_1]}(\sigma'_0,w)$ and by state equivalence (2) $\hat{y}/_{(t_0,\hat{t}_1]} = \bar{A}^*_{(t_0,\hat{t}_1]}(\sigma''_0,w)$. (2) $\Longrightarrow$ $\exists (u_0,y_0) \in R_{(\hat{t}_0,\hat{t}_1]} \ni u_0/(\hat{t}_0,t_0] \in C_{t_0}(\sigma''_0)$ and $u_0/(t_0,\hat{t}_1] = w$. As $u''_0/(\hat{t}_0,t_0] \in C_{t_0}(\sigma''_0)$, $u_0/(\hat{t}_0,t_0] \simeq u''_0/(\hat{t}_0,t_0] \Longrightarrow u''_0/(\hat{t}_0,t_0]0w$ is admissible.

The same proof can be used to show: $u''_0/(\hat{t}_0,t_0]0v$ is admissible $\Longrightarrow u'_0/(\hat{t}_0,t_0]$ is admissible.

(b) Equations (1) and (2) show that whenever $u'_0/(\hat{t}_0,t_0]0w$ and $u''_0/(\hat{t}_0,t_0]0w$ are admissible then the part of the response corresponding to w is $\hat{y}/_{(t_0,\hat{t}_0]}$ due to state equivalence.

Hence $u'_0/_{(\hat{t}_0,t_0]} \simeq u''_0/_{(\hat{t}_0,t_0]}$ proving $\sigma'_0 \simeq \sigma''_0 \Longrightarrow \sigma'_0 = \sigma''_0$.

<u>THM.II.4.2</u>: For a given object $\mathcal{O}$, any two reduced State Spaces have the same cardinality at time t.

<u>PROOF</u>: By Thm.II.3.1 an object $\mathcal{O}$ has a State Description iff $(\Sigma^*,\bar{A}^*)$ is one. By Thm.II.4.3 $\Sigma^*(t_0)$ is a reduced State Description for each t_0. What we will do then is to establish a well defined, one-to-one, onto correspondence between any reduced State Space $\Sigma(t_0)$ and the equivalence classes of inputs which will imply a well-defined, one-to-one, onto correspondence between $\Sigma(t_0)$ and $\Sigma^*(t_0)$.

Given any $\sigma_0 \in \Sigma(t_0)$, σ_0 must be reachable from $\sigma_{\hat{t}_0}$ since $\Sigma(t_0)$ is reduced, i.e.,

$$\exists\, u_0 \in U_{(\hat{t}_0,t_0]} \ni \begin{cases} (1)\ u_0 0 u \in U_{(\hat{t}_0,t]} \Longrightarrow (\sigma_0,u) \in D_{(t_0,t]} \\ \\ (2)\ \bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},u_0 0 u)/_{(t_0,t]} = \\ \qquad \bar{A}_{(t_0,t]}(\sigma_0,u) \end{cases}$$

<u>We define</u>: $B_{t_0} : \Sigma(t_0) \to \mathcal{H}_{t_0}$ to be : $B_{t_0}(\sigma_0) \triangleq H_{t_0}[u_0]$, where u_0 is the above input guaranteed to exist by reachability.

B_{t_0} is well-defined: i.e., $\sigma'_0 = \sigma''_0 \Rightarrow B_{t_0}(\sigma'_0) = B_{t_0}(\sigma''_0)$ or equivalently $\sigma'_0 = \sigma''_0 \Rightarrow H_{t_0}(u'_0) = H_{t_0}(u''_0)$ where u'_0 and u''_0 are the inputs that take $\sigma_{\hat{t}_0}$ into σ'_0 and σ''_0 respectively. In fact:

$$u'_0 0 u \in U_{(\hat{t}_0,t]} \Leftrightarrow (\sigma'_0,u) \in D_{(t_0,t]} \quad \text{by (1)}$$
$$\Leftrightarrow (\sigma''_0,u) \in D_{(t_0,t]} \quad \text{since } \sigma'_0 = \sigma''_0$$
$$\Leftrightarrow u''_0 0 u \in U_{(\hat{t}_0,t]} \quad \text{by (1).}$$

$$\bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},u'_0 0 u)/_{(t_0,t]} = \bar{A}_{(t_0,t]}(\sigma'_0,u) \quad \text{by (2).}$$
$$= \bar{A}_{(t_0,t]}(\sigma''_0,u) \quad \text{since } \sigma'_0 = \sigma''_0$$
$$= \bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},u''_0 0 u)/_{(t_0,t]} \quad \text{by (2).}$$

Hence by Fact II.3.1, $u'_0 \approx u''_0$.

B_{t_0} is one to one: i.e., $B_{t_0}(\sigma'_0) = B_{t_0}(\sigma''_0) \Rightarrow \sigma'_0 = \sigma''_0$ which is true iff $H_{t_0}(u'_0) = H_{t_0}(u''_0) \Rightarrow \sigma'_0 = \sigma''_0$ which is true iff $u'_0 \simeq u''_0 \Rightarrow \sigma'_0 = \sigma''_0$, where u'_0 and u''_0 are as before. We prove $\sigma'_0 \simeq \sigma''_0$: $(\sigma'_0,u) \in D_{(t_0,t]} \Leftrightarrow u'_0 0 u \in U_{(\hat{t}_0,t]} \Leftrightarrow u''_0 0 u \in U_{(\hat{t}_0,t]} \Leftrightarrow (\sigma''_0,u) \in D_{(t_0,t]}$ and $\bar{A}_{(t_0,t]}(\sigma'_0,u) = [\bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},u'_0 0 u) = \bar{A}_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0},u''_0 0 u)]/_{(t_0,t]} = \bar{A}_{(t_0,t]}(\sigma''_0,u)$. So $\sigma'_0 \simeq \sigma''_0$ and $\Sigma(t_0)$ being a reduced State Space $\sigma'_0 = \sigma''_0$.

B_{t_0} is onto: i.e., given any $H_{t_0}[u_0]$, $\exists \sigma_0 \in \Sigma(t_0)$ $\ni B_{t_0}(\sigma_0) = H_{t_0}[u_0]$. This means given any $u_0 \in U_{(\hat{t}_0, t_0]}$ is leaves the object in some state. This is guaranteed by (S2') applied to $I_0 = (\hat{t}_0, t_0]$ and $(\sigma_{\hat{t}_0}, u_0) \in D_{(\hat{t}_0, t_0]}$ (or by (M2')).

NOTE II.4.2: The following theorem is a simple corollary of Thm.II.4.2. But yet it is a very important one in that it shows for a given object Θ the reduced State Description is unique in a sense eliminating the ambiguities posed by the example of Section II.2 about the reduced descriptions.

THM.II.4.3: Any reduced State Description $(\Sigma, \bar{A})$ of Θ is nothing but $(\Sigma^*, \bar{A}^*)$ defined by use of equivalence classes of inputs.

PROOF: We have seen in Thm. II.4.2 the existence of a one to one, onto mapping $B_{t_0} : \Sigma(t_0) \to \mathcal{H}_{t_0}$ $\ni B_{t_0}(\sigma_0) \triangleq H_{t_0}[u_0]$, where u_0 was the input that took $\sigma_{\hat{t}_0}$ into σ_0. The existence of u_0 was guaranteed $\Sigma(t_0)$ being reduced. Actually this makes B_{t_0}, the mapping C_{t_0} and $\Sigma(t_0)$, the State Space $\Sigma^*(t_0)$ since by definition $\Sigma^*(t_0)$ is any set of the same cardinality as $\mathcal{H}_{t_0}$ and C_{t_0} any one of many one to one, onto mappings between two sets of same cardinality (Def.I.3.2).

Let now $((t_0,t],\sigma_0,u,y)\in\bar{A}_{(t_0,t]}$, consider that u_0 taking $\sigma_{\hat{t}_0}$ into σ_0. By definition of reachability $u_0 0 u\in U_{(\hat{t}_0,t_0]}$ and $\exists y_0 \ni (u_0 0 u, y_0 0 y)\in R_{(\hat{t}_0,t]}$. Clearly $(u_0 0 u, y_0 0 y)/_{(t_0,t]} = (u,y)$ and $u_0 0 u/_{(\hat{t}_0,t_0]} = u_0 \in B_{t_0}(\sigma_0)$. The definition of $\bar{A}^*_I$ being satisfied $\bar{A}_{(t_0,t]} = \bar{A}^*_{(t_0,t]}$ and hence $(\Sigma,\bar{A}) = (\Sigma^*,\bar{A}^*)$.

COR.II.4.1: A State Description of $\mathcal{G}$ is reduced iff it is the description $(\Sigma^*,\bar{A}^*)$ obtained by use of equivalence classes of inputs.

PROOF: The corollary is a direct result of Thms.II.4.1 and II.4.3.

NOTE II.4.3: As we go along, in Chapter IV especially, we will see that some results about Half Reduced State Descriptions can be conveniently used to prove some descriptions are State Descriptions and half reduced. We first give the definition of a particular State Description (already defined in [RE5]) then prove it is a half reduced description. But unfortunately the converse will *not* necessarily be true as will be explained in Note II.4.6.

DEF.II.4.1:

For $t_0 \in \hat{I}$, a partitioning of $U_{(\hat{t}_0, t_0]}$ into classes $H'_{t_0}[u_0]$ of inputs is called a HALF EQUIVALENCE PARTITIONING iff:

(i) $u_0 \in U_{(\hat{t}_0,t_0]} \Longrightarrow u_0 \in H'_{t_0}[u_0]$

(ii) $u_0, u'_0 \in U_{(\hat{t}_0,t_0]} \Longrightarrow$ either $H'_{t_0}[u_0] = H'_{t_0}[u'_0]$ or $H'_{t_0}[u_0] \cap H'_{t_0}[u'_0] = \Phi$.

(iii) $u'_0 \in H'_{t_0}[u_0] \Longrightarrow u'_0 \simeq u_0$.

The family $\mathcal{H}'_{t_0}$ is: $\mathcal{H}'_{t_0} \triangleq \{H'_{t_0}[u_0] : u_0 \in U_{(\hat{t}_0,t_0]}\}$, for $t_0 > \hat{t}_0$. As before we take $\Sigma'(t_0)$ to be any set with the same cardinality as $\mathcal{H}'_{t_0}$, for $t_0 > \hat{t}_0$ and assign $\sigma_0 \in \Sigma'(t_0)$ to $H'_{t_0}[u_0]$ by $C'_{t_0} : \Sigma'(t_0) \to H'_{t_0}[u_0]$, where C'_{t_0} is one of the many one to one, onto mappings between two sets of same cardinality.

The I-O-S-R is defined as before: for $I_0 = (t_0,t]$, $t_0 > \hat{t}_0$, $(I,\sigma_0,u,y) \in \bar{A}'$ iff $u_0 0 u$ is admissible for some $u_0 \in C'_{t_0}(\sigma_0)$ and $y = y_0/_I$, where $(u_0 0 u, y_0) \in R_I$ for $t_0 = \hat{t}_0$, $\Sigma'(\hat{t}_0)$ is any unit set with $((\hat{t}_0,t],\sigma_{\hat{t}_0},u,y)$ being such that $\sigma_{\hat{t}_0} \in \Sigma(\hat{t}_0)$ and $(u,y) \in R_{(\hat{t}_0,t]}$.

NOTE II.4.4:

What we did in Def.II.4.1 was to partition the equivalence classes of inputs into some mutually exclusive classes of pre-t_0 inputs, and then take these classes as describing the states, everything else remaining the same. This is a partition based on a sufficient condition, rather than a necessary and sufficient condition for equivalence of pre-t_0 inputs. That we are now provided with a Half Reduced State Description is the why of the next theorems.

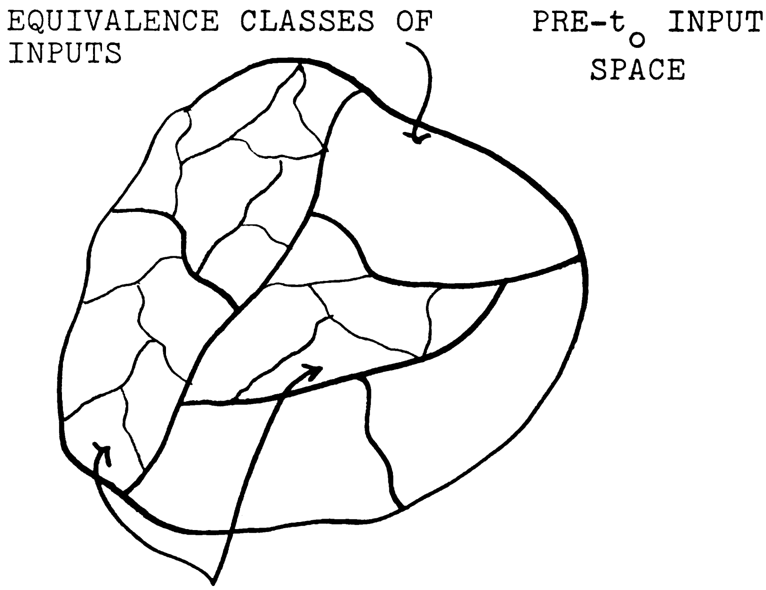


Figure II.4.1

<u>THM.II.4.4:</u> The description $(\Sigma', \bar{A}')$ of Def.II.4.1 is a Half Reduced State Description $(\Sigma_{HR}, \bar{A}_{HR})$.

<u>PROOF</u>: First of all it is a State Description since:

(M1) Let $I = (t_0, t] \subset \hat{I}$ (the case $t_0 = \hat{t}_0$ being trivial we consider $t_0 > \hat{t}_0$):

$$(u,y) \in R_I \Longrightarrow \exists (\hat{u},\hat{y}) \in R_{(\hat{t}_0,t]} \ni (\hat{u},\hat{y})/_I = (u,y).$$

Consider $\sigma_0 = C_{t_0}^{-1}(H'_{t_0}[\hat{u}/_{(\hat{t}_0,t_0]}])$. Then by definition of $\bar{A}'$, $(I, \sigma_0, u, y) \in \bar{A}'$.

$(I,\sigma_0,u,y)\in\bar{A}' \Longrightarrow (u,y)\in R_I$ is obvious.

(M2) is fulfilled by definition of $\Sigma'(\hat{t}_0)$. (S1) is automatically satisfied, since for $I = (t_0,t]$, $\sigma_0\in\Sigma'(t_0)$ and $u\in U_I$ we have a unique output that can correspond to $u_0 0 u$, where $C'_{t_0}(\sigma_0) = H'_{t_0}[u_0]$, if $u_0 0 u$ is admissible.

(S2') As in Fact II.3.2 the domain D'_I of $\bar{A}'_I$ is: $D'_{(t_0,t]} = \{(\sigma_0,u) : \sigma_0\in\Sigma'(t_0)$ and $u_0 0 u\in U_{(\hat{t}_0,t]}$ for $u_0\in C'_{t_0}(\sigma_0)\}$, $t_0 > \hat{t}_0$, and $D'_{(\hat{t}_0,t]} =$ $\{(\sigma_{\hat{t}_0},u) : \sigma_{\hat{t}_0}\in\Sigma'(\hat{t}_0)$ and $u\in U_{(\hat{t}_0,t]}\}$. For $I_0 = (t_0,t_1]$ and $(\sigma_0,u_0)\in D'_I$ the state required by (S2') becomes: $\sigma_1 = C'^{-1}_{t_1}(H'_{t_1}[u'_0 0 u_0])$ where $u'_0\in C'_{t_0}(\sigma_0)$, if $t_0 > \hat{t}_0$, or $\sigma_1 = C'^{-1}_{t_0}(H'_{t_1}[u_0])$ if $t_0 = \hat{t}_0$. In fact:

$$
\begin{aligned}
\text{(i)}\ (\sigma_0,u_0 0 u)\in D'_{(t_0,t]} &\Longleftrightarrow \sigma_0\in\Sigma'(t_0) \text{ and } \\
&\quad u'_0 0(u_0 0 u)\in U_{(\hat{t}_0,t]} \text{ for } \\
&\quad u'_0\in C'_{t_0}(\sigma_0) \\
&\Longleftrightarrow (u'_0 0 u_0) 0 u\in U_{(\hat{t}_0,t]} \\
&\quad \text{and } \sigma_1 = \\
&\quad C'^{-1}_{t_1}(H'_{t_1}[u'_0 0 u_0])\in\Sigma'(t_1) \\
&\Longleftrightarrow (\sigma_1,u)\in D'_{(t_1,t]} \text{ for } \\
&\quad t_0 > \hat{t}_0.
\end{aligned}
$$

$$(\sigma_{\hat{t}_0}, u_0 0 u) \in D'(\hat{t}_0, t] \Longleftrightarrow \sigma_{\hat{t}_0} \in \Sigma'(\hat{t}_0) \text{ and }$$
$$u_0 0 u \in U(\hat{t}_0, t]$$
$$\Longleftrightarrow u_0 0 u \in U(\hat{t}_0, t] \text{ and } \sigma_1 =$$
$$C'^{-1}_{t_1}(H'_{t_1}[u_0]) \in \Sigma'(t_1)$$
$$\Longleftrightarrow (\sigma_1, u) \in D'(t_1, t], \text{ for}$$
$$t_0 = \hat{t}_0.$$

(ii) $A'_{(t_0,t]}(\sigma_0, u_0 0 u) = y_0/_{(t_0,t]}$ where $(u'_0 0(u_0 0 u), y_0) \in R_{(\hat{t}_0,t]}$ with $u'_0 \in C'_{t_0}(\sigma_0)$. $\bar{A}'_{(t_1,t]}(\sigma_1, u) = \tilde{y}_0/_{(t_1,t]}$ where $(\tilde{u}_0 0 u, \tilde{y}_0) \in R_{(\hat{t}_0,t]}$ for some $\tilde{u}_0 \in C'_{t_1}(\sigma_1) = H'_{t_1}[u'_0 0 u_0] \Longrightarrow \tilde{u}_0 \simeq u'_0 0 u_0 \Longrightarrow \tilde{y}_0/_{(t_1,t]} = y_0/_{(t_1,t]}$ by definition of equivalence. Thus we have: $\bar{A}'_{(t_0,t]}(\sigma_0, u_0 0 u)/_{(t_1,t]} = \bar{A}'_{(t_1,t]}(\sigma_1, u)$ for $t_0 > \hat{t}_0$. For $t_0 = \hat{t}_0$, $\bar{A}'_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0}, u_0 0 u) = y_0$ where $(u_0 0 u, y_0) \in R_{(\hat{t}_0,t]}$ and $\sigma_{\hat{t}_0} \in \Sigma'(\hat{t}_0)$. $\bar{A}'_{(t_1,t]}(\sigma_1, u) = \tilde{y}_0/_{(t_1,t]}$ where $(\tilde{u}_0 0 u, \tilde{y}_0) \in R_{(\hat{t}_0,t]}$ for some, hence for any $\tilde{u}_0 \in C'_{t_1}(\sigma_1) = H'_{t_1}[u_0] \Longrightarrow \tilde{u}_0 \simeq u_0$ and hence $y_0/_{(t_1,t]} = \tilde{y}_0/_{(t_1,t]}$ giving us (S2').

Now we prove that $(\Sigma', \bar{A}')$ is half reduced. To show there are no singular states in $\Sigma'(t_0)$ for any t_0, we consider $\sigma_0 \in \Sigma'(t_0)$. Then $\exists u_0 \in U_{(\hat{t}_0,t]} \ni C'_{t_0}(\sigma_0) = H'_{t_0}[u_0]$. u_0 is the input that takes $\sigma_{\hat{t}_0}$ into σ_0. For:

(i) Clearly $(\sigma_{\hat{t}_0}, u_0) \in D'_{(\hat{t}_0, t_0]}$.

(ii) $(\sigma_{\hat{t}_0}, u_0 0 u) \in D'_{(\hat{t}_0,t]} \Longleftrightarrow u_0 0 u \in U_{(\hat{t}_0,t]}$ and $\sigma_{\hat{t}_0} \in \Sigma(\hat{t}_0)$ by def. of D'_I.

$\Longleftrightarrow u_0 0 u \in U_{(\hat{t}_0,t]}$ for $u_0 \in C'_{t_0}(\sigma_0)$ from above

$\Longleftrightarrow (\sigma_0, u) \in D'_{(t_0,t]}$ by definition of D'_I.

(iii) $\bar{A}'_{(\hat{t}_0,t]}(\sigma_{\hat{t}_0}, u_0 0 u) = y_0$, for $y_0 \ni (u_0 0 u, y_0) \in R_{(\hat{t}_0,t]}$ and $\bar{A}'_{(t_0,t]}(\sigma_0, u) = y_0/_{(t_0,t]}$ since $u'_0 0 u$ is admissible for any $u'_0 \in C'_{t_0}(\sigma_0) = H'_{t_0}[u_0]$ and since $u'_0 \simeq u_0$, the $(t_0,t]$ portions of the outputs corresponding to $u'_0 0 u$ and $u_0 0 u$ are equal.

NOTE II.4.5: That the State Description $(\Sigma', \bar{A}')$ is not a reduced one can be observed by the following fact, if $(\Sigma', \bar{A}')$ is based on a $\mathcal{H}'_{t_0}$ that is finer than $\mathcal{H}_{t_0}$ (the condition "finer" is necessary because equivalence classes of inputs are by definition a Half Reduced Partitioning). The fact now is there are states that are equivalent, namely the ones corresponding to the classes $H'_{t_0}[u_0]$, $H'_{t_0}[u_1]$, . . . , $H'_{t_0}[u_n]$, . . . such that $u_0 \simeq u_1 \simeq . . . \simeq u_n \simeq . . . .$

<u>NOTE II.4.6</u>: The converse of Thm.II.4.4 (the counterpart of Thm.II.4.2) is unfortunately not true unless some extra hypothesis is added. Any Half Reduced State Description is not based on a Half Reduced Partitioning for the following reason. Consider a Half Reduced State Description $(\Sigma_{HR},\bar{A}_{HR})$ and define $\Sigma_{\sigma_0}(t_0) \triangleq \{\sigma'_0 \in \Sigma_{HR}(t_0) : \sigma'_0 \simeq \sigma_0\}$ for a fixed $\sigma_0 \in \Sigma_{HR}(t_0)$. If we had thrown all of $\Sigma_{\sigma_0}(t_0)$ but σ_0 from $\Sigma_{HR}(t_0)$ then σ_0 would correspond to an equivalence class $H_{t_0}[u_0]$ for some u_0, making $\Sigma_{\sigma_0}(t_0)$ correspond to the same class. Now we compare card $\Sigma_{\sigma_0}(t_0)$ with card $H_{t_0}[u_0]$. We can always assume card $\Sigma_{\sigma_0}(t_0)$ > card $H_{t_0}[u_0]$ (if it were not we could always add as many equivalent states to σ_0 as we wish without disturbing the State Description) for the purposes of our note. Then card $\Sigma_{\sigma_0}(t_0)$ > card $H_{t_0}[u_0]$ would make impossible the existence of a one to one, onto correspondence between $\Sigma(t_0)$ and any Half Reduced Partitioning $\mathcal{H}'_{t_0}$, since equivalent states can come only from the partitioning of an equivalence class $H_{t_0}[\cdot]$ and the most $H_{t_0}[\cdot]$ can be partitioned to, is into its individual inputs.

<u>NOTE II.4.7</u>: We close this section, and the chapter, with a theorem that constitutes an answer to a problem posed in [RE2], concerning a property that State Descriptions should have.

<u>THM.II.4.5</u>: The following is true for a Half Reduced State Description $(\Sigma_{HR},\bar{A}_{HR})$

(i) $(\sigma_0,u_0 0 u)\in D_{HR(t_0,t]} \Longrightarrow (\sigma_0,u_0)\in D_{HR(t_0,t_1]}$

where $t_1\in(t_0,t]$ is arbitrary.

(ii) $\bar{A}_{HR(t_0,t]}(\sigma_0,u_0 0 u)/_{(t_0,t_1]} = \bar{A}_{HR(t_0,t_1]}(\sigma_0,u_0)$

<u>PROOF</u>: As all states are non singular, $\exists u'_0 \ni \sigma_0$ is reachable from $\sigma_{\hat{t}_0}$ by u'_0 and

$(\sigma_0,u_0 0 u)\in D_{HR(t_0,t]} \Longleftrightarrow u'_0 0 u_0 0 u\in U_{(t_0,t]}$ by definition of reachability

$\Longrightarrow u'_0 0 u_0\in U_{(\hat{t}_0,t_1]}$ the restriction of an input to $I\subset\hat{I}$ is admissible.

$\Longrightarrow (\sigma_0,u_0)\in D_{HR(t_0,t_1]}$ by reachability again.

$$\left.\begin{array}{l} \bar{A}_{HR(t_0,t]}(\sigma_0,u_0 0 u) = \\ \bar{A}_{HR(\hat{t}_0,t]}(\sigma_{\hat{t}_0},u'_0 0 u_0 0 u)/_{(t_0,t]} \\ \\ \bar{A}_{HR(t_0,t_1]}(\sigma_0,u_0) = \\ \bar{A}_{HR(\hat{t}_0,t_1]}(\sigma_{\hat{t}_0},u'_0 0 u_0)/_{(t_0,t_1]} \end{array}\right\}\quad \text{both by reachability}$$

$$\bar{A}_{HR(\hat{t}_0,t]}(\sigma_{\hat{t}_0},u'_0 0 u_0 0 u)/_{(t_0,t_1]} = \bar{A}_{HR(\hat{t}_0,t_1]}(\sigma_{\hat{t}_0},u'_0 0 u_0)$$

by causality (Thm.II.3.1). Hence:

$$\bar{A}_{HR(t_0,t]}(\sigma_0,u_0 0 u)/_{(t_0,t_1]} = \bar{A}_{HR(t_0,t_1]}(\sigma_0,u_0).$$

CHAPTER III

LINEAR, TIME-INVARIANT OBJECTS

III.1--Introduction

Many authors including Zadeh and Balakrishnan give the definition of "Linearity," "Time-Invariance," etc., for objects, in terms of the State Descriptions of objects [ZA2], [BA4]. However, whether an object has these properties or not, does not depend on its State Description, a machinery introduced by us. In fact state descriptions are ambivalent: even if an object is linear or time-invariant, there are state descriptions in which the State Space and the I-O-S-R are not linear or time-invariant. Consider:

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t)$$

$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + u(t) \; ;$$

This certainly is a Linear Object, but the state description is non-linear (for an example of time-invariant case, substitute x_1 by t in the 2 x 2 matrix).

In Sections 2 and 3 of this chapter, we start with the basic definitions of "Linearity" and "Time-Invariance" and proceed to show that state descriptions can be choosen to provide the object with "Linear" and "Time Invariant" Reduced Descriptions. Then Zadeh's definitions are obtained as results of these natural definitions. Also some nice results, such as "Separation Property of the I-O-S-R" for linear systems, "$\Sigma(t)$ is the same set at any time t" for time-invariant objects, are attained among others.

As another application of equivalence classes of inputs, the linear, time invariant system given by the equations:

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$

is investigated and conditions are found for its State Space to be reduced when A is in Jordan form. These conditions will prove useful in the last section of Chapter IV.

To close this section we would like to add that the whole chapter illustrates the importance of "the equivalence classes of inputs" and exhibits how much can be accomplished with the help of this concept without going into deep mathematical analysis.

III.2--Linear Objects and Properties of the State Description

NOTE III.2.1: We start the basic definition of "Linear Objects" and after stating some facts, we show that equivalence classes of inputs can be given some linear structure, which will be very useful in showing the "Linearity of the State Description." As usual, we have to stand some tedious Lemmas.

DEF.III.2.1:

The object $\mathcal{O}$, given by the list $R_{\hat{I}}$ of I-O pairs, is a LINEAR OBJECT iff:

$$\left.\begin{array}{c}(u_1,y_1)\in R_{\hat{I}}\\ \text{and}\\ (u_2,y_2)\in R_{\hat{I}}\end{array}\right\} \Longrightarrow (u_1 + au_2, y_1 + ay_2)\in R_{\hat{I}}, \text{ for } a\in\mathbb{R}.$$

FACTS III.2:

1. $\mathcal{O}$ is linear iff $\left.\begin{array}{c}(u_1,y_1)\in R_I\\ (u_2,y_2)\in R_I\end{array}\right\} \Longrightarrow$ $(u_1 + au_2, y_1 + ay_2)\in R_I \quad \forall I\subset\hat{I},\ a\in\mathbb{R}$.

2. $(0,0)\in R_I$, $\forall I\subset\hat{I}$. But $(0,0)\in R_{(\hat{t},t]}$ is the unique pair with 0 first element, for all t due to causality.

3. $\mathcal{O}$ is Linear $\Longrightarrow$ U_I and Y_I are linear function spaces $\forall I$.

DEF.III.2.2:

$$aH_{t_0}[u_0] \triangleq \{au_0^1 : u_0^1 \in H_{t_0}[u_0] \text{ and } a\in\mathbb{R}\} \text{ for } a \neq 0$$

$$aH_{t_0}[u_0] \triangleq H_{t_0}[0] \quad \text{for } a = 0.$$

$$H_{t_0}[u'_0] + aH_{t_0}[u''_0] = \{u_0^1 + au_0^2 : u_0^1 \in H_{t_0}[u'_0],$$
$$u_0^2 \in H_{t_0}[u''_0] \text{ and } a \in \mathbb{R}\}$$

<u>LEMMA III.2.1</u>: Let Θ be linear, u_0 and $u'_0 \in U_{(\hat{t}_0, t_0]}$ and $w \in U_{(t_0, \hat{t}_1]}$. Then $(u_0 + u'_0)0w \in U_{(\hat{t}_0, \hat{t}_1]} \Longrightarrow \exists w_1$ and $w_2 \in U_{(t_0, \hat{t}_1]} \ni u_0 0 w_1$ and $u'_0 0 w_2 \in U_{(\hat{t}_0, \hat{t}_1]}$, and $u_0 0 w_1 + u'_0 0 w_2 = (u_0 + u'_0)0w$.

<u>PROOF</u>: "$\Longleftarrow$" is trivial.

"$\Longrightarrow$" $u_0 \in U_{(\hat{t}_0, t_0]} \Longleftrightarrow \exists w_1 \in U_{(t_0, \hat{t}_1]} \ni u_0 0 w_1 \in U_{(\hat{t}_0, \hat{t}_1]}$. As the input space is linear (Fact III.2.1), $(u_0 + u'_0)0w - u_0 0 w_1 \in U_{(\hat{t}_0, \hat{t}_1]} \Longrightarrow u'_0 0(w - w_1) \in U_{(\hat{t}_0, \hat{t}_1]}$. Now if we let $w_2 \triangleq w - w_1 \in U_{(t_0, \hat{t}_1]}$ then obviously $u'_0 0 w_2 \in U_{(\hat{t}_0, t_1]}$ and $(u_0 + u'_0)0w = u_0 0 w_1 + u'_0 0 w_2$.

<u>LEMMA III.2.2</u>: For a linear object Θ, $u'_0 \simeq u''_0$ and $u^1{}_0 \simeq u^2{}_0 \Longrightarrow u'_0 + u^1{}_0 \simeq u''_0 + u^2{}_0$, for u'_0, u''_0, $u^1{}_0$ and $u^2{}_0 \in U_{(\hat{t}_0, t_0]}$.

<u>PROOF</u>:

(i) Let $(u'_0 + u^1{}_0)0w$ be admissible, by Lemma III.2.1.

$\exists\, w'$ and $w^1 \in U_{(t_0, \hat{t}_1]} \ni (u'_0 + u^1{}_0)0w = u'_0 0 w' + u^1{}_0 0 w^1$.

$$\left.\begin{array}{l} u'{}_0 0w' \in U_{(\hat{t}_0,\hat{t}_1]} \text{ iff} \\ u''{}_0 0w' \in U_{(\hat{t}_0,\hat{t}_1]}, \text{ since } u'_0 \simeq u''_0 \\ \\ u^1{}_0 0w^1 \in U_{(\hat{t}_0,\hat{t}_1]} \text{ iff} \\ u^2{}_0 0w^1 \in U_{(\hat{t}_0,\hat{t}_1]}, \text{ since } u^1_0 \simeq u^2_0 \end{array}\right\} \text{Using linearity and}$$

Lemma III.2.1 $u''{}_0 0w' + u^2{}_0 0w^1 \in U_{(\hat{t}_0,\hat{t}_1]} \Longrightarrow$ $(u''_0 + u^2_0)0w \in U_{(\hat{t}_0,\hat{t}_1]}$. In exactly the same manner we can show $(u''_0 + u^2_0)0w \in U_{(\hat{t}_0,\hat{t}_1]} \Longrightarrow$ $(u'_0 + u^1_0)0w \in U_{(\hat{t}_0,\hat{t}_1]}$.

(ii) We must now show that the portion corresponding to w, of the response to $(u''_0 + u^2_0)0w$ is identical with the portion corresponding to w of the response to $(u'_0 + u^1_0)0w$. In the following all concatenations occur at t_0. There exists unique, since causal (Thm.I.3.2), $y'{}_0 0y'$ and $y^1{}_0 0y^1 \in Y_{(\hat{t}_0,\hat{t}_1]} \ni$ $(u'{}_0 0w', y'{}_0 0y') \in R_{(\hat{t}_0,\hat{t}_1]}$ and $(u^1{}_0 0w^1, y^1{}_0 0y^1) \in R_{(\hat{t}_0,\hat{t}_1]}$. Using linearity $(u'{}_0 0w' + u^1{}_0 0w^1, y'{}_0 0y' + y^1{}_0 0y^1) \in R_{(\hat{t}_0,\hat{t}_1]}$. As $u'{}_0 0w' + u^1{}_0 0w^1 = (u'_0 + u^1_0)0w$, by uniqueness of the outputs for $\hat{I} = (\hat{t}_0,\hat{t}_1]$ to given inputs, we can write $((u'_0 + u^1_0)0w, (u'_0 + u^1_0)0y) \in R_{(\hat{t}_0,\hat{t}_1]}$ where $y \triangleq y' + y^1$. Using $u'_0 \simeq u''_0$ and $u^1_0 \simeq u^2_0$, the outputs to $u''{}_0 0w'$ and $u^2{}_0 0w^1$ are such that $(u''{}_0 0w', y''{}_0 0y') \in R_{(\hat{t}_0,\hat{t}_1]}$ and $(u^2{}_0 0w^1, y^2{}_0 0y^1) \in R_{(\hat{t}_0,\hat{t}_1]}$

where y' and y^1 are as above due to equivalence of inputs. As $u''_0 0 w' + u^2_0 0 w^1 = (u''_0 + u^2_0) 0 w$ we have $((u''_0 + u^2_0) 0 w, y''_0 0 y' + y^2_0 0 y^1) =$ $((u''_0 + u^2_0) 0 w, (y''_0 + y^2_0) 0 y) \in R_{(\hat{t}_0, \hat{t}_1]}$ proving the Lemma.

<u>LEMMA III.2.3</u>: For a linear object $\mathcal{O}$, $u'_0 \simeq u''_0 \Longrightarrow$ $au'_0 \simeq au''_0 \quad \forall a \in \mathbb{R}$, u'_0 and $u''_0 \in U_{(\hat{t}_0, t_0]}$.

<u>PROOF</u>: For $a = 0$ the Lemma is trivially true. So let $a \neq 0$, then:

$$(au'_0) 0 w \in U_{(\hat{t}_0, \hat{t}_1]} \Longleftrightarrow a^{-1}[(au'_0) 0 w] = u'_0 0 (a^{-1} w) \in U_{(\hat{t}_0, \hat{t}_1]}$$

by linearity.

$$\Longleftrightarrow u''_0 0 (a^{-1} w) \in U_{(\hat{t}_0, \hat{t}_1]} \quad \text{by } u'_0 \simeq u''_0$$

$$\Longleftrightarrow a[u''_0 0 (a^{-1} w)] = (au''_0) 0 w \in U_{(\hat{t}_0, \hat{t}_1]}$$

by linearity.

Then there exists a unique $y'_0 0 y \in Y_{(\hat{t}_0, \hat{t}_1]} \ni$ $((au'_0) 0 w, y'_0 0 y) \in R_{(\hat{t}_0, \hat{t}_1]}$. Now:

Linearity $\Longrightarrow (u'_0 0 (a^{-1} w),\ a^{-1}(y'_0 0 y)) \in R_{(\hat{t}_0, \hat{t}_1]}$

$u'_0 \simeq u''_0 \Longrightarrow (u''_0 0 (a^{-1} w),\ a^{-1}(y''_0 0 y)) \in R_{(\hat{t}_0, \hat{t}_1]}$ where y''_0 is $\ni (au''_0, y''_0) \in R_{(\hat{t}_0, t_0]}$

Linearity $\Longrightarrow ((au''_0) 0 w,\ y''_0 0 y) \in R_{(\hat{t}_0, \hat{t}_1]}$.

That the t_0 to $\hat{t}_1$ portions of the responses, to $(au'_0) 0 w$ and $(au''_0) 0 w$, corresponding to w are equal proves this Lemma.

NOTE III.2.2: Our aim is to show the linearity of the collection of equivalence classes of inputs and all the machinery of the previous three Lemmas has been introduced for this purpose. Although $H_{t_0}[u'_0] + aH_{t_0}[u''_0]$ is well defined for each t_0, in Def.III.2.2, there is nothing that guarantees it is an equivalence class. Once this is established, the linearity of $\mathcal{H}_{t_0}$ for each t_0 is then reached with respect to the operations defined in Def. III.2.2. On the other hand, $u'_0 + au''_0$ being an input for $a\in\mathbb{R}$, $H_{t_0}[u'_0 + au''_0]$ is a well defined equivalence class. The next theorem is central, in that it establishes the linearity of $\mathcal{H}_{t_0}$.

THM.III.2.1: For a linear object Θ, $H_{t_0}[u'_0 + au''_0] = H_{t_0}[u'_0] + aH_{t_0}[u''_0]$ $\forall t_0 \in \hat{I}$. Hence the equivalence classes of inputs form a linear space $\mathcal{H}_{t_0}$ $\forall t_0 \in \hat{I}$.

PROOF:

(i) First we show $H_{t_0}[u'_0 + u''_0] = H_{t_0}[u'_0] + H_{t_0}[u''_0]$.

Let $u_0 \in H_{t_0}[u'_0 + u''_0]$ then $u_0 \simeq u'_0 + u''_0$. We can write u_0 as follows: $u_0 = u^1_0 + u^2_0$ where $u^1_0 = u'_0$ and $u^2_0 = u_0 - u'_0$. Then $u^1_0 \in H_{t_0}[u'_0]$ obviously. Also $u_0 \simeq u'_0 + u''_0 \Longrightarrow u_0 - u'_0 \simeq u''_0$ by Lemma III.2.2 $\Longrightarrow u^2_0 \simeq u_0 - u'_0 \in H_{t_0}[u''_0]$. So we obtained $u_0 = u^1_0 + u^2_0$ where $u^1_0 \in H_{t_0}[u'_0]$,

$u^2_0 \in H_{t_0}[u''_0]$. By Def.III.2.2 $u_0 \in H_{t_0}[u'_0] + H_{t_0}[u''_0] \Longrightarrow H_{t_0}[u'_0 + u''_0] \subset H_{t_0}[u'_0] + H_{t_0}[u''_0]$.

Now we let $u_0 \in H_{t_0}[u'_0] + H_{t_0}[u''_0]$. Then $u_0 = u^1_0 + u^2_0 \ni u^1_0 \simeq u'_0$ and $u^2_0 \simeq u''_0$, (Def.III.2.2). By Lemma III.2.2 $u_0 = u^1_0 + u^2_0 \simeq u'_0 + u''_0 \Longrightarrow u_0 \in H_{t_0}[u'_0 + u''_0]$. Thus $H_{t_0}[u'_0 + u''_0] = H_{t_0}[u'_0] + H_{t_0}[u''_0]$.

(ii) Second we show $H_{t_0}[au_0] = aH_{t_0}[u_0]$ for $0 \neq a \in \mathbb{R}$, since for $a = 0, H_{t_0}[au_0] = H_{t_0}[0]$ and $aH_{t_0}[u_0] = H_{t_0}[0]$ by Def. III.2.2.

Let $u'_0 \in H_{t_0}[au_0] \Longrightarrow u'_0 \simeq au_0$. Define $u^1_0 \triangleq a^{-1}u'_0$. Then $u'_0 = au^1_0$ and $u^1_0 = a^{-1}u'_0 \simeq u_0$ by Lemma III.2.3. Hence $u'_0 \in aH_{t_0}[u_0] \Longrightarrow H_{t_0}[au_0] \subset aH_{t_0}[u_0]$.

Now let $u'_0 \in aH_{t_0}[u_0] \Longrightarrow u'_0 = au^1_0$ for $u^1_0 \in H_{t_0}[u_0]$. Using Lemma III.2.3 $u^1_0 \simeq u_0 \Longrightarrow au^1_0 \simeq au_0 \Longrightarrow u'_0 \simeq au_0 \Longrightarrow u'_0 \in H_{t_0}[au_0]$. Thus $H_{t_0}[au_0] = aH_{t_0}[u_0]$.

(iii) Finally $H_{t_0}[u'_0 + au''_0] = H_{t_0}[u'_0] + H_{t_0}[au''_0] = H_{t_0}[u'_0] + aH_{t_0}[u''_0]$ proving the theorem.

DEF.III.2.3:

$\Sigma^*(t_0)$ will be called a COMPATIBLE STATE SPACE for the linear object $\mathcal{O}$ iff $\Sigma^*(t_0)$ is a vector space isomorphic to $\mathcal{H}_{t_0}$ for each $t_0 \in \hat{I}$.

NOTE III.2.3: When defining $\Sigma^*(t_0)$, the one to one, onto mapping $C_{t_0} : \Sigma^*(t_0) \to \mathcal{H}_{t_0}$ was chosen to be any one of such mappings among many others existing between two sets of the same cardinality (Def.I.3.2). To get a compatible state space we also require C_{t_0} be chosen in such a way that, it be an isomorphism between $\Sigma^*(t_0)$ and $\mathcal{H}_{t_0}$, which is always possible due to Thm.III.2.1. Now we are sure, at least, that a linear object can be provided with a reduced state space that is linear.

NOTE III.2.4: The answer to the question of whether a half reduced state space, based on a half reduced partitioning (Def.II.4.1), can be chosen to be a linear space, may not be affirmative. The reason is the lack of a theorem similar to Thm.III.2.1 providing us with some linear structure for $\mathcal{H}'_{t_0}$. For example, if we take as half reduced partitioning, the equivalence classes of inputs $\mathcal{H}_{t_0}$ with only one equivalence class partitioned into two nonempty subclasses (the others remaining the same), then this $\mathcal{H}'_{t_0}$ has not a linear structure in the sense of Def.III.2.2. However, we can with no difficulty assert that there are special ways of partitioning $\mathcal{H}_{t_0}$

in a manner that can provide us with a linear structure. One way of doing it is to take each input as a half reduced partitioning; other ways will be seen in Chapter IV.

The next theorem is a result about the dimension of the state space and illustrates the usefulness of equivalence classes of inputs.

<u>THM.III.2.2</u>: If the input space of $\mathcal{O}$ is finite dimensional of dimension n, then the dimension of the compatible state space $\Sigma^*(t_0)$, dim $\Sigma^*(t_0)$, is such that dim $\Sigma^*(t_0) \leqslant$ dim $U_{(\hat{t}_0,t_0]} \leqslant n$, for all t_0.

<u>PROOF</u>: It is enough to show dim $\mathcal{H}_{t_0} \leqslant$ dim $U_{(\hat{t}_0,t_0]}$ $\forall t_0 \in \hat{I}$, since $\Sigma^*(t_0)$ is isomorphic to $\mathcal{H}_{t_0} \Longrightarrow$ dim $\Sigma^*(t_0)$ = dim $\mathcal{H}_{t_0}$. It is also clear that dim $U_{(\hat{t}_0,t_0]}$ is less than dim $U_{(\hat{t}_0,\hat{t}_1]}$ $\forall t_0 \in \hat{I}$, since every $u_0 \in U_{(\hat{t}_0,t_0]}$ is the restriction of some input $u \in U_{(\hat{t}_0,\hat{t}_1]}$ to $(\hat{t}_0,t_0]$. Let $u_1, u_2, \ldots, u_n \in U_{(\hat{t}_0,t_0]}$ be a basis for $U_{(\hat{t}_0,t_0]}$. Then the classes $H_{t_0}[u_1], H_{t_0}[u_2], \ldots, H_{t_0}[u_n]$ spans $\mathcal{H}_{t_0}$. For any equivalence class $H_{t_0}[u_0]$ in $\mathcal{H}_{t_0}$ we have: $u_0 \in U_{(\hat{t}_0,t_0]}$ and u_0 can be expressed in terms of the basis $u_0 = \sum_{i=1}^{n} a_i u_i$. Using Thm.III.2.1 $H_{t_0}[u_0] = H_{t_0}[\sum_{i=1}^{n} a_i u_i] = \sum_{i=1}^{n} a_i H_{t_0}[u_i]$. Hence dim $\mathcal{H}_{t_0} \leqslant$ dim $U_{(\hat{t}_0,t_0]}$.

NOTE III.2.5: Fortunately, $\dim \Sigma^*(t_0) = \dim U_{(\hat{t}_0,t_0]}$ does not generally hold, since there is a multitude of known examples where the dimension of the state space is less than the dimension of the input space. In order to show $\dim \mathcal{H}_{t_0} = \dim U_{(\hat{t}_0,t_0]}$ we need to show that the set $H_{t_0}[u_1], \ldots, H_{t_0}[u_n]$ is linearly independent. Equivalently then, we must show that $H_{t_0}[u_k] = \sum_{\substack{i=1\\i\neq k}}^{n} a_i H_{t_0}[u_i]$ for at least one $a_i \neq 0 \Longrightarrow u_k = \sum_{\substack{i=1\\i\neq k}}^{n} a_i u_i$. However we can only infer $u_k \approx \sum_{\substack{i=1\\i\neq k}}^{n} a_i u_i$, which does not necessarily give equality.

NOTE III.2.6: Thm.III.2.3 that follows is as close as we can get to Zadeh's definition of linearity [ZA2] without further assumptions on our object. It also demonstrates a linearity property of the I-O-S-R, $\bar{A}_I$.

THM.III.2.3: The object $\mathcal{O}$ is linear iff it can be given a reduced state description $(\Sigma_R,\bar{A}_R)$ such that the following are true:

(i) $\Sigma_R(t_0)$ is the compatible state space (Def.III.2.3), $\forall t_0 \in \hat{I}$.

(ii) $D_{R(t_0,t]}$ is a linear space for $t_0 > t_0$, i.e., $(\sigma'_0,u'_0) \in D_{R(t_0,t]}$ and $(\sigma''_0,u''_0) \in D_{R(t_0,t]} \Longrightarrow$

$(\sigma'_0 + a\sigma''_0, u'_0 + au''_0) \in D_{R(t_0,t]}$ for

$(t_0,t] \subsetneq (\hat{t}_0,\hat{t}_1]$ and for any $a \in \mathbb{R}$.

(iii) $\bar{A}_{R(t_0,t]} : D_{R(t_0,t]} \to Y_{(t_0,t]}$ is a linear transformation, i.e., $\bar{A}_{R(t_0,t]}(\sigma'_0 + a\sigma''_0, u' + au'') =$ $\bar{A}_{R(t_0,t]}(\sigma'_0,u') + a\bar{A}_{R(t_0,t]}(\sigma''_0,u'')$ for (σ'_0,u') and $(\sigma''_0,u'') \in D_{R(t_0,t]}$ for $(t_0,t] \subsetneq (\hat{t}_0,\hat{t}_1]$ and for all $a \in \mathbb{R}$.

PROOF:

" $\Longleftarrow$ " Only (ii) and (iii) are enough to imply that $\mathcal{O}$ is linear. In fact (ii) implies that the input space is linear and (iii) implies that the object is linear.

" $\Longrightarrow$" is somewhat tedious to prove.

(i) is true by Thm.II.2.1 and Def.III.2.2. By Cor.II.4.1 any reduced state description is nothing but (Σ^*, A^*) and its properties will be used in the proof of (ii) and (iii).

(ii) Let $(\sigma'_0,u') \in D_{R(t_0,t]}$ and $(\sigma''_0,u''_0) \in D_{R(t_0,t]}$. By Fact II.3.2 we can write

$$
\begin{aligned}
(\sigma'_0,u'_0) \in D_{R(t_0,t]} &\Longleftrightarrow \sigma'_0 \in \Sigma_R(t_0) \text{ and} \\
&\qquad u_0 0 u' \in U_{(\hat{t}_0,t]} \text{ for any} \\
&\qquad u_0 \in C_{t_0}(\sigma'_0) \\
&\Longleftrightarrow \exists\, u'_0 \in U_{(\hat{t}_0,t_0]} \ni C_{t_0}(\sigma'_0) = \\
&\qquad H_{t_0}[u'_0] \text{ and } u'_0 0 u' \in U_{(\hat{t}_0,t]}
\end{aligned}
$$

$$(\sigma''_0, u''_0) \in D_{R(t_0,t]} \iff \exists\, u''_0 \in U_{(\hat{t}_0,t_0]} \ni C_{t_0}(\sigma''_0) = H_{t_0}[u''_0] \text{ and } u''_0 0 u'' \in U_{(\hat{t}_0,t]}$$

The object being linear $u'_0 0 u' + a(u''_0 0 u'') = (u'_0 + au''_0) 0 (u' + au'') \in U_{(\hat{t}_0,t]}$. The equivalence class $H_{t_0}[u'_0 + au''_0] = H_{t_0}[u'_0] + aH_{t_0}[u''_0]$ corresponds to the state $\sigma'_0 + a\sigma''_0$ since $C_{t_0} : \Sigma_R(t_0) \to H_{t_0}$ is defined to be an isomorphism. Hence $\exists\, u_0 \in U_{(\hat{t}_0,t_0]}$, namely $u_0 = u'_0 + au''_0 \ni C_{t_0}(\sigma'_0 + a\sigma''_0) = H_{t_0}[u'_0 + au''_0]$ and $(u'_0 + au''_0) 0 (u' + au'') \in U_{(\hat{t}_0,t]}$. This implies $(\sigma'_0 + a\sigma''_0,\ u'_0 + au''_0) \in D_{R(t_0,t]}$ by Fact II.3.2.

(iii) Let $y' \triangleq \bar{A}_{R(t_0,t]}(\sigma'_0, u')$ and $y'' \triangleq \bar{A}_{R(t_0,t]}(\sigma''_0, u'')$. From part (ii) $(\sigma'_0 + a\sigma''_0,\ u' + au'') \in D_{R(t_0,t_1]}$. Therefore we will be done if we can show $y' + ay'' = \bar{A}_{R(t_0,t]}(\sigma'_0 + a\sigma''_0,\ u' + au'')$, i.e., we have to show: $\exists (\hat{u},\hat{y}) \in R_{(\hat{t}_0,t]} \ni$

$$\hat{u}/_{(\hat{t}_0,t_0]} \in C_{t_0}(\sigma'_0 + a\sigma''_0) \text{ and } (\hat{u},\hat{y})/_{(t_0,t]} = (u' + au'',\ y' + ay'').$$

$y' = \bar{A}_{R(t_0,t]}(\sigma'_0, u') \iff \exists (\hat{u}',\hat{y}') \in R_{(\hat{t}_0,t]} \ni \hat{u}'/_{(\hat{t}_0,t]} \in C_{t_0}(\sigma'_0)$ and $(\hat{u}',\hat{y}')/_{(t_0,t]} = (u',y')$. $y'' = \bar{A}_{R(t_0,t]}(\sigma''_0, u'') \iff \exists (\hat{u}'',\hat{y}'') \in R_{(\hat{t}_0,t]} \ni \hat{u}''/_{(\hat{t}_0,t]} \in C_{t_0}(\sigma''_0)$ and $(\hat{u}'',\hat{y}'')/_{(t_0,t]} = (u'',y'')$ by definition of $\bar{A}^*_I$.

If we take $\hat{u} \triangleq \hat{u}' + a\hat{u}'' \in U_{(\hat{t}_0,t]}$ then $\hat{y}' + a\hat{y}'' \in Y_{(\hat{t}_0,t]}$ will be the unique response such that $(\hat{u}' + a\hat{u}'', \hat{y}' + a\hat{y}'') \in R_{(\hat{t}_0,t]}$, by linearity and by causality. Now, $(\hat{u}' + a\hat{u}'')/_{(\hat{t}_0,t_0]} \in C_{t_0}(\sigma'_0 + a\sigma''_0)$ since:

$$\left.\begin{aligned} &\hat{u}'/_{(\hat{t}_0,t_0]} \in C_{t_0}(\sigma'_0) \Longrightarrow H_{t_0}[\hat{u}'/_{(\hat{t}_0,t_0]}] = \\ &\qquad C_{t_0}(\sigma'_0) \\ &\hat{u}''/_{(\hat{t}_0,t_0]} \in C_{t_0}(\sigma''_0) \Longrightarrow H_{t_0}[\hat{u}''/_{(\hat{t}_0,t_0]}] = \\ &\qquad C_{t_0}(\sigma''_0) \end{aligned}\right\} \Longrightarrow$$

$H_{t_0}[\hat{u}'/_{(\hat{t}_0,t_0]} + a(\hat{u}''/_{(\hat{t}_0,t_0]})] = H_{t_0}[\hat{u}'/_{(\hat{t}_0,t_0]}] + aH_{t_0}[\hat{u}''/_{(\hat{t}_0,t_0]}] = C_{t_0}(\sigma'_0 + a\sigma''_0)$. Hence

$\hat{u}'/_{(\hat{t}_0,t_0]} + a\hat{u}''/_{(\hat{t}_0,t_0]} =$

$(\hat{u}' + a\hat{u}'')/_{(\hat{t}_0,t_0]} \in C_{t_0}(\sigma'_0 + a\sigma''_0)$ and also

$[(\hat{u},\hat{y}) = (\hat{u}' + a\hat{u}'', \hat{y}' + a\hat{y}'')]/_{(t_0,t_1]} =$

$(u' + au'', y' + ay'')$ which is what we had to show.

<u>NOTE III.2.7</u>: As we said in the previous note Thm.II.2.3 is the closest result to Zadeh's definition of linear objects under State Axioms $\mathcal{A}3$ and the basic definition of "Linear Objects." One extra condition on the nature of Θ, namely:

(<u>C1</u>) "$U_{(\hat{t}_0,\hat{t}_1]}$ is so that $_{\hat{t}_0}u_{t_0}0_{\hat{t}_1}$ is admissible for any $u \in U_{(\hat{t}_0,t_0]}$ and any t_0, as well as $_{\hat{t}_0}0_{t_0}u_{\hat{t}_1}$ is admissible for any $u \in U_{(t_0,\hat{t}_1]}$" gives us the property of the I-O-S-R, $\bar{A}_I$, that Zadeh uses to define his linear objects [ZA2]. For a linear object (C1) is equivalent to:

(<u>C2</u>) "$U_{(\hat{t}_0,\hat{t}_1]}$ is so that $_{\hat{t}_0}u_{t_0}u'_{\hat{t}_0}$ is admissible for any $u \in U_{(\hat{t}_0,t_0]}$ and for any $u' \in U_{(t_0,\hat{t}_1]}$, for any t_0."

That (C1) is equivalent to (C2) is easy to show:

(C2) $\Longrightarrow$ (C1) trivially

(C1) $\Longrightarrow$ (C2) since $_{\hat{t}_0}u_{t_0}0_{\hat{t}_1}$ and $_{\hat{t}_0}0_{t_0}u'_{\hat{t}_0}$ are admissible, by linearity $_{\hat{t}_0}u_{t_0}0_{\hat{t}_1} + {}_{\hat{t}_0}0_{t_0}u_{\hat{t}_1} = {}_{\hat{t}_0}u_{t_0}u'_{\hat{t}_1}$ is admissible for any $u \in U_{(\hat{t}_0,t_0]}$, $u' \in U_{(\hat{t}_0,\hat{t}_1]}$.

(C1) or (C2) imply via equivalence classes of input the condition (C3).

(<u>C3</u>) "Any reduced state description of a linear object Θ is such that $(\sigma,u) \in D_{R(t_0,t]}$ for any $\sigma \in \Sigma_R(t_0)$ and any $u \in U_{(t_0,t]}$, t arbitrary in $\hat{I}$, if (C1) (or (C2)) is satisfied by Θ" i.e. briefly "$D_{R(t_0,t]} = \Sigma_R(t_0) \times U_{(t_0,t]}$."

The proof of (C3) is simple: $(\Sigma_R,\bar{A}_R)$ is a reduced state description and $\sigma \in \Sigma_R(t_0) \Longrightarrow \exists u_0 \in U_{(\hat{t}_0,t_0]} \ni u_0$ takes $\sigma_{\hat{t}_0}$ into σ, since σ must be reachable. As u_0 can be followed by any $u \in U_{(t_0,t]} \Longrightarrow (\sigma,u) \in D_{R(t_0,t]}$ by definition of reachability, for any $u \in U_{(t_0,t]}$.

From all these discussions for a linear object $\mathcal{O}$ satisfying (C1), we obtain the separation property and the linearity of the I-O-S-R as defined by Zadeh [ZA2]. This is summarized in the next theorem.

THM.III.2.4: Let $\mathcal{O}$ be a linear object that satisfies ((C1) Note III.2.7). Then the object $\mathcal{O}$ can be given a reduced state description $(\Sigma_R, \bar{A}_R)$ with the following properties:

(i) $\bar{A}_R$ has "the separation property" i.e.
$\bar{A}_{R(t_0,t]}(\sigma,u) = \bar{A}_{R(t_0,t]}(\sigma,0) + \bar{A}_{R(t_0,t]}(0,u)$
for all $\sigma \in \Sigma(t_0)$ and for all $u \in U_{(t_0,t]}$,
$(t_0,t] \subset \hat{I}$ is arbitrary.

(ii) $\bar{A}_R$ is "zero input linear" i.e.
$\bar{A}_{R(t_0,t]}(\sigma_1 + a\sigma_2, 0) = \bar{A}_{R(t_0,t]}(\sigma_1,0) + a\bar{A}_{R(t_0,t]}(\sigma_2,0)$

(iii) $\bar{A}_R$ is "zero state linear" i.e.
$\bar{A}_{R(t_0,t]}(0,u_1 + au_2) = \bar{A}_{R(t_0,t]}(0,u_1) + a\bar{A}_{R(t_0,t]}(0,u_2)$ both (ii) and (iii) are for all σ_1, $\sigma_2 \in \Sigma_R(t_0)$, for all u_1, $u_2 \in U_{(t_0,t]}$ and for all $a \in \mathbb{R}$.

PROOF: The proof follows directly from Thm.III.2.3, the condition (C1) and the discussion of Note III.2.7.

NOTE III.2.8: That (C1) has to be assumed is shown by the example of the object "with complete memory." Let $\mathcal{O}$ be given by: $U_{(t_0,t_1]} \triangleq \{u(t) : u(t) = a \in \mathbb{R}$ $\forall t \in (\hat{t}_0,\hat{t}_1]$ and $\{a : \exists u(t) \in U_{(\hat{t}_0,\hat{t}_1]} \ni u(t) = a\} = \mathbb{R}\}$ $R_{(\hat{t}_0,\hat{t}_1]} \triangleq \{(u,ku) : u \in U_{(\hat{t}_0,\hat{t}_1]}$ and $k \in \mathbb{R}$ is fixed$\}$. The object $\mathcal{O}$ defined by $R_{(\hat{t}_0,t_1]}$ is linear but clearly $\nexists\, u_0 \in U_{(\hat{t}_0,t_0]} \ni u_0 0 \in U_{(\hat{t}_0,t]}$, $t > t_0$ unless $u_0 = 0$.

III.3--Time-Invariant Objects and Properties of the State Description

NOTE III.3.1: According to Zadeh and Balakrishnan, and although Zadeh defines the concept of "weak time-invariance," the same way we define our "time invariant objects," the definition of a time invariant object is based on the I-O-S-R. That $\Sigma(t)$ is the same set for all t is part of this definition [BA4]. Our task here is again to start with the more basic definition of time-invariant objects and get the afore-mentioned definition as a result under $\mathcal{A}3$. Contrary to the definition of linearity, Def.III.2.1, where the existence interval could be any finite or semi-finite interval and not hurt linearity, the concept of time invariance requires from the object, $\hat{I} = (-\infty,\infty)$ as the existence interval. From a time invariant object we at least expect that it does not change its main properties as regard to the inputs and outputs, i.e., for example an input u admissible from t_1 on, must be admissible from any t_2 on and must yield the same output y whenever applied. As the existence interval is finite means that all basic I-O pairs are defined on $\hat{I} = (\hat{t}_0,\hat{t}_1]$, where both $\hat{t}_0$ and $\hat{t}_1$ are finite, one cannot speak of an input being admissible for $t_2 < \hat{t}_0$ or for $t_2 > \hat{t}_1$. Starting to exist and dying are such important properties reflecting the time varying aspect of anything that even theoretical objects possessing these properties must be expected to be time varying. Thus, although some "semi time-invariance" can be defined for objects with finite existence interval, the only real (expected) definition of time invariance can be given for objects that exist forever.

<u>NOT.III.3.1</u>: From now on, the existence interval $\hat{I}$ will be $(-\infty,\infty)$ for the objects under consideration (this was already mentioned in Note I.3.7).

<u>NOT.III.3.2</u>: Let $f(\cdot)$ be a function defined on the domain $D \subset \mathbb{R}$. Δ_τ is the operator defined on the space of functions with domain D by $\Delta_\tau f(t) \triangleq f(t-\tau)$ $\forall t \in D$ and where τ can be any finite real number. The domain of $\Delta_\tau f$ is the set $D + \tau \triangleq \{t + \tau : t \in D\}$.

<u>DEF.III.3.1</u>:

An object $\mathcal{O}$ whose existence interval $\hat{I} = (-\infty,\infty)$ is called TIME-INVARIANT iff $(u,y)\in R_{\hat{I}} \Longleftrightarrow (\Delta_\tau u,\Delta_\tau y)\in R_{\hat{I}}$.

<u>FACT III.3.1</u>: The object $\mathcal{O}$ is time invariant iff $(u,y)\in R_{(t_0,t_1]} \Longrightarrow (\Delta_\tau u,\Delta_\tau y)\in R_{(t_0+\tau,t_1+\tau]}$ for all intervals $(t_0,t_1]$.

<u>PROOF</u>:

" $\Longleftarrow$ " is trivial.

" $\Longrightarrow$ " If $\mathcal{O}$ is time invariant then: $(u,y)\in R_{(t_0,t_1]} \Longrightarrow \exists(\hat{u},\hat{y})\in R_{\hat{I}} \ni (\hat{u},\hat{y})/_{(t_0,t_1]} = (u,y)$. But we have $(\hat{u},\hat{y})\in R_{\hat{I}} \Longrightarrow (\Delta_\tau\hat{u},\Delta_\tau\hat{y})\in R_{\hat{I}}$ by time invariance. Hence $(\Delta_\tau\hat{u},\Delta_\tau\hat{y})/_{(t_0+\tau,t_1+\tau]} = (\Delta_\tau u,\Delta_\tau y)\in R_{(t_0+\tau,t_1+\tau]}$.

<u>DEF.III.3.2</u>:

The TRANSLATE Δ_τ OF AN EQUIVALENCE CLASS is defined to be: $\Delta_{-\tau}H_{t_0}[u_0] \triangleq \{u'\in U_{(-\infty,t_0+\tau]} : u' \simeq \Delta_\tau u_0\}$

<u>NOTE III.3.2</u>: The minus sign in $\Delta_{-\tau}H_{t_0}[u_0]$ is strictly notational, we could as well have used Δ_τ. The reason which made us choose $\Delta_{-\tau}$ is, when Δ_τ is applied to a function $f(t)$ it changes its argument to $f(t-\tau)$, however as we shall soon see in Thm.III.3.1, $\Delta_{-\tau}H_{t_0}[u_0]$ is the same equivalence class as $H_{t'_0}[\Delta_\tau u_0]$ where this time $t'_0 = t_0 + \tau$, i.e., the argument has been modified by τ instead of $-\tau$.

NOTE III.3.3: It will make more sense to speak of "the translate of an equivalence class" once the next theorem is proved. We have to show that $\Delta_\tau H_{t_0}[u_0]$ is an equivalence class, however, it may be clear intuitively for a time invariant object. So we prove that when an equivalence class is shifted, it still contains nothing but the shifted version of the inputs it had before the translation.

THM.III.3.1: For a time invariant object $\mathcal{O}$, $u'_0 \in H_{t_0}[u_0] \iff \Delta_\tau u'_0 \in \Delta_{-\tau} H_{t_0}[u_0]$, i.e., $\Delta_\tau H_{t_0}[u_0] = H_{t_0} + {}_\tau[\Delta_\tau u_0]$, for any $\tau \in \mathbb{R}$.

PROOF: Let $u'_0 \in H_{t_0}[u_0]$. By definition of $H_{t_0}[u_0]$, $u'_0 \in H_{t_0}[u_0] \Longrightarrow u'_0 \simeq u_0$. Then

(i) $(\Delta_\tau u'_0) 0 u \in U_{(-\infty, t]} \iff u'_0 0 (\Delta_{-\tau} u) \in U_{(-\infty,\ t-\tau]}$ by time invariance.

$\iff u_0 0 (\Delta_{-\tau} u) \in U_{(-\infty, t-\tau]}$ since $u_0 \simeq u'_0$

$\iff (\Delta_\tau u_0) 0 u \in U_{(-\infty, t]}$ by time invariance

for $u \in U_{(t_0 + \tau, t]}$, is $t > t_0 + \tau$, proving $(\Delta_\tau u'_0) 0 u$ is admissible iff $(\Delta_\tau u_0) 0 u$ is.

(ii) Since they are admissible there exists y' and y such that: $((\Delta_\tau u'_0)0u, y') \in R_{(-\infty,t]} \Longleftrightarrow (u'_0 0(\Delta_{-\tau}u), \Delta_{-\tau}y') \in R_{(-\infty,t-\tau]}$ and $((\Delta_\tau u_0)0u, y) \in R_{(-\infty,t]} \Longleftrightarrow (u_0 0(\Delta_{-\tau}u), \Delta_{-\tau}y) \in R_{(-\infty,t-\tau]}$ both by time invariance. But $\Delta_{-\tau}y'/_{(t_0,t-\tau]} = \Delta_{-\tau}y/_{(t_0,t-\tau]}$ since $u'_0 \simeq u_0$. Thus $y'/_{(t_0+\tau,t]} = y/_{(t_0+\tau,t]}$ which proves $\Delta_\tau u'_0 \simeq \Delta_\tau u'$ and therefore $\Delta_\tau u' \in \Delta_{-\tau}H_{t_0}[u_0]$.

Letting $\Delta_\tau u'_0 \in \Delta_{-\tau}H_{t_0}[u_0]$, which is true iff $\Delta_\tau u'_0 \simeq \Delta_\tau u_0$ (Def.III.3.2), we can proceed as above to show $u'_0 \simeq u_0$ giving us $u'_0 \in H_{t_0}[u_0]$.

That $\Delta_{-\tau}H_{t_0}[u_0] = H_{t_0+\tau}[\Delta_\tau u_0]$ clearly follows from above.

<u>NOTE III.3.4</u>: For a time invariant object $\mathcal{O}$, a shifted equivalence class is still an equivalence class, justifying Def.III.3.2. It would be nice to show that this property alone makes $\mathcal{O}$ time invariant, that is a converse to Thm.III.3.1. However, this is not true in general as the following counter example shows: Let $\mathcal{O}$ be given by the unique pair, $R_{(-\infty,\infty)} = (u(t), e^{-|t|})$ where $u(t) = C, \forall t$. Then $H_{t_0}[u_0] = \{c/_{(-\infty,t_0]}\}$ is the unique

equivalence class $\forall t_0 \in \hat{I}$, and trivially $u \in H_{t_0}[u_0] \Longleftrightarrow \Delta_\tau u \in \Delta_\tau H_{t_0}[u_0]$. But the object is not time invariant, for $\Delta_\tau e^{-|t|}$ is not an admissible output.

Of the following two corollaries, the second one is a result we were aiming for.

<u>COR.III.3.1</u>: $\Delta_{-\tau}(\Delta_\tau H_{t_0}[u_0]) = \Delta_\tau(\Delta_{-\tau} H_{t_0}[u_0]) = H_{t_0}[u_0]$.

<u>COR.III.3.2</u>: The reduced state space $\Sigma(t)$ for a time invariant object can be taken the same set $\forall t \in (-\infty, \infty)$.

<u>PROOF</u>: All we need to show is that H_{t_0} and H_{t_1} have the same cardinality for any t_0 and t_1. That there exists a well-defined and 1-1 (since invertible) mapping $T : \mathcal{H}_{t_0} \to \mathcal{H}_{t_1}$ defined by $T(H_{t_0}[u_0]) \triangleq \Delta_{-\tau} H_{t_0}[u_0]$, where $\tau = t_1 - t_0$ is clear by Thm.III.3.1 and Cor.III.3.1. It is also clear that T is onto, since any $H_{t_0}[u_0] \in \mathcal{H}_{t_1}$ is the image of the equivalence class $H_{t_0}[\Delta_{-\tau} u_0] \in H_{t_0}$ because of $T(H_{t_0}[\Delta_{-\tau} u_0]) = \Delta_{-\tau} H_{t_0}[\Delta_{-\tau} u_0] = H_{t_1}[u_0]$. This proves card $\mathcal{H}_{t_0}$ = card $\mathcal{H}_{t_1}$. As by Thm.II.4.3 any reduced state space is nothing but $\Sigma^*(t)$, the one obtained by use of equivalence classes of inputs, the same set can be put into 1-1, onto correspondence with both $\mathcal{H}_{t_0}$ and $\mathcal{H}_{t_1}$ for any $t_0 \neq t_1$, i.e., a unique set suffices to be taken the state space $\forall t \in \hat{I}$.

<u>NOTE III.3.5</u>: Of course we are completely free to select the set $\Sigma_R(t)$ for each t as long as it has the same cardinality as H_t, but any other choice, than the same set $\forall t$ seems to be artificial, unless there be a necessity.

We achieve our next goal with Thm.III.3.2 by showing the form the I-O-S-R takes when the object is time-invariant. It is here that we have to remember the distinction made between A_I and $\bar{A}_I$ in Con.I.3.1.

<u>NOT.III.3.3</u>: Let $\sigma_0(t_0)\in\Sigma(t_0)$ denote the state corresponding to the class $H_{t_0}[u_0]$ for the time-invariant object $\mathcal{O}$, where $\Sigma(t)$ is reduced $\forall t\in\hat{I}$. Then $\sigma_0(t_0 + \tau)\in\Sigma(t_0 + \tau)$ will denote the state corresponding to the class $\Delta_{-\tau}H_{t_0}[u_0]$, $\forall\tau\in(-\infty,\infty)$, i.e., $\sigma_0(t_0 + \tau) = C^{-1}_{t_0 + \tau}(\Delta_{-\tau}H_{t_0}[u_0])$.

<u>THM.III.3.2</u>: An object $\mathcal{O}$ existing over $(-\infty,\infty)$ is time invariant iff it has a reduced state description $(\Sigma_R,\bar{A}_{IR})$ such that:

(i) $(\sigma_0(t_0),u)\in D_{R(t_0,t]} \Longleftrightarrow$

$(\sigma_0(t_0 + \tau),\Delta_\tau u)\in D_{R(t_0 + \tau,t_1 + \tau)}$ $\forall t_1$ and $\tau\in(-\infty,\infty)$

(ii) $A_{R(t_0,t_1]}(\sigma_0(t_0),u) =$

$A_{R(t_0 + \tau,t_1 + \tau]}(\sigma_0(t_0 + \tau),\Delta_\tau u)$ $\forall t\in(t_0,t_1]$

PROOF:

"$\Longrightarrow$" Let Θ be time invariant. Then:

(i) Using Thm.II.4.3, as usual, we do the proof for $(\Sigma^*,\bar{A}^*)$. By Fact II.3.2 $(\sigma(t_0),u)\in D^*_{(t_0,t_1]} \Longleftrightarrow$ $u_0 \circ u \in U_{(-\infty,t_1]}$ for any $u_0 \in C_{t_0}(\sigma_0) = H_{t_0}[\tilde{u}]$.
By Fact III.3.1 $u_0 \circ u \in U_{(-\infty,t_1]} \Longleftrightarrow \Delta_\tau(u_0 \circ u) =$ $(\Delta_\tau u_0)\circ(\Delta_\tau u) \in U_{(-\infty,t_1+\tau]}$ and by Thm.III.3.1 $u_0 \in H_{t_0}[\tilde{u}] \Longleftrightarrow \Delta_\tau u_0 \in \Delta_{-\tau}H_{t_0}[\tilde{u}_0] \triangleq C_{t_0+\tau}(\sigma(t_0+\tau))$.
Thus we have $(\Delta_\tau u'_0)\circ(\Delta_\tau u)\in U_{(-\infty,t_1]}$ for any u'_0, the quantifier "for any" being well-placed due to Thm.III.3.1. Thus, again by Fact II.3.2 $(\sigma(t_0+\tau),\Delta_\tau u)\in D^*_{(t_0+\tau,t_1+\tau]}$.

(ii) $y = \bar{A}^*_{(t_0,t_1]}(\sigma(t_0),u) \Longleftrightarrow \exists(\hat{u},\hat{y})\in R_{(-\infty,t_1]} \ni$ $\hat{u}/_{(-\infty,t_0]}\in C_{t_0}(\sigma_0)$ and $(\hat{u},\hat{y})/_{(t_0,t]} = (u,y)$ by definition of $\bar{A}^*_I$. By Fact III.3.1 $(\hat{u},\hat{y})\in R_{(-\infty,t_1]} \Longleftrightarrow (\Delta_\tau\hat{u},\Delta_\tau\hat{y})\in R_{(-\infty,t_1+\tau]}$ and hence $(\Delta_\tau\hat{u},\Delta_\tau\hat{y})/_{(t_0+\tau,t_1+\tau]} = (\Delta_\tau u,\Delta_\tau y)$. Moreover, $\exists\, \tilde{u}\in U_{(-\infty,t_0]} \ni \hat{u}/_{(-\infty,t_0]} \in C_{t_0}(\sigma_0) = H_{t_0}[\tilde{u}]$. By Thm.III.3.1, $\Delta_\tau(\hat{u}/_{(-\infty,t_0]}) \in \Delta_{-\tau}H_{t_0}[\tilde{u}] =$ $C_{t_0+\tau}(\sigma(t_0+\tau))$. And again by definition of $\bar{A}^*_I$, $\Delta_\tau y = \bar{A}^*_{(t_0+\tau,t_1+\tau]}(\sigma(t_0+\tau),\Delta_\tau u)$ which implies $A^*_{(t_0,t_1]}(\sigma(t_0),u) =$

$A^*_{(t_0 + \tau,t_1 + \tau]}(\sigma(t_0 - \tau),\Delta_\tau u)$.

"$\Longleftarrow$" Let (i) and (ii) be true. Then:

$(u,y)\in R_{(t_0,t_1]} \Longleftrightarrow \exists\sigma\in\Sigma^*(t_0) \ni y(t) = A^*_{(t_0,t_1]}(\sigma(t_0),u)$ by axiom (M1). Then (ii) $\Longrightarrow A^*_{(t_0,t_1]}(\sigma(t_0),u) =$ $A^*_{(t_0 + \tau,t_1 + \tau]}(\sigma(t_0 + \tau),\Delta_\tau u)$ for any $\tau\in(-\infty,\infty)$, or that $\Delta_\tau y = \bar{A}^*_{(t_0,t_1]}(\sigma(t_0 + \tau),\Delta_\tau u) \Longrightarrow$ $(\Delta_\tau u,\Delta_\tau y)\in R_{(t_0 + \tau,t + \tau]}$ proving the theorem.

<u>NOTE III.3.6</u>: A word about half reduced state descriptions closes this section. A proof in the same lines as Thm.III.3.1 can be given to show that the family $\mathcal{H}'_{t_0}$ based on any half reduced partitioning can be translated by τ, to yield the family $\mathcal{H}'_{t_0 + \tau}$ which has the same structure as $\mathcal{H}'_{t_0}$. More precisely, any $H'_{t_0}[u_0]\in\mathcal{H}'_{t_0}$ can give rise to a $\Delta_{-\tau}H'_{t_0}[u_0]$, which can be shown to be equal to $H'_{t_0 + \tau}[\Delta_\tau u_0]\in\mathcal{H}'_{t_0 + \tau}$ and to constitute a half reduced partitioning of $U_{(-\infty,t_0 + \tau]}$. This then will allow us to keep the same half reduced state space $\Sigma_{HR}(t)$, $\forall t\in(-\infty,\infty)$ and to have a property of A_{IHR} that is similar to the one in Thm.III.3.2.

III.4--An Application of Equivalence Classes of Inputs to Lumped Objects

NOTE III.4.1: In this section we shall deal with objects which have a representation of the form:

$$\frac{dX(t)}{dt} = AX(t) + Bu(t)$$
$$y(t) = CX(t) + Du(t) \qquad \text{III.4.1}$$

where A_{nxn}, B_{nx1}, C_{1xn}, D_{1x1} are constant matrices. Our concern here will be to see under what conditions on A, B, C and D equations III.4.1 yield a reduced state description with minimal dimension. This result will be useful in section IV.5. We start with the precise definition of the object under consideration.

DEF.III.4.1:

A linear, time-invariant object will be called $\mathcal{O}_L$ iff it satisfies:

(i) $U_{L(-\infty,\infty)} \triangleq \{u(t)$: u(t) is a regular distribution with support bounded at left and which is summable on $(-\infty,b)$, for all finite $b \in \mathbb{R}\}$, is the input space.

(ii) $R_{L(-\infty,\infty)} \triangleq \{(u_0,y_0) : u_0 \in U_{L(-\infty,\infty)}$ and y_0 satisfies III.4.1 for this $u_0\}$. For a given $u \in U_L$, we will denote by T, the point such that $u(t) \equiv 0 \;\; \forall t < T$.

<u>FACT III.4.1</u>: It follows from our definition of U_L that any input can follow and can be followed by any other, i.e., all concatenations are permissible. For: $u_0 \in U_{L(-\infty,t_0]}$ and $u \in U_{L(t_0,t]}$ implies both are summable making $u_0 0 u$ summable on $(-\infty,t]$.

<u>LEMMA III.4.1</u>: Let $(u_0 0 u, y_0 0 y) \in R_{L(-\infty,t_1]}$ where $(u_0,y_0) \in R_{L(-\infty,t_0]}$ with $t_0 < t_1$. Then: $y(t) =$

$$Ce^{A(t-t_0)}X(t_0) + \int_{t_0}^{t} Ce^{A(t-\tau)}Bu(\tau)d\tau + Du(t) \qquad \text{III.4.2}$$

$\forall t \in (t_0,t_1]$ where the integrals are in the Lebesgue sense

$$\text{and } X(t_0) = \int_{-\infty}^{t_0} e^{A(t_0-\tau)}Bu_0(\tau)d\tau \qquad \text{III.4.3}$$

<u>PROOF</u>:
$$\frac{dX(t)}{dt} = AX(t) + B(u_0 0 u)(t)$$

III.4.1 gives: III.4.4

$$(y_0 0 y)(t) = CX(t) + D(u_0 0 u)(t)$$

Since we are talking of distributions we can write (see [ZE1] or [SC2]):

$$I\delta'(t)*X(t) = A\delta(t)*X(t) + B(u_0 0 u)(t) \qquad \text{III.4.5}$$

where $\delta(t)$ is the delta-distribution, I the identity-matrix and * denotes the convolution operation (A.2.9). Then III.4.5 yields:

$$[I\delta'(t) - A\delta(t)]*X(t) = B(u_0 0 u)(t) \qquad \text{III.4.6}$$

$1(t)e^{At}$, where $1(t)$ is the unit step distribution, is the convolution inverse of $I\delta'(t) - A\delta(t)$ as one can easily verify. Therefore:

$$X(t) = 1(t)e^{At}B*(u_0 0 u)(t) = {}_T\int^t e^{A(t-\tau)}B(u_0 0 u)(\tau)d\tau \qquad \text{III.4.7}$$

since the convolution of two locally summable regular distributions can be written as the right hand side of III.4.7 (Thm.A.2.7). Moreover:

$$\begin{aligned}(y_0 0 y)(t) &= Ce^{At}B1(t)*(u_0 0 u)(t) + D(u_0 0 u)(t) = \\ &\qquad C{}_T\int^t e^{A(t-\tau)}B(u_0 0 u)(\tau)d\tau + D(u_0 0 u)(t), \\ &\qquad \forall t \in (-\infty, t_1] \\ &= Ce^{At}{}_T\int^{t_0} e^{-A\tau}Bu_0(\tau)d\tau + \\ &\qquad C{}_{t_0}\int^t e^{A(t-\tau)}Bu(\tau)d\tau + Du(t)\ , \forall t \in (t_0, t_1] \\ &= Ce^{A(t-t_0)}{}_T\int^{t_0} e^{A(t_0-\tau)}Bu_0(\tau)d\tau + \\ &\qquad C{}_{t_0}\int^t e^{A(t-\tau)}Bu(\tau)d\tau + Du(t). \qquad \text{III.4.8}\end{aligned}$$

Using III.4.7 at $t = t_0$ we obtain III.4.2 and III.4.3.

<u>NOTE III.4.2</u>: The expression III.4.8 can also be written as:

$$y(t) = [Ce^{At}B1(t) + D\delta(t)]*u(t) \text{ for any } (u,y) \in R_L, \qquad \text{III.4.9}$$

where this form, namely $Ce^{At}B1(t) + D\delta(t)$, can be viewed as the convolution representation of the object $\mathcal{O}_L$. This gave us the initial idea in Chapter IV about how to find a state description for more general objects of the form $y = w*u$.

NOTE III.4.3: The restriction in Def.III.4.1 that any input has support bounded at left, is necessary for the existence of the integral in III.4.7 and alike, since A may have positive eigenvalues. Another alternative then would be to assume that A has negative eigenvalues and then let the input space be the space of regular distributions that are summable on $(-\infty,b]$ for all finite b.

We would also like to note that III.4.2 and III.4.3 are true for any $t\in(t_0,\infty)$ when $(u_0 0 u, y_0 0 y)\in R_{L(-\infty,\infty)}$ are such that $(u,y)\in R_{L(t_0,\infty)}$.

LEMMA III.4.2: Let $u_0, u'_0 \in U_{L(-\infty,t_0]}$. Then $u_0 \simeq u'_0 \iff$

$$\int_{-\infty}^{t_0} Ce^{A(t-\tau)}B[u_0(\tau)-u'_0(\tau)]d\tau = 0, \ \forall t > t_0, \qquad \text{III.4.10}$$

PROOF: $u_0 0 u$ and $u'_0 0 u$ are admissible for any $u\in U_{L(t_0,\infty)}$ by Fact III.4.1.

Now let $(u_0 0 u, y_0)$ and $(u'_0 0 u, y_0')\in R_{L(-\infty,\infty)}$. Then expression III.4.9 gives: $y_0(t) = Ce^{At}B1(t)*(u_0 0 u)(t) + D\delta(t)*(u_0 0 u)(t)$ and $y'_0(t) = Ce^{At}B1(t)*(u'_0 0 u)(t) + D\delta(t)*(u'_0 0 u)(t) \ \forall t$. As the object is linear:

$$y_0(t) - y'_0(t) = Ce^{At}B1(t)*[(u_0 0 u)(t)-(u'_0 0 u)(t)] + D[(u_0 0 u)(t)-(u'_0 0 u)(t)], \ \forall t\in(-\infty,\infty).$$

$$= Ce^{At}B1(t)([(u_0-u'_0)_0 0](t) + D[(u_0-u'_0)_0 0](t) \ \forall t\in(-\infty,\infty).$$

In our case, as indicated above, $u_0 \approx u'_0 \Longrightarrow y_0/(t_0,\infty) = y'_0/(t_0,\infty)$, i.e.,

$$u_0 \approx u'_0 \Longleftrightarrow Ce^{At}B1(t)*[(u_0-u'_0)_0 0](t) = 0 \quad \forall t > t_0 \qquad \text{III.4.11}$$

$$u_0 \approx u'_0 \Longleftrightarrow \int_{-\infty}^{t} Ce^{A(t-\tau)}B[(u_0-u'_0)_0 0](\tau)d\tau = 0 \quad \forall t > t_0 \text{ or}$$

$$\Longleftrightarrow \int_{-\infty}^{t_0} Ce^{A(t-\tau)}B[u_0(\tau)-u'_0(\tau)]d\tau = 0 \quad \forall t > t_0$$

which is III.4.10

<u>THM.III.4.1</u>: Let $\Sigma_L(t_0) \triangleq \{X(t_0)\in\mathbb{R}^n : \exists u_0\in U_{(-\infty,t_0]} \ni X(t_0) = \int_{-\infty}^{t_0} e^{A(t_0-\tau)}Bu_0(\tau)d\tau\}$ for $t_0 > -\infty$ and $\Sigma_L(t_0) \triangleq \{0\}$ for $t_0 = -\infty$. Also let $y(t) = A_{L(t_0,t]}(X(t_0),u)$, for $X(t_0)\in\Sigma_L(t_0)$ and $u\in U_{L(t_0,t]}$, be the expression III.4.2. Then (Σ_L, A_{L_I}) is a half reduced state description of $\mathcal{O}_L$ and $\Sigma_L(t_0)$ is a linear subspace of $\mathbb{R}^n$.

<u>PROOF</u>: To prove the theorem we show that (Σ_L, A_L) is based on a half reduced partitioning. Then Thm.II.4.4 completes our task.

Consider $\mathcal{H}'_{t_0} = H'_{t_0}[u_0] : u_0\in U_{L(-\infty,t_0]}$ where we define:

$$H'_{t_0}[u_0] \triangleq \{u'_0\in U_{(-\infty,t_0]} : \int_{-\infty}^{t_0} e^{A(t_0-\tau)}Bu'_0(\tau)d\tau = \int_{-\infty}^{t_0} e^{A(t_0-\tau)}Bu_0(\tau)d\tau\} \qquad \text{III.4.12}$$

(i) $u_0\in H'_{t_0}[u_0]$ trivially.

(ii) If u_0 and $u'_0\in U_{L(-\infty,t_0]}$ then either

$$\int_{-\infty}^{t_0} e^{A(t_0-\tau)}Bu_0(\tau)d\tau = \int_{-\infty}^{t_0} e^{A(t_0-\tau)}Bu'_0(\tau)d\tau$$

in which case $H'_{t_0}[u_0] = H'_{t_0}[u'_0]$, or

$$\int_{-\infty}^{t_0} e^{A(t_0-\tau)}Bu_0(\tau)d\tau \neq \int_{-\infty}^{t_0} e^{A(t_0-\tau)}Bu'_0(\tau)d\tau$$

in which case $H'_{t_0}[u_0] \cap H'_{t_0}[u'_0] = \phi$.

(iii) Let $u'_0 \in H'_{t_0}[u_0]$. Then

$$\int_{-\infty}^{t_0} e^{A(t_0-\tau)}Bu_0(\tau)d\tau = \int_{-\infty}^{t_0} e^{A(t_0-\tau)}Bu'_0(\tau)d\tau$$

$$\Longrightarrow \int_{-\infty}^{t_0} e^{-A\tau}Bu_0(\tau)d\tau = \int_{-\infty}^{t_0} e^{-A\tau}Bu'_0(\tau)d\tau$$

$$\Longrightarrow Ce^{At}\int_{-\infty}^{t_0} e^{-A\tau}Bu_0(\tau)d\tau =$$

$$Ce^{At}\int_{-\infty}^{t_0} e^{-A\tau}Bu'_0(\tau)d\tau \ \forall t \Longrightarrow$$

$\int_{-\infty}^{t_0} Ce^{A(t-\tau)}B[u_0(\tau)-u'_0(\tau)]d\tau = 0$ in particular $\forall t > t_0 \Longrightarrow u_0 \simeq u'_0$ by Lemma III.4.2 proving that $\mathcal{H}'_{t_0}$ is indeed a half reduced partitioning.

Moreover each $X(t_0) \in \Sigma_L(t_0)$ represents one and only one $H'_{t_0}[u_0]$ due to expressions III.4.3 and III.4.12. This is to say that there exists a one to one, onto mapping $C'_{t_0} : \Sigma_L(t_0) \to \mathcal{H}'_{t_0}[u_0]$. Thus using Thm.III.4.4, $\Sigma_L(t_0)$ qualifies for the state space of $\mathcal{O}_L$.

The I-O-S-R, $y(t) = A_{L(t_0,t]}(X(t_0),u)$, defined by III.4.2 is such that clearly $y(t)$ is the (t_0,∞) portion of the response to $u_0 0 u$, for any $u_0 \in C'_{t_0}(X(t_0))$, due to III.4.8. Hence $A_{L(t_0,t]}$ qualifies for the I-O-S-R based on the half reduced partitioning III.4.12.

Finally it is simple to see that $\Sigma_L(t_0)$ is a linear space since: $X_1(t_0)$, $X_2(t_0)\in\Sigma_L(t_0) \Longrightarrow$ $\exists u_0$, $u'_0\in U_{(-\infty,t_0]}$ $\ni X_1(t_0) = \int_{-\infty}^{t_0} e^{A(t_0-\tau)}Bu_0(\tau)d\tau$ and $X_2(t_0) = \int_{-\infty}^{t_0} e^{A(t_0-\tau)}Bu'_0(\tau)d\tau$. As $\mathcal{O}_L$ is linear $u_0 + au'_0\in U_{(-\infty,t_0]}$, Fact III.2.3, which implies $X_1(t_0) + aX_2(t_0)\in\Sigma_L(t_0)$, for $a\in\mathbb{R}$.

NOTE III.4.4: However it is not generally true that $\Sigma_L(t_0)$ as defined in Thm.III.4.1 is a reduced state space. The necessary and sufficient condition III.4.10 does not require the defining relation of III.4.12 to hold, for the inputs u_0 and u'_0 to be equivalent. Condition III.4.12 is a sufficient one for III.4.10 or for that matter for $u_0 \approx u'_0$. The following theorem and its corrollaries tell us when III.4.12 becomes also necessary for III.4.10, i.e., when $\Sigma_L(t_0)$ becomes reduced, or if not, what is the dimension of the reduced state description, etc.

NOTE III.4.5: We assume that in the equations III.4.1 the matrix A is in, or has been brought to, its Jordan Canonical form. This is no restriction at all, at least theoretically, since every matrix has a Jordan equivalent [HO]. We further assume that:

$$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_k \end{bmatrix}, \quad A_k = \begin{bmatrix} J_1^{(k)} & 0 & \cdots & 0 \\ 0 & J_2^{(k)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{n_k}^{(k)} \end{bmatrix}, \quad J_j^{(k)} = \begin{bmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & 1 \\ 0 & 0 & \cdots & & \lambda_k \end{bmatrix}$$

for $k = 1, 2, \ldots, k$ for $j = 1, 2, \ldots, n_k$

III.4.13

In A_k the size of $J_j^{(k)}$ decreases as j increases and each A_k corresponds to a different eigenvalue λ_k. If the matrix A has a diagonalizable part, with or without distinct eigenvalues, we view each entry on the diagonal as a 1 x 1 Jordan block. We also partition the row and column vectors C and B conformal to the partition of A, i.e.

$$C = [C_1 C_2 \ . \ . \ . \ C_k] \text{ and } B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_k \end{bmatrix} \qquad \text{III.4.12}$$

such that the product $C_k A_k B_k$ is defined for k = 1, 2,...,K.

<u>THM.III.4.2</u>: Suppose that:

(i) A_k is formed of a unique elementary Jordan block $J^{(k)}$, and

(ii) $c_k^{(1)} b_k^{(d_k)} \neq 0$, where $c_k^{(1)}$ is the first element of C_k and $b_k^{(d_k)}$ is the last element of B_k, d_k being the multiplicity of λ_k because of (i).

Then $(\Sigma_L, \bar{A}_{L_I})$ is a reduced state description, where $\Sigma_L(t_0)$ as defined in Thm.III.4.1 is $\mathbb{R}^n$ $\forall t_0 \in (-\infty,\infty)$, and $\bar{A}_{L_I}$ is as defined in III.4.2.

<u>PROOF</u>: Using Lemma III.4.2 for u_0 and $u'_0 \in U_{(-\infty,t_0]}$, the expression III.4.10 takes the form $u_0 \simeq u'_0 \Longrightarrow$

$$\int_{-\infty}^{t_0} \sum_{k=1}^{k} C_k e^{A_k(t-\tau)} B_k [u_0(\tau) - u'_0(\tau)]d\tau = 0, \ \forall t>t_0 \qquad \text{III.4.14}$$

since the matrix e^{At} has the submatrices $e^{A_k t}$ on its diagonal and 0 submatrices elsewhere, like A had the A_k's on its diagonal. $C_k e^{A_k t} B_k$ is conformal by Note III.4.5. It was pointed out in the same note that each A_k corresponds to a different eigenvalue λ_k, so that two different terms of the summation in III.4.13 corresponding to, say k_1 and k_2 will yield terms containing $e^{\lambda_{k_1} t}$ and $e^{\lambda_{k_2} t}$ as factors. By summing up such terms there is no chance of cancelling one $e^{\lambda_{k_1} t}$ by another $e^{\lambda_{k_2} t}$ for all $t > t_0$. Thus:

$$u_0 \simeq u'_0 \Longleftrightarrow \int_{-\infty}^{t_0} C_k e^{A_k(t-\tau)} B_k [u_0(\tau)-u'_0(\tau)]d\tau = 0$$

$$\forall t > t_0 \qquad \text{III.4.15}$$

$$\Longleftrightarrow \int_{-\infty}^{t_0} C_k e^{J^{(k)}(t-\tau)} B_k [u_0(\tau)-u'_0(\tau)]d\tau = 0$$

$$\forall t > t_0 \qquad \text{III.4.16}$$

since by (i) each A_k is constituted of a single elementary Jordan block.

Using the matrix form of $e^{J^{(k)}t}$ as given by [CO], III.4.16 reduces to:

$$u_0 \simeq u'_0 \Longleftrightarrow \sum_{\substack{j,i=1\\ i\leqslant j}}^{d_k} c_k^{(i)} b_k^{(j)} \int_{-\infty}^{t_0} \frac{(t-\tau)^{j-i}}{(j-i)!} e^{\lambda_k(t-\tau)}[u_0(\tau) -$$

$$u'_0(\tau)]d\tau = 0 \quad \forall t > t_0 \quad k=1,2,\ldots,K \qquad \text{III.4.17}$$

We consider now the term which contains the highest power of t in the summation of III.4.16. t has the highest power when $j = d_k$ and $k = 1$, yielding the term

$$c_k^{(1)} b_k^{(d_k)} \int_{-\infty}^{t_0} \frac{(t-\tau)}{(d_k-1)!}^{d_k-1} e^{\lambda_k(t-\tau)}[u_0(\tau)-u'_0(\tau)]d\tau,$$

where $c_k^{(1)} b_k^{(d_k)} \neq 0$ by hypothesis.

If we expand $(t-\tau)^{d_k-1}$ and consider the term that contains t^{d_k-1}, it is of the form

$$\alpha_{d_k-1} t^{d_k-1} e^{\lambda_k t} \int_{-\infty}^{t_0} e^{-\lambda_k\tau}[u_0(\tau) - u'_0(\tau)]d\tau, \ \alpha_{d_k-1} \neq 0.$$

and it is the only term in III.4.17 with t^{d_k-1} as factor. Thus if the left hand side of III.4.17 has to be zero $\forall t > t_0$, the only way this can happen is to have:

$$\int_{-\infty}^{t_0} e^{-\lambda_k\tau}[u_0(\tau) - u'_0(\tau)]d\tau = 0 \qquad \text{III.4.18}$$

Proceeding in this manner at each step we will be left with a term of the form $\int_{-\infty}^{t_0} \tau^m e^{-\lambda_k \tau}[u_0(\tau)-u'_0(\tau)]d\tau$ that must equal zero, considering the results of the previous steps. Thus we will obtain:

$$u_0 \simeq u'_0 \Longleftrightarrow \int_{-\infty}^{t_0} \tau^m e^{-\lambda_k \tau}[u_0(\tau)-u'_0(\tau)]d\tau = 0 \qquad \text{III.4.18}$$

$$\text{for } m = 0,1\ldots,d_k-1$$

$$k = 1,2,\ldots,K$$

The defining relation III.4.12 for $H'_{t_0}[u_0]$, after cancellation of e^{At_0}, can be rewritten as:

$$H'_{t_0} = \{u'_0 \in U_{(-\infty,t_0]} : \int_{-\infty}^{t_0} e^{-A\tau}B[u_0(\tau)-u'_0(\tau)]d\tau = 0\} \qquad \text{III.4.19}$$

i.e., $u'_0 \in H'_{t_0}[u_0] \Longleftrightarrow \int_{-\infty}^{t_0} e^{-J^{(k)}\tau}B_k[u_0(\tau)-u'_0(\tau)]d\tau =$

$$0 \quad \text{for } k = 1,2,\ldots,K \qquad \text{III.4.20}$$

since $A_k = J^{(k)}$ by hypothesis. Writing the column vector $\int_{-\infty}^{t_0} e^{-J(k)\tau}B_k[u_0(\tau)-u'_0(\tau)]d\tau$, we get

$$u'_0 \in H'_{t_0}[u_0] \Longleftrightarrow \begin{bmatrix} \int_{-\infty}^{t_0} e^{-\lambda_k \tau} \sum_{i=1}^{dk} b_k^{(i)} \frac{(-\tau)^{i-1}}{(i-1)!}[u_0(\tau)-u'_0(\tau)]d\tau \\ \int_{-\infty}^{t_0} e^{-\lambda_k \tau} \sum_{i=2}^{dk} b_k^{(i)} \frac{(-\tau)^{i-2}}{(i-2)!}[u_0(\tau)-u'_0(\tau)]d\tau \\ \vdots \\ \int_{-\infty}^{t_0} e^{-\lambda_k \tau} \sum_{i=d_k}^{dk} b_k^{(i)} \frac{(-\tau)^{i-d_k}}{(i-d_k)!}[u_0(\tau)-u'_0(\tau)]d\tau \end{bmatrix} = 0$$

$$\text{for } k = 1,2,\ldots,K \qquad \text{III.4.21}$$

Starting with the last row of III.4.21 which has the unique term $b_k^{(d_k)} \int_{-\infty}^{t_0} e^{-\lambda_k \tau}[u_0(\tau)-u'_0(\tau)]d\tau$ with $b_k^{(d_k)} \neq 0$ by hypothesis, we see that $\int_{-\infty}^{t_0} e^{-\lambda_k \tau}[u_0(\tau)-u'_0(\tau)]d\tau$ must equal zero. Moving upwards, each previous step eliminating all terms except, $b_k^{(d_k)} \int_{-\infty}^{t_0} e^{-\lambda_k \tau}\tau^{m-1}[u_0(\tau)-u'_0(\tau)]d\tau$ at the m-th step, $m = 1,\ldots,d_k$ we finally obtain (changing m-1 to m):

$$u'_0 \in H'_{t_0}[u_0] \Longleftrightarrow \int_{-\infty}^{t_0} \tau^m e^{-\lambda_k \tau}[u_0(\tau)-u'_0(\tau)]d\tau \qquad \text{III.4.22}$$

for $m = 0,1,\ldots,d_k-1$

$k = 1,2,\ldots,K$

Combining III.4.18 with III.4.22 we see that $u'_0 \in H'_{t_0}[u_0] \Longleftrightarrow u'_0 \simeq u_0$, making the classes $H'_{t_0}[u_0]$ equivalence classes of inputs, i.e., the sufficient condition has turned out to be a necessary condition in this case, as it was indicated by Note III.4.4. Thus under the hypothesis (Σ_L, A_{L_I}) becomes based on equivalence classes of inputs. As equivalence classes are also a half reduced partitioning Thm.III.4.1 proves that (Σ_L, A_{L_I}) is a reduced state description under (i) and (ii).

That $\Sigma_L(t_0)$ is the n-dimensional Euclidean Space $\forall\, t_0 \in (-\infty,\infty)$ is quite obvious. It was proved in Thm. III.4.1 that $\Sigma_L(t_0)$ was linear (it had dimension n by definition), it became reduced here, giving these properties to the reduced state description $\forall t_0 \in (-\infty,\infty)$, since Θ_L is time invariant (Cor.III.3.2).

NOTE III.4.6: The following corollaries follow Thm. III.4.2. The proofs follow the same lines as the proof of Thm.III.4.2 and are not given. The results are for more general cases, the last one being the most general. Let the vectors C and B be partitioned as in III.4.12. The vectors C_k and B_k that pre- and post-multiply A_k are partitioned into submatrices:

$$C_k = [C_{k,1} C_{k,2} \cdots C_{k,n_k}] \text{ and } B_k = \begin{bmatrix} B_{k,1} \\ B_{k,2} \\ \vdots \\ B_{k,n_k} \end{bmatrix} \qquad \text{III.4.23}$$

COR.III.4.1: If the first entry $c_{k,1}^{(1)}$ of $C_{k,1}$ and the last entry of the vector $B_{k,1}$ are nonzero for $k = 1, 2, \ldots, K$ then:

$$\dim \Sigma_{LR}(t_0) = \sum_{k=1}^{K} \text{size } J_1^{(k)} \qquad \text{III.4.24}$$

i.e. the dimension of the reduced state space is equal to the sum of the sizes of the largest elementary Jordan blocks for different eigenvalues.

COR.III.4.2: Let again each A_k consist of a single elementary Jordan block and let the $\gamma_k + 1$ th element of C_k be nonzero, the first γ_k being zero and the

$\beta_k + 1$ th element of B_k be nonzero, the last β_k elements being zero, i.e.

$$C_k = [\underbrace{00\cdots\cdots0}_{\gamma_k\text{-zeroes}}\ c_k^{(\gamma_k+1)}\cdots\cdots c_k^{(d_k)}] \text{ and } B_k = \begin{bmatrix} b_k^{(1)} \\ \vdots \\ b_k^{(\beta_k+1)} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{III.4.25}$$

$c_k^{(\gamma_k+1)}$ and $b_k^{(\beta_k+1)} \neq 0$.

Then: $$\dim \Sigma_{LR}(t_0) = \sum_{k=1}^{K} [d_k-\gamma_k-\beta_k\} \qquad \text{III.4.26}$$

where $[d_k-\gamma_k-\beta_k\} \triangleq \begin{cases} d_k-\gamma_k-\beta_k & \text{if } d_k-\gamma_k-\beta_k > 0 \\ 0 & \text{otherwise.} \end{cases}$

<u>COR.III.4.3</u>: In the most general case, let for each k, $k = 1,2,\ldots,K$, d'_k be the size of $J_j^{(k)}$ such that $C_{k,j}J_j^{(k)}B_{k,j}$ is non-zero and $C_{k,j}e^{J_j(i)t}B_{k,j}$ gives rise to the highest power of t. If the numbers $\dot{\gamma}_k$ and β_k denote the number of first consecutive and last consecutive zeros in $C_{k,j}$ and $B_{k,j}$, as in Cor.III.4.2 then:

$$\dim \Sigma_{LR}(t_0) = \sum_{k=1}^{K} [d'_k-\gamma_k-\beta_k\} \qquad \text{III.4.27}$$

CHAPTER IV

SOME CANONICAL FORMS AND PROPERTIES OF THE STATE DESCRIPTIONS FOR LINEAR, TIME INVARIANT, CONTINUOUS OBJECTS

IV.1--Introduction

In the previous chapters we have only dealt with the gross properties of the state description, without trying to generate any analytic description of the I-O-S-R, except maybe in section III.4. So, Chapter IV gives us some analytical forms for the I-O-S-R and a good knowledge about the interesting properties of the state space, when, as the title indicates, the object under consideration is a linear, time-invariant and continuous one.

Of the two strategy procedures available to reach the goal, the less mathematically sophisticated and more engineering approach, of first guessing what the I-O-S-R and the state space might be and then showing that they satisfy the axioms, is chosen, rather than building up to the result by using the state axioms and mathematical tools as does Balakrishnan in [BA 1-4]. However, we would also like to point out that by proceeding as such, it should not be understood that we are being mathematically irrigorous. Our main tool in these investigations is the theory of distributions and their orthogonal series expansion as developed by Zemanian in [ZE2], a brief exposé of which is given in Appendix A.

The next section of the present chapter tries to justify the use of convolutional objects as our starting point by means of arguments that stem from the references [ZE1, 4].

In section 3, we give an infinite but countable state description of a large class of convolutional objects, namely the ones with an impulse response which has an infinite series representation. Then we investigate and prove some very important properties of the I-O-S-R and the state space such as: "The infinite A-matrix associated with the I-O-S-R is a Hilbert Matrix," "The state space $\Sigma(t)$ is a closed linear subspace of the Hilbert Space ℓ^2," etc.

The last section deals with the most general convolutional objects and shows that it is possible to approximate any such object with objects that have a finite dimensional reduced state space. This, as noted in [ZA1], happens to be a very important problem in that it may provide us with some tools of approximating a large class of distributed systems with passive, lumped RLC networks.

IV.2--Convolution Representation of Linear, Time Invariant and Continuous Objects

NOTE IV.2.1: In Fact I.2.1 we noted that a uniform object was completely defined when the input-output list $R_{\hat{I}}$ over the existence interval $\hat{I}$ was known. Furthermore, Thm.I.3.2 stated that the object $\mathcal{O}$ had a state description iff it was causal. We thus have a single valued mapping from the input space $U_{\hat{I}}$ into the output space $Y_{\hat{I}}$ due to causality.

Taking $\hat{I} = (-\infty,\infty)$, if we restrict our attention, for the moment, to objects with inputs in the space $\mathcal{D}$ of testing functions and outputs from the space $\mathcal{D}'$ of distributions over $\mathcal{D}$, we then have a single valued mapping from $\mathcal{D}$ into $\mathcal{D}'$ (for the definition of $\mathcal{D}$, $\mathcal{D}'$ and notions related to distributions see Appendix A). Moreover we have a linear, time invariant mapping from $\mathcal{D}$ into $\mathcal{D}'$ if we let our object be linear and time invariant. To these properties of single valuedness, linearity and time invariance possessed by many systems we will add one more property, "continuity," which is more difficult to interpret physically and which can crudely be described by: "in the input-output list $R_{\hat{I}}$ of the object $\mathcal{O}$, to two different inputs that are almost the same, correspond two outputs that are almost the same." A precise definition of "almost the same" would require a discussion of the neighborhood concept in $\mathcal{D}$ and $\mathcal{D}'$, and that would lead us to Topological Vector Spaces (see, e.g., [TR], [HOR]). To avoid that, and as we already have a concept of convergence in $\mathcal{D}'$, we define "continuity" as follows (a slightly modified version of the definition in [ZE4]).

DEF.IV.2.1:

An object is said to be CONTINUOUS (or a CONTINUOUS MAPPING FROM $\mathcal{D}$ INTO $\mathcal{D}'$) iff the convergence of $\{u_n\}_{n=1}^{\infty}$ to u, in $\mathcal{D} \Longrightarrow$ the convergence of $\{y_n\}_{n=1}^{\infty}$ to y, in $\mathcal{D}'$ with $(u,y) \in R_{\hat{I}}$.

THM.IV.2.1: SCHWARTZ'S KERNEL THEOREM [TR]. The mapping from $\mathcal{D}$ into $\mathcal{D}'$, that the object $\mathcal{O}$ given by $R_{\hat{I}}$ describes, is single valued, linear and continuous iff there exists a unique $w(t,\tau) \in \mathcal{D}'$ defined on the real plane such that $(u,y) \in R_{\hat{I}} \Longleftrightarrow y(t) = w(t,\tau) x u(\tau)$ $\forall u \in \mathcal{D}$, where $w(t,\tau) x u(\tau) \in \mathcal{D}'$ is defined by $\langle w(t,\tau) x u(\tau), \phi(t)\rangle \triangleq \langle w(t,\tau), u(\tau)\phi(t)\rangle$ $\forall \phi \in \mathcal{D}$.

THM.IV.2.2: [SC1, vol.II, pp. 53-54] The object satisfies the hypothesis of Thm.IV.2.1 and is time invariant iff there exists a unique $w(t) \in \mathcal{D}'$ such that $(u,y) \in R_{\hat{I}} \Longleftrightarrow y(t) = w(t) * u(t)$ $\forall u \in \mathcal{D}$, where $w(t) * u(t)$ is defined in Def.A.2.7.

NOTE IV.2.2: [ZE4, p. 8] Now, because of this convolution representation, the input space $U_{\hat{I}}$ of $\mathcal{O}$ can be extended to the space $\mathcal{E}'$ of distributions with compact support, i.e., for $u \in U_{\hat{I}} = \mathcal{E}'$ $(u,y) \in R_{\hat{I}} \Longleftrightarrow y(t) = w(t)*u(t)$. As $\mathcal{D}$ is dense in $\mathcal{E}'$ this extension of the convolution representation is unique. Moreover, if w happens to be suitably restricted the input space $U_{\hat{I}}$ can be further extended to larger spaces of distributions. If for example $w \in \mathcal{D}'_R$ the space of distributions with support bounded at left then $U_{\hat{I}}$ can be taken as all of $\mathcal{D}'_R$, or for that matter any subspace of it. Also, if $w \in \mathcal{E}'$, then $U_{\hat{I}}$ can be taken to be all of $\mathcal{D}'$. In both cases, as $\mathcal{D}$ is dense in $\mathcal{D}'_R$ and $\mathcal{D}'$, the extensions are unique.

In case the object $\mathcal{O}$ is not time invariant, the same extensions can be made by using the kernel representation of $\mathcal{O}$.

<u>THM.IV.2.3</u>: [ZE4, p. 9, THM.3] Let $\mathcal{O}$ be defined by $y(t) = w(t,\tau)xu(\tau)$, where u belongs to the extended $U_{\hat{I}}$; then $\mathcal{O}$ is causal (Def.I.2.4) iff supp $y(t,\tau)$ is contained in the half plane $\{(t,\tau) : t \geqslant \tau\}$.

If in addition $\mathcal{O}$ is time invariant then $\mathcal{O}$ is causal iff supp $w(t) \subset [0,\infty)$.

<u>NOTE IV.2.3</u>: In the light of the above discussions, the following two sections concentrate on convolutional objects with supp $w(t) \subset [0,\infty)$. These will be defined more precisely at the beginning of each section.

IV.3--A Countable-Differential State Description

NOTE IV.3.1: In this section we will obtain the state description of a linear, time invariant continuous object whose impulse response w(t) is a distribution in $\mathfrak{a}'$ (Def.A.2.9). This state description can be viewed as the generalization of the familiar state equations in III.4, to include state descriptions for distributed systems. In fact they ressemble very much the form in III.4.1 except that the matrices and the vectors involved are infinite in size.

The idea in developing the state description is simple and its root lies in the fact that for a lumped network we obtain a state description via the decomposition of w(t) into different exponomial terms (see Def. IV.4.2 for a description of "exponomial term"), as we have already remarked in Note III.4.2.

First we precisely define the object, for which the description can be given then proceed to obtain the state description.

DEF.IV.3.1:

The convolutional object under consideration is called an O_D object iff its input space $U_{D\hat{I}} \triangleq$ $\{u(t)$: U(t) is real and square summable over $(-\infty,b]$ for any real finite b$\}$. We further assume that w(t)

is a real distribution in $\mathcal{O}\!\mathcal{l}'$ and that the I-O list is:
$R_{D\hat{I}} \triangleq \{(u,y) : u \in U_{D\hat{I}} \text{ and } y(t) = w(t)*u(t)\}$.

NOTE IV.3.2: The convolution of $w(t) \in \mathcal{O}\!\mathcal{l}'$ with $u(t) \in U_{D\hat{I}}$ is well defined (Thm.A.3.1). Furthermore supp $w(t) \subset [0,\infty)$ due to causality by Thm.IV.2.3. As the input space, $w(t)$ and the eigenvalues λ_n (Not.A.2.1) are real, we can take $\{\Psi_n(t)\}_{n=1}^{\infty}$ to be real.

THM.IV.3.1: Any $\mathcal{O}_D$ object can be given the following dynamic description (a conjectured I-O-S-R):

$$\frac{dx_n(t)}{dt} = \sum_{m=1}^{\infty} a_{nm} x_m(t) + b_n u(t) \quad n = 1,2,\ldots \qquad \frac{dX(t)}{dt} = AX(t) + Bu(t) \qquad \text{IV.3.1}$$

i.e.

$$y(t) = \sum_{n=1}^{\infty} c_n x_n(t) + \sum_{k=0}^{K} d_k u^{(k)}(t) \qquad y(t) = CX(t) + DU^*(t) \qquad \text{IV.3.2}$$

where in IV.3.1 the convergence is pointwise and in IV.3.2 in $\mathcal{D}'$ together with

$$X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ \vdots \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \\ \vdots & \vdots & \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \end{bmatrix}, \quad C^T = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ \vdots \end{bmatrix},$$

$$D^T = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{bmatrix} \text{ and } U^*(t) = \begin{bmatrix} u(t) \\ u^{(1)}(t) \\ \vdots \\ u^{(k)}(t) \end{bmatrix}$$

PROOF: As $w(t) \in \mathcal{O}\!\mathcal{L}'$, $w(t) = \sum_{n=1}^{\infty} \langle w, \Psi_n \rangle \Psi_n(t)$ by Thm. A.2.15 where $\{\Psi_n\}_{n=1}^{\infty}$ are as given by Not.A.2.1. Using Thm.A.3.3, $y(t) = w(t) * u(t) =$

$$\left[\sum_{n=1}^{\infty} \langle w, \Psi_n \rangle \Psi_n(t) 1(t) + \sum_{k=0}^{K} d_k \delta^{(k)}(t) \right] * u(t). \qquad \text{IV.3.3}$$

By Thm.A.3.2, the infinite summation in IV.3.3 can be taken outside and Thm.A.2.8 can be used to yield:

$$y(t) = \sum_{n=1}^{\infty} [\langle w, \Psi_n \rangle \Psi_n(t) 1(t) * u(t)] + \sum_{k=0}^{K} d_k u^{(k)}(t). \qquad \text{IV.3.4}$$

Now we define $x_n(t) \triangleq \Psi_n(t) 1(t) * u(t)$ for $n=1,2,\ldots$ IV.3.5

Differentiating IV.3.5 according to Thm.A.2.9 we obtain

$$\frac{dx_n(t)}{dt} = \left[\frac{d}{dt}(\Psi_n(t) 1(t))\right] * u(t) = \Psi'_n(t) 1(t) * u(t) + \Psi_n(t)\delta(t) * u(t)$$

$$= \Psi'_n(t) 1(t) * u(t) + \Psi_n(0) u(t). \qquad \text{IV.3.6}$$

But $\Psi'_n(t) \in \mathcal{O}\!\mathcal{L}$ (Lemma A.3.3) and can be expressed as $\Psi'_n(t) = \sum_{m=1}^{\infty} \langle \Psi'_n, \Psi_m \rangle \Psi_m(t)$, Thm.A.3.11. This time the convergence being in $\mathcal{O}\!\mathcal{L}$, and certainly in $L^2_{(-\infty,\infty)}$. So for each n we can write

$$\frac{dx_n(t)}{dt} = \left[\sum_{m=1}^{\infty} \langle \Psi'_n, \Psi_m \rangle \Psi_m(t) 1(t) \right] * u(t) + \Psi_n(0) u(t) \qquad \text{IV.3.7}$$

$$= \int_{-\infty}^{t} \sum_{m=1}^{\infty} \langle \Psi'_n, \Psi_m \rangle \Psi_m(t-\tau) u(\tau) d\tau + \Psi_n(0) u(t)$$

$$\frac{dx_n(t)}{dt} = \langle \sum_{m=1}^{\infty} \langle \Psi'_n, \Psi_m \rangle \Psi_m(t-\tau), u(\tau)/_{(-\infty,t]}0^0 \rangle + \Psi_n(0)u(t)$$

for each t. IV.3.8

Now as convergence in L^2 defined by the norm is continuous with respect to the inner product, i.e., as strong convergence implies weak convergence [RI, p. 69], the infinite summation can be taken outside the inner product in IV.3.8 and

$$\frac{dx_n(t)}{dt} = \sum_{m=1}^{\infty} \langle \Psi'_n, \Psi_m \rangle \langle \Psi_m(t-\tau), u(\tau)/_{(-\beta,t]}0^0 \rangle + \Psi_n(0)u(t)$$

for each t.

$$= \sum_{m=1}^{\infty} \langle \Psi'_n, \Psi_m \rangle \int_{-\infty}^{t} \Psi_m(t-\tau)u(\tau)d\tau + \Psi_n(0)u(t)$$

$$= \sum_{m=1}^{\infty} \langle \Psi'_n, \Psi_m \rangle [\Psi_m(t)1(t) * u(t)] + \Psi_n(0)u(t)$$

for each t, IV.3.9

can be obtained. Using IV.3.5 in IV.3.4 and IV.3.9 and defining the coefficients

$$a_{nm} \triangleq \langle \Psi'_n, \Psi_m \rangle, \; b_n \triangleq \Psi_n(0) \text{ and } c_n \triangleq \langle w, \Psi_n \rangle \qquad \text{IV.3.10}$$

$$\frac{dx_n(t)}{dt} = \sum_{m=1}^{\infty} \langle \Psi'_n, \Psi_m \rangle x_m(t) + \Psi_n(0)u(t)$$

$$= \sum_{m=1}^{\infty} a_{nm} x_m(t) + b_n u(t) \qquad \text{IV.3.1}$$

$$y(t) = \sum_{n=1}^{\infty} \langle w,\Psi_n\rangle x_n(t) + \sum_{k=0}^{K} d_k u^{(k)}(t)$$

$$= \sum_{n=1}^{\infty} c_n x_n(t) + \sum_{k=0}^{K} d_k u^{(k)}(t). \qquad \text{IV.3.2}$$

That the convergence, in IV.3.1 is pointwise is clear since IV.3.9 converges for each t and that in IV.3.2 it is in $\mathcal{D}'$ is clear by IV.3.4, where the convergence is in $\mathcal{D}'$.

NOTE IV.3.3: The following theorem obtains a certain half reduced partitioning (Def.II.4.1) that is compatible with the dynamic description of Thm.IV.3.1. The theorem after that using this half reduced partitioning and Thm.II.4.4 shows that the expressions IV.3.1 and IV.3.2 provide us with a half reduced state description.

THM.IV.3.2: The family $\mathcal{H}'_{t_0} \triangleq \{H'_{t_0}[u_0] : u_0 \in U_{D(-\infty,t_0]}\}$ of classes of inputs where for any $t_0 \in (-\infty,\infty)$ $H'_{t_0}[u_0] \triangleq$

$$\{u'_0 \in U_{D(-\infty,t_0]} : \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u'_0(\tau)d\tau = \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u_0(\tau)d\tau \text{ for } u_0 \in U_{D(-\infty,t_0]} \text{ and } n = 1,2,\ldots\} \qquad \text{IV.3.11}$$

is a half reduced partitioning.

PROOF: We verify (Def.II.4.1):

(i) $H'_{t_0}[u_0] \neq \phi$, if nothing else one such class may contain only the defining input.

(ii) Let u'_0 and $u''_0 \in U_{D(-\infty,t_0]}$ and $H'_{t_0}[u'_0] \cap H'_{t_0}[u''_0) \neq \phi$. Then $\exists \hat{u}_0 \in U_{D(-\infty,t_0]} \ni \hat{u}_0 \in H'_{t_0}[u'_0] \cap H'_{t_0}[u''_0]$.

For any $u_0 \in H'_{t_0}[u'_0]$ we have:

$$\int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u_0(\tau)d\tau = \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u'_0(\tau)d\tau =$$

$$\int_{-\infty}^{t_0} \Psi_n(t_0-\tau)\hat{u}_0(\tau)d\tau = \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u''_0(\tau)d\tau$$

$n = 1,2,\ldots$

Thus $u_0 \in H'_{t_0}[u''_0]$ and $H'_{t_0}[u'_0] \subset H'_{t_0}[u''_0]$. Similarly $H'_{t_0}[u''_0] \subset H'_{t_0}[u'_0]$ can be shown, giving $H'_{t_0}[u'_0] = H'_{t_0}[u''_0]$.

(iii) Let $u^1_0 \in H_{t_0}[u^2_0]$. We have to show $u^1_0 \simeq u^2_0$, i.e., as all concatenations are allowed all we need to prove is: for any $u \in U_{D(t_0,\infty)}$, $t_0 < \infty$ such that $(u^1_0 0 u, y^1)$ and $(u^2_0 0 u, y^2) \in R_{D(-\infty,\infty)}$ we must have $y^1/_{(t_0,\infty)} = y^2/_{(t_0,\infty)}$. Going back to expression IV.3.4

$$\begin{aligned} y^i(t) &= w(t) * (u^i_0 0 u)(t) \\ &= \sum_{n=1}^{\infty} [\langle w,\Psi_n\rangle \Psi_n(t) 1(t) * (u^i_0 0 u)(t)] + \\ &\quad \sum_{k=0}^{K} d_k (u^i_0 0 u)^{(k)}(t) \\ &= \sum_{n=1}^{\infty} \langle w,\Psi_n\rangle [\int_{-\infty}^{t} \Psi_n(t-\tau)(u^i_0 0 u)(\tau)d\tau + \\ &\quad \sum_{k=0}^{K} d_k (u^i_0 0 u)^{(k)}(t) \quad \text{for } i = 1,2. \end{aligned} \qquad \text{IV.3.12}$$

$$y^i(t)/_{(t_0,\infty)} = \sum_{n=1}^{\infty} \langle w,\Psi_n\rangle \Big[\int_{-\infty}^{t_0} \Psi_n(t-\tau)u^i{}_0(\tau)d\tau + \int_{t_0}^{t} \Psi_n(t-\tau)u(\tau)d\tau \Big] + \sum_{k=0}^{K} d_k u^{(k)}(t) \text{ for } t > t_0 \qquad \text{IV.3.13}$$

As $\Psi_n(t-\tau)$ is in $L^2{}_{(-\infty,\infty)}$ and as $\Psi_n(t_0-\tau)$ forms a complete orthonormal basis for $L^2{}_{(-\infty,\infty)}$ we can use Fact A.3.2 in IV.3.13 to obtain:

$$y^i(t)/_{(t_0,\infty)} =$$

$$\sum_{n=1}^{\infty} \langle w,\Psi_n\rangle \Big[\int_{-\infty}^{t_0} \sum_{m=1}^{\infty} \langle \Psi_n(t-\tau),\Psi_m(t_0-\tau)\rangle \Psi_m(t_0-\tau)u^i{}_0(\tau)d\tau + \int_{t_0}^{t} \Psi_n(t-\tau)u(\tau)d\tau \Big] + \sum_{k=0}^{K} d_k u^{(k)}(t)$$

$$\text{for } i=1,2 \text{ and } t > t_0 \qquad \text{IV.3.14}$$

Using once more the continuity of inner product with respect to convergence in $L^2{}_{(-\infty,\infty)}$

$$\int_{-\infty}^{t_0} \sum_{m=1}^{\infty} \langle \Psi_n(t-\tau),\ \Psi_m(t_0-\tau)\rangle \Psi_m(t_0-\tau)u^i{}_0(\tau)d\tau =$$

$$= \Big\langle \sum_{m=1}^{\infty} \langle \Psi_n(t-\tau),\ \Psi_m(t_0-\tau)\rangle \Psi_m(t_0-\tau),\ (u^i{}_0 0)(\tau)\Big\rangle$$

$$= \sum_{m=1}^{\infty} <\Psi_n(t-\tau),\Psi_m(t_0-\tau)><\Psi_m(t_0-\tau),(u^1{}_0 0 0)(\tau)>$$

$$= \sum_{m=1}^{\infty} <\Psi_n(t-\tau),\Psi_m(t_0-\tau)> \int_{-\infty}^{t_0} \Psi_m(t_0-\tau)u^1{}_0(\tau)d\tau$$

IV.3.15

Using IV.3.14 in IV.3.14 we finally obtain:

$$y^1/_{(t_0,\infty)} = \sum_{n=1}^{\infty} <w,\Psi_n>[\sum_{m=1}^{\infty} <\Psi_n(t-\tau),$$
$$\Psi_m(t_0-\tau)> \int_{-\infty}^{t_0} \Psi_m(t_0-\tau)u^1{}_0(\tau)d\tau +$$
$$\int_{t_0}^{t} \Psi_n(t-\tau)u(\tau)d\tau] + \sum_{k=0}^{K} d_k u^{(k)}(t) \quad \text{IV.3.16}$$

As $\int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u^1{}_0(\tau)d\tau = \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u^2{}_0(\tau)d\tau$

for n=1,2,... by IV.3.11, it follows that

$y^1{}_0/_{(t_0,\infty)} = y^2{}_0/_{(t_0,\infty)}$.

<u>DEF.IV.3.2</u>:

We define the set $\Sigma_D(t)$ (a conjectured state space) by:

$\Sigma_D(t_0) \triangleq \{X(t_0) : X(t_0) = (x_1(t_0),x_2(t_0),\ldots)$ where

$x_n(t_0) = \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u_0(\tau)d\tau$ for each n and for some

$u_0 \in U_{D(-\infty,t_0]}\}$ IV.3.16

The mapping $(C'_{t_0})^{-1} : \mathcal{H}'_{t_0} \to \Sigma_D(t_0)$ is defined as follows, for any class $H'_{t_0}[u_0] \in \mathcal{H}'_{t_0}$: $(C'_{t_0})^{-1}(H'_{t_0}[u_0]) = X(t_0)$ where the defining input u'_0 for $X(t_0)$ is any $u'_0 \in H'_{t_0}[u_0]$.

DEF.IV.3.3:

We define the INFINITE TRANSITION MATRIX by $\Phi(t,t_0) \triangleq [\langle\Psi_n(t-\tau),\Psi_m(t_0-\tau)\rangle]_{nm}$ and the conjectured I-O-S-R, where $\sigma_0 \triangleq X(t_0) \in \Sigma'(t_0)$ and C is the infinite vector whose components are defined by the expression IV.3.10, with:

$$\bar{A}_{D(t_0,\infty)}(\sigma_0,u) \triangleq C\Phi(t,t_0)X(t_0) + w(t)*u(t) \text{ on } (t_0,\infty). \tag{IV.3.17}$$

THM.IV.3.3: $(\Sigma_D,\bar{A}_D)$ as given by Def.IV.3.2 and IV.3.3 is a half reduced state description of the object $\mathcal{O}_D$.

PROOF: We verify Def.II.4.1 where the half reduced partitioning is that of Thm.IV.3.2.

(i) For $\Sigma_D(t_0)$ to be a half reduced state description all we have to show is that $(C'_{t_0})^{-1}$ (Def.IV.3.2) is one to one and onto.

It is onto: by definition any $X(t_0) \in \Sigma(t_0)$ is such that there exists a $u_0 \in U_{D(-\infty,t_0]}$ for which $x_n(t_0) = \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u_0(\tau)d\tau$ for $n=1,2,\ldots$ thus $X(t_0) = (C'_{t_0})^{-1}(H'_{t_0}[u_0])$.

It is one to one: let u'_0 and $u''_0 \in U_{D(-\infty,t_0]}$ be such that $H'_{t_0}[u'_0] \neq H'_{t_0}[u''_0]$. By definition of $H'_{t_0}[u_0]$ this means that there exists n_1 such that: $\int_{-\infty}^{t_0} \Psi_{n_1}(t_0-\tau)u'_0(\tau)d\tau \neq \int_{-\infty}^{t_0} \Psi_{n_1}(t_0-\tau)u''_0(\tau)d\tau$, which then implies: $X'(t_0) = (C'_{t_0})^{-1}(H'_{t_0}[u'_0]) \neq (C'_{t_0})^{-1}(H'_{t_0}[u''_0]) = X''(t_0)$. With the above proof the use of the inverse notation for C'_{t_0} is also justified.

(ii) That $\bar{A}_{D(t_0,\infty)}$ as in Def.IV.2.3 is the I-O-S-R can easily be shown. In the previous theorem, expression IV.3.16 gave us the (t_0,∞) portion of the response to $u_0 0 u$.

$$y/_{(t_0,\infty)} = \sum_{n=1}^{\infty} \langle w,\Psi_n\rangle[\sum_{m=1}^{\infty} \langle\Psi_n(t-\tau), \Psi_m(t_0-\tau)\rangle \int_{-\infty}^{t_0} \Psi_m(t_0-\tau)u_0(\tau)d\tau + \int_{t_0}^{t} \Psi_n(t-\tau)u(\tau)d\tau] + \sum_{k=0}^{K} d_k u^{(k)}(t),\ t > t_0$$

$$= \sum_{n=1}^{\infty} \langle w,\Psi_n\rangle[\sum_{m=1}^{\infty} \langle\Psi_n(t-\tau), \Psi_m(t_0-\tau)\rangle x_m(t_0)] + \sum_{n=1}^{\infty} \langle w,\Psi_1\rangle[\Psi_n(t)1(t)*u(t)] + \sum_{k=1}^{K} d_k \delta^{(k)}(t)*u(t)$$

$$= \sum_{n=1}^{\infty} c_n [\sum_{m=1}^{\infty} <\Psi_n(t-\tau),\Psi_m(t_0-\tau)> x_m(t_0)] +$$

$$[\sum_{n=1}^{\infty} <w,\Psi_n>\Psi_n(t)1(t) + \sum_{k=0}^{K} d_k \delta^{(k)}(t)]*u(t)$$

can be achieved using the definitions of $x_m(t_0)$ and c_n, and Thm.A.3.2. Finally using the infinite matrix notation and Thm.A.3.3:

$$y/_{(t_0,\infty)} = C\Phi(t,t_0)X(t_0) + w(t)*u(t) =$$

$$\bar{A}_{D(t_0,\infty)}(\sigma_0,u). \qquad \text{IV.3.17}$$

Clearly for any $u_0 \epsilon C_{t_0}(\sigma_0)/_{\sigma_0=X(t_0)}$, $u_0 0 u$ is admissible and $\bar{A}_{D(t_0,\infty)}$ is the (t_0,∞) portion of the response to $u_0 0 u$, thus making IV.3.17 an I-O-S-R by Def.II.4.4 and $(\Sigma_D,\bar{A}_D)$ a half reduced state description by Thm.II.4.4.

NOTE IV.3.4: Thm.IV.3.2 and Thm.IV.3.3 have shown that $(\Sigma_D,\bar{A}_D)$ constitutes a state description and in their light, the dynamic equations IV.3.1 and IV.3.2 can be viewed as the state equations of the object $\mathcal{O}_D$. To make the tie between the state description and the state equations stronger and to justify the name "Infinite Transition Matrix" for $\Phi(t,t_0)$, we prove the following two theorems which also improve the I-O-S-R, IV.3.17.

THM.IV.3.4: The infinite transition matrix $\Phi(t,t_0)$, Def.IV.3.3, is the FUNDAMENTAL MATRIX of the infinite differential equation system $\frac{dX(t)}{dt} = AX(t)$, with $\Phi(t_0,t_0) = I$.

PROOF: $\Phi(t_0,t_0) \triangleq [<\Psi_n(t-\tau),\Psi_m(t_0-\tau)>]_{mn} = [\partial_{mn}]$, where ∂_{mn} is the Kroenecker delta, since $\{\Psi_n(t_0-\tau)\}_{n=1}^{\infty}$ forms a complete orthonormal system (Fact A.3.2). To show $\Phi(t,t_0)$ is a fundamental matrix:

(i) we first show that every column of $\Phi(t,t_0)$ is a vector solution of the infinite differential system. To pick a column of $\Phi(t,t_0) = [<\Psi_n(t-\tau),\Psi_m(t_0-\tau)>]_{mn}$ we fix the column index m at an arbitrary m_0; then we substitute the vector so obtained by X(t) in $\frac{dX}{dt} = AX(t)$ to get:

$$\frac{d}{dt}<\Psi_n(t-\tau),\Psi_{m_0}(t_0-\tau)> \equiv \sum_{m=1}^{\infty} <\Psi'_n,\Psi_m><\Psi_m(t-\tau),\Psi_{m_0}(t_0-\tau)> \tag{IV.3.18}$$

using IV.3.10, definition of $\bar{A}$. Now we have to verify the identity IV.3.18. In fact:

$$\frac{d}{dt}<\Psi_n(t-\tau),\Psi_{m_0}(t_0-\tau)> \equiv \frac{d}{dt} \int_{-\infty}^{\infty} \Psi_n(t-\tau)\Psi_{m_0}(t_0-\tau)d\tau$$

$$\equiv \int_{-\infty}^{\infty} \Psi'_n(t-\tau)\Psi_{m_0}(t_0-\tau)d\tau$$

since $\Psi_n(\cdot)$ is infinitely smooth.

$$\equiv <\Psi'_n(t-\tau),\Psi_{m_0}(t_0-\tau)> \equiv$$

$$< \sum_{m=1}^{\infty} <\Psi'_n(t-\tau),\Psi_m(t-\tau)>\Psi_m(t-\tau),\Psi_{m_0}(t_0-\tau)> \tag{IV.4.19}$$

Using the continuity of the inner product and noting that: $<\Psi'_n(t-\tau),\Psi_m(t-\tau)> =$

$\int_{-\infty}^{\infty}\Psi'_n(t-\tau)\Psi_m(t-\tau)d\tau = \int_{-\infty}^{\infty}\Psi'_n(\tau)\Psi_m(\tau)d\tau =$

$<\Psi'_n,\Psi_m>$, IV.3.19 becomes $\frac{d}{dt}<\Psi_n(t-\tau),\Psi_{m_0}(t_0-\tau)> \equiv$

$\sum_{m=1}^{\infty}<\Psi'_n,\Psi_m><\Psi_m(t-\tau),\Psi_{m_0}(t_0-\tau)>$ verifying the identity IV.3.18.

(ii) We now show that the columns of $\Phi(t,t_0)$ are linearly independent for all t. To do this select any k (k also arbitrary) columns of $\Phi(t,t_0)$ and suppose for some t, there exists scalars $\alpha_1,\alpha_2,\ldots,\alpha_k$ such that not all α_k are zero and $\alpha_1<\Psi_i(t-\tau),\Psi_{n_1}(t_0-\tau)> +$ $\alpha_2<\Psi_i(t-\tau),\Psi_{n_2}(t_0-\tau)> + \ldots + \alpha_k<\Psi_i(t-\tau),$ $\Psi_{n_k}(t_0-\tau)> = 0$ or

$$<\Psi_i(t-\tau),\sum_{j=1}^{k}\alpha_j\Psi_{n_j}(t_0-\tau)> = 0 \quad \text{for } i=1,2,\ldots \qquad \text{IV.3.20}$$

As $\{\Psi_i(t-\tau)\}_{i=1}^{\infty}$ is a complete orthonomal basis, IV.3.20 gives $\sum_{j=1}^{k}\alpha_j\Psi_{n_j}(t_0-\tau) = 0$; but this cannot be true since $\{\Psi_j(t_0-\tau)\}_{j=1}^{\infty}$ also forms a complete orthonormal basis. This contradiction implies the columns of $\Phi(t,t_0)$ are linearly independent for any t.

<u>THM.IV.3.5</u>: The I-O-S-R, $\bar{A}_D$ of the object Θ_D is also given by:

$$y/_{(t_0,\infty)} = C\Phi(t,t_0)X(t_0) + C\int_{t_0}^{t} \Phi(t,t_0)Bu(\tau)d\tau + DU^*(t) \qquad \text{IV.3.21}$$

<u>PROOF</u>: Combining IV.3.16 with IV.3.17 we can write:

$$y/_{(t_0,\infty)} = C\Phi(t,t_0)X(t_0) + \sum_{n=1}^{\infty} \langle w,\Psi_n\rangle \int_{t_0}^{t} \Psi_n(t-\tau)u(\tau)d\tau + \sum_{k=0}^{K} d_k u^{(k)}(t) \qquad \text{IV.3.22}$$

For fixed n consider $\Psi_n(t-\tau+z)$ which can be written for each t,τ as: $\Psi_n(t-\tau+z) = \sum_{m=1}^{\infty} \langle\Psi_n(t-\tau+s),\Psi_m(s)\rangle\Psi_m(z)$ where s is the dummy variable, i.e.,

$$\Psi_n(t-\tau+z) = \sum_{m=1}^{\infty} [\int_{-\infty}^{\infty} \Psi_n(t-\tau+s)\Psi_m(s)ds]\Psi_m(z) \qquad \text{IV.3.23}$$

by Thm.A.2.12, $\Psi_n(t-\tau+z)$ and the convergence being in $\mathfrak{A}$. Then by Cor.A.2.1 it follows that the convergence is uniform or compact subsets of $(-\infty,\infty)$; therefore we can evaluate IV.3.23 at $z = 0$. Also doing the change $-x = -\tau+s$ of variables

$$\Psi_n(t-\tau) = \sum_{m=1}^{\infty} [\int_{-\infty}^{\infty} \Psi_n(t-x)\Psi_m(\tau-x)dx]\Psi_m(0)$$

$$= \sum_{m=1}^{\infty} \langle\Psi_n(t-x),\Psi_m(\tau-x)\rangle\Psi_m(0) \qquad \text{IV.3.24}$$

follows from IV.3.23. Substituting IV.3.24 in IV.3.22:

$$y/_{(t_0,\infty)} = C\Phi(t,t_0)X(t_0) + \sum_{n=1}^{\infty} \langle w,\Psi_n\rangle \int_{t_0}^{t} \sum_{m=1}^{\infty} \langle \Psi_n(t-x),$$

$$\Psi_m(\tau-x)\rangle \Psi_m(0)u(\tau)d\tau + \sum_{k=0}^{K} d_k u^{(k)}(t)$$

and using IV.3.10 and Def.IV.3.3

$$y/_{(t_0,\infty)} = C\Phi(t,t_0)X(t_0) + C\int_{t_0}^{t} \Phi(t,t_0)Bu(\tau)d\tau + DU^*(t) \qquad \text{IV.3.21}$$

is finally obtained.

NOTE IV.3.5: We, thus, have shown the strong resemblance between the exponential transition matrix $e^{A(t-t_0)}$ for a square A of finite size and our transition matrix $\Phi(t,t_0)$. Now we will investigate the nature of our state space and an essential property of our infinite A matrix that may bear strong relation to the stability of the object $\mathcal{O}_D$ under consideration. The following theorem is the main reason for all the work we had to go through in A.3 when defining the convolution of a distribution in $\mathfrak{A}'$, with an input from U_D; it makes it possible to show that $\Sigma(t_0)$ is a closed linear subspace of the classical Hilbert space ℓ^2, which could not be proved if U_D was not taken as the space of square summable functions over $(-\infty,b)$ for any finite b.

THM.IV.3.6: For each $t_0 \in (-\infty,\infty)$, the state space $\Sigma_D(t_0)$ of the object $\mathcal{O}_D$ is a closed linear subspace of the classical Hilbert space ℓ^2.

PROOF: $\Sigma_D(t_0) \subset \ell^2 \triangleq \{\{a_i\}_{i=1}^{\infty} : \sum_{i=1}^{\infty} |a_i|^2 < \infty\}$. To prove this let $X(t_0) \in \Sigma'(t_0)$, then by Def.IV.3.2 $x_n(t_0) = \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u_0(\tau)d\tau$ for some $u_0 \in U_{(-\infty,t_0]}$, $n=1,2,\ldots$ where $x_n(t_0)$ is a component of $X(t_0)$. Thus: $x_n(t_0) = \langle\Psi_n(t_0-\tau),(u_0{}_0 0)(\tau)\rangle$ since $u_0{}_0 0 \in L^2_{(-\infty,\infty)}$. Thus $x_n(t_0)$ is the Fourier coefficient of $u_0{}_0 0$ for each n with respect to the complete orthonormal basis $\{\Psi_n(t_0-\tau)\}_{n=1}^{\infty}$. Invoking a classical theorem (see e.g., [PO, p. 36]) we can write: $\sum_{n=1}^{\infty} |\langle\Psi_n(t_0-\tau),(u_0{}_0 0)(\tau)\rangle|^2 = \sum_{n=1}^{\infty} |x_n(t_0)|^2 < ||u_0{}_0 0||^2 < \infty$ since $u_0{}_0 0 \in L^2_{(-\infty,\infty)}$ making $X(t_0)$ a square summable sequence and proving $\Sigma_D(t_0) \in \ell^2$.

$\Sigma'(t_0)$ is a linear subspace. Let $X^1(t_0)$ and $X^2(t_0) \in \Sigma_D(t_0)$ and a, a scalar. Then there exists $u^1{}_0$ and $u^2{}_0 \in U_{(-\infty,t_0]}$ such that: $ax^1{}_n(t_0) + x^2{}_n(t_0) = a\int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u^1{}_0(\tau)d\tau + \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u^2{}_0(\tau)d\tau = \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)[au^1{}_0(\tau) + u^2{}_0(\tau)]d\tau$. Since the input space is linear, $au^1{}_0 + u^2{}_0 \in U_{(-\infty,t_0]}$ (Fact III.2.3) which gives that $aX^1(t_0) + X^2(t_0) \in \Sigma_D(t_0)$.

$\Sigma'(t_0)$ is closed. Let $\{X^i(t_0)\}_{i=1}^{\infty} = \{(x_1{}^i(t_0), x_2{}^i(t_0), \ldots.)\}$ be a sequence in $\Sigma_D(t_0)$ and let $\lim_{i\to\infty} X^i(t_0) = X$ in ℓ^2. We want to show that X is a state, i.e., that there exists $u \in U_{(-\infty,t_0]}$ such that $x_n = \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u(\tau)d\tau$ for n=1,2,... where x_n is the n^{th} component of the infinite limit vector X.

$$\lim_{i\to\infty} X^i(t_0) = X(t_0) \Longleftrightarrow \lim_{i\to\infty} ||X(t_0) - X^i(t_0)|| = 0$$

by definition of convergence in ℓ^2.

$$\Longleftrightarrow \lim_{i\to\infty} \sum_{n=1}^{\infty} |x_n(t_0) - x^i{}_n(t_0)| = 0$$

by definition of the norm in ℓ^2.

IV.3.24

For each i, as $X^i(t_0) \in \Sigma_D(t_0)$, there exists an input $u_i \in U_{(-\infty,t_0]}$ such that $x^i{}_n(t_0) = \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u_i(\tau)d\tau = \int_{-\infty}^{\infty} \Psi_n(t_0-\tau)(u_i 0 0)(\tau)d\tau$ for n=1,2,... Since ℓ^2 is isomorphic to $L^2_{(-\infty,\infty)}$, given any element of ℓ^2, e.g., $X(t_0)$, there is a corresponding element, $u \in L^2_{(-\infty,\infty)}$, such that [BE]: $x_n(t_0) = \int_{-\infty}^{\infty} \Psi_n(t_0-\tau)u(\tau)d\tau$ for n=1,2,... Then:

$$\lim_{i\to\infty} \sum_{n=1}^{\infty} |x_n(t_0) - x^i{}_n(t_0)|^2 = \lim_{i\to\infty} \sum_{n=1}^{\infty} \left|\int_{-\infty}^{\infty} \Psi_n(t_0-\tau)[u(\tau) - (u_i 0 0)(\tau)]d\tau\right|^2$$

$$= \lim_{i\to\infty} \sum_{n=1}^{\infty} |\langle \Psi_n(t_0-\tau), u(\tau) - (u_i 0 0)(\tau)\rangle|^2$$

$$= \lim_{i\to\infty} ||u(\tau) - (u_i 0 0)(\tau)||^2$$

by Parseval's equality.

$$= \lim_{i\to\infty} \int_{-\infty}^{\infty} |u(\tau) - (u_i 0 0)(\tau)|^2 d\tau$$

by definition of $L^2_{(-\infty,\infty)}$ norm.

$$= \lim_{i\to\infty} \int_{-\infty}^{t_0} |u(\tau) - u_i(\tau)|^2 d\tau + \int_{t_0}^{\infty} |u(\tau)|^2 d\tau = 0$$

Therefore $u(t)$ must equal zero i.e., for $t > t_0$, so that the convergence in IV.3.24 holds. Thus, the function $\tilde{u}(t) = u(t)/_{(-\infty,t_0]}$ will certainly be in $U_{(-\infty,t_0]}$ and will be such that $x_n(t_0) = \int_{-\infty}^{t_0} \Psi_n(t_0-\tau)u(\tau)d\tau$, $n=1,2,\ldots$ proving that $\Sigma_D(t_0)$ is closed for any $t_0 \in (-\infty,\infty)$.

NOTE IV.3.6: In concluding this section, our aim now is to show that the infinite matrix A in the state equations IV.3.1 is a Hilbert matrix (Def.B.3). As our infinite state vectors are from ℓ^2, that makes A a bounded operator mapping ℓ^2 into ℓ^2 (Note B.2), thus enabling us in the future to investigate about the spectrum of A and its other properties and carry out some important analysis of the object Θ_D such as its stability.

However we could only prove that A is a Hilbert matrix, under an assumption for the eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ of the operator $\mathcal{N}$, in Not.A.2.1 used to generate the testing function space Θ (Def.A.2.8) and their dual space $\mathcal{O}'$ of distributions, to which w belongs. The eigenvalues

$\{\lambda_n\}_{n=1}^{\infty}$ were already real and no value of λ_n was assumed more than a finite number of times. The numbering was so chosen that $|\lambda_1|\leqslant|\lambda_2|\leqslant \ldots$. This clearly implies that $|\lambda_n|\to\infty$ as $n\to\infty$. Now, although we conjecture that Thm. IV.3.7 is true without any further assumption, we assume that there exists a finite integer p_0 such that $\sum_{\substack{n=1\\ \lambda_n\neq 0}}^{\infty}|\lambda_n|^{-p_0}<\infty$ converges. This assumption is not too restrictive, since the eigenvalues of many $\mathcal{N}$ operators seem to possess this property (see [ZE2] for some examples).

<u>THM.IV.3.7</u>: In the dynamic equations IV.3.1 and IV.3.2 of $\mathcal{O}_D$, the infinite matrix A is a Hilbert matrix and the coefficients c_n are such that there exists an integer $q_0 \geqslant 0$ for which $\sum_{\substack{n=1\\ \lambda_n\neq 0}}^{\infty} |\lambda_n|^{-2q_0}|c_n|^2<\infty$.

<u>PROOF</u>: That the c_n's are such is given by Thm.A.2.16 since $c_n \triangleq \langle w,\Psi_n\rangle$, $n=1,2,\ldots$ To show that A is a Hilbert matrix we proceed in five steps:

(i) $\langle\Psi'_n,\Psi_m\rangle = \int_{-\infty}^{\infty}\Psi'_n(t)\Psi_m(t)d \;= \Psi_n(t)\Psi_m(t)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\Psi_n(t)\Psi'_m(t)dt = -\langle\Psi_n,\Psi'_m\rangle$ since $\Psi_n(\pm\infty) = 0$ for $n=1,2,\ldots$ by Fact A.3.3.

(ii) From (i) it follows that $\sum_{m=1}^{\infty}\langle\Psi_n,\Psi'_m\rangle\Psi_m(t)$ converges in $\mathcal{O}\!\mathcal{l}$ as well as $\sum_{m=1}^{\infty}\langle\Psi'_n,\Psi_m\rangle\Psi_m(t)$. Then using Thm.A.2.13 we have that $\sum_{m=1}^{\infty}|\lambda_m|^{2k}|a_{nm}|^2$ converges for every $k = 1,2,\ldots$ and for every $n = 1,2,\ldots$ whereas $\sum_{n=1}^{\infty}|\lambda_m|^{2k}|a_{nm}|^2$ converges for every $k = 1,2,\ldots$ and for every $m = 1,2,\ldots$, since $a_{nm} \triangleq \langle \Psi'_n, \Psi_m \rangle$.

(iii) Now we will show that $\sum_{m=1}^{\infty}|a_{nm}| \leq M$, where M is independent of n. By Note IV.3.6 there exists a finite k_0 such that $\sum_{\substack{m=1 \\ \lambda_n \neq 0}}^{\infty} |\lambda_m|^{-2k_0}$ converges. But $\sum_{m=1}^{\infty}|\lambda_m|^{4k_0}|a_{mn}$ converges for every n, since it is convergent for any integer power of the λ_m's. Thus given $\varepsilon = 1$, there exists a finite m_0 such that $\sum_{\substack{m=m_0 \\ \lambda_m \neq 0}}^{\infty}|\lambda_m|^{4k_0}|a_{mn}|^2 < 1$ as the λ_m's have finite multiplicity and no finite point of accumulation. It then follows that $|\lambda_m|^{4k}|a_{mn}|^2 \leq 1$ and that $|a_{mn}| < |\lambda_m|^{-2k_0}$ for all $m \geq m_0$. Thus, $\sum_{m=m_0}^{\infty}|a_{mn}| \leq \sum_{m=m_0}^{\infty} |\lambda_m|^{-2k_0}$ which is certainly bounded independent of n. So we have $\sum_{m=m_0}^{\infty}|a_{mn}| \leq M_1$ independent of n. Now we have to show that $\sum_{m=1}^{m_0}|a_{mn}|$ is bounded independent of n. This time looking at the series $\sum_{\substack{n=1 \\ \lambda_n \neq 0}}^{\infty}|\lambda_n|^{4k_0}|a_{mn}|^2$ which converges for every m and in particular for $m = 1,2,\ldots,m_0$ we can proceed as above to obtain that $|a_{mn}| \leqslant |\lambda_{n_0}|^{-2k_0}$ for all n greater than some n_0. As the λ_n's are so ordered that $|\lambda_{n_1}| \leqslant |\lambda_{n_2}|$ whenever $n_1 \leqslant n_2$ (Not.A.2.1) we can write $|a_{mn}| \leqslant |\lambda_{n_0}|^{-2k_0}$ for all $n \geqslant n_0$. Now: for $n > n_0$, $\sum_{m=1}^{m_0} |a_{mn}| < m_0 |\lambda_{n_0}|^{-2k_0}$, and for $n \leqslant n_0$, we have finite number of terms of the form $\sum_{m=1}^{m_0} |a_{mn}|$ for each n. So defining $M_2 \triangleq \max\{m_0 |\lambda_{n_0}|^{-2k_0}, \sum_{m=1}^{m_0} |a_{m1}|, \ldots, \sum_{m=1}^{m_0} |a_{mn_0}|\}$ we find the bound on $\sum_{m=1}^{\infty} |a_{mn}|$ independent of n. Finally $\sum_{m=1}^{\infty} |a_{mn}| < M_1 + M_2 = M$ proves (iii).

(iv) In exactly the same fashion, that $\sum_{n=1}^{\infty} |a_{mn}| < M'$ independent of m, can be shown.

(v) The hypothesis of Thm.B.2 being satisfied it follows that A is a Hilbert matrix.

IV.4--Approximation of a Large Class of Objects Having Finite Dimensional State Description

NOTE IV.4.1: In this section we are dealing with a very general class of linear, time invariant, continuous objects, i.e., with convolutional objects (Thm.IV.2.2 and Note IV.2.2), without any restriction on their impulse response w(t), but with some restrictions on their input space. The approximating objects all have reduced state descriptions of the form:

$$\frac{dX(t)}{dt} = AX(t) + Bu(t) \qquad \text{IV.4.1}$$

$$y(t) = CX(t) + \sum_{k=0}^{K} d_k u^{(k)}(t) \qquad \text{IV.4.2}$$

where all vectors and matrices are finite.

Again we start with the definition of the object under consideration and continue with the definition of the term exponomial, an expression which is the combination of "exponential" with "polynomial."

DEF.IV.4.1:

The linear, time invariant, continuous (therefore convolutional) object under consideration is called an $\mathcal{O}_G$ object iff its input space: $U_{G\hat{I}} \triangleq \{u(t) : u(t)$ is a regular distribution with support bounded from the left$\}$ and the I-O pairs are given by: $R_{G\hat{I}} \triangleq \{(u,y) : u \in U_{G\hat{I}},$ $y(t) = w(t)*u(t)$ and supp $w(t) \subset [0,\infty)\}$.

DEF.IV.4.2:

A SIMPLE EXPONOMIAL IN t is a polynomial in t multiplied by the exponential in t (e.g., EXPOL(t) = $e^{\gamma t}P(t) = a_0 e^{\gamma t} + a_1 t e^{\gamma t} + \ldots. + a_n t^n e^{\gamma t}$). An EXPONOMIAL is the sum of simple exponomials.

THM.IV.4.1: Let Ω be an open subset of $\mathbb{R}^n$. Any distribution in Ω is the limit of a sequence of exponomial functions.

PROOF: The proof is really trivial if we consider Thm.A.2.3 and that the polynomials are a subset of exponomials. Therefore exponomials are dense in $\mathcal{D}'$, in the topology of $\mathcal{D}'$, since the polynomials are dense in $\mathcal{D}'$.

NOTE IV.4.2: Due to Thm.IV.4.1, and if we take $\Omega = (-\infty,\infty)$, w(t) can be written as:

$$w(t) = \lim_{i\to\infty} \sum_{\mu=1}^{q_i} \sum_{\nu=1}^{k_\mu^i} c^i_{\mu\nu}\, t^{\nu-1} e^{\gamma_\mu^i t}, \text{ where the convergence is in } \mathcal{D}'. \qquad \text{IV.4.3}$$

Using Note A.3.5, as supp $w(t) \subset [0,\infty)$, w(t) can also be written as:

$$w(t) = \lim_{i\to\infty} \sum_{\mu=1}^{q_i} \sum_{\nu=1}^{k^i_\mu} c^i_{\mu\nu} t^{\nu-1} e^{\gamma_\mu^i t} 1(t) + \sum_{k=0}^{K} d_k \delta^{(k)}(t) \text{ in } \mathcal{D}'. \qquad \text{IV.4.4}$$

We would like to note that in IV.4.4, the form that we will be using for w(t), the integers q_i and k_μ^i are finite for each i and for each μ.

NOTE IV.4.3: It would be much nicer if w(t) was given by a summation of the form $w(t) = \sum_{\mu=1}^{\infty} \sum_{\nu=1}^{k_\mu} c_{\mu\nu} t^{\nu-1} e^{\gamma_\mu t} 1(t)$ where either k_μ is finite for each μ, in which case we have an infinite number of poles with finite multiplicity, or k_μ is allowed to be infinite for some μ, in which case we have poles of infinite multiplicity. In such a case of infinite series expansion for $w(t)$, the development is much in the same lines as the following development for IV.4.4. The question is to find out what impulse responses $w(t)$ have such a series expansion.

<u>THM.IV.4.2</u>: Let an $\mathcal{O}_G$ object be given with

$$w(t) = \lim_{i\to\infty} \sum_{\mu=1}^{q_i} \sum_{\nu=1}^{k_\mu^i} c^i_{\mu\nu} t^{\nu-1} e^{\gamma_\mu^i t} 1(t) + \sum_{k=0}^{K} d_k \delta^{(k)}(t) \qquad \text{IV.4.4}$$

Then the object $\mathcal{O}_G$ can be given the following generalized dynamic description. For each i, where i denotes a superscript and not an exponent:

$$\frac{dx^i_{\mu j}(t)}{dt} = x^i_{\mu(j+1)}(t) + \gamma_\mu^i x^i_{\mu j}(t) + b^i_{\mu j} u(t) \text{ where}$$

$$j=1,2,\ldots,k_\mu^i \text{ and } \mu=1,2,\ldots,q_i \qquad \text{IV.4.5}$$

$$y(t) = \lim_{i\to\infty} \sum_{\mu=1}^{q_i} \Big[\sum_{j=1}^{k_\mu^i} x^i_{\mu j}(t) \Big] + \sum_{k=1}^{K} d_k u^{(k)}(t). \qquad \text{IV.4.6}$$

<u>PROOF</u>: By Def.IV.4.1 of an $\mathcal{O}_G$ object and IV.4.4 the output to $u(t) \in U_{(-\infty,\infty)}$ is given by:

$$y(t) = w(t)*u(t) = \lim_{i\to\infty} \Big[\sum_{\mu=1}^{q_i} \sum_{\nu=1}^{k_\mu^i} c^i_{\mu\nu} t^{\nu-1} e^{\gamma_\mu^i t} 1(t)*u(t) \Big] +$$

$$\sum_{k=0}^{K} d_k u^{(k)}(t) \qquad \text{IV.4.7}$$

where the limit can be taken after the convolution due to Thm.A.2.10 since all distributions involved have support bounded from the left.

For each i, we define the coefficients $b^i_{\mu\nu}$ as follows:

$$b^i_{\mu\nu} \triangleq (\nu-1)!c^i_{\mu\nu} - \nu!c^i_{\mu\nu+1} \text{ where } c^i_{\mu(\nu+1)} \triangleq 0 \text{ for } \nu+1>k_\mu{}^i \qquad \text{IV.4.8}$$

with these new coefficients, remembering that i denotes a superscript not an exponent the summation over ν in IV.4.7 takes the form:

$$\sum_{\nu=1}^{k_\mu{}^i} c^i_{\mu\nu} t^{\nu-1} e^{\gamma_\mu{}^i t} 1(t) = \left[\sum_{\nu=1}^{k_\mu{}^i} \frac{b^i{}_{\mu\nu} t^{\nu-1}}{(\nu-1)!} + \sum_{\nu=2}^{k_\mu{}^i} \frac{b^i{}_{\mu\nu} t^{\nu-2}}{(\nu-2)!} + \ldots + b_{\mu_{k_\mu^i}}\right] e^{t\gamma_\mu{}^i} 1(t) \qquad \text{IV.4.9}$$

$$= \left[\sum_{j=1}^{k_\mu{}^i} \sum_{\nu=j}^{k_\mu{}^i} \frac{b^i{}_{\mu\nu} t^{\nu-j}}{(\nu-j)!}\right] e^{t\gamma_\mu{}^i} 1(t) \qquad \text{IV.4.10}$$

To verify that the right hand side of IV.4.9 with IV.4.8 gives the left hand side of IV.4.9, we note that for every coefficient of every power of t, i.e., the coefficients of $t^{\nu-1}$ from $\nu = 1$ to $\nu = k_\mu{}^i$, we have:

$$\frac{t^{j-1}}{(j-1)!}[b^i_{\mu j}+b^i_{\mu(j+1)}+\ldots+b^i_{\mu k^i_\mu}]= \frac{t^{j-1}}{(j-1)!}[(j-1)!c^i_{\mu j}- j!c^i_{\mu(j+1)}+j!c^i_{\mu(j+1)}- (j+1)!c^i_{\mu(j+2)}+\ldots+ (k^i_\mu-1)!c^i_{\mu k^i_\mu}]$$

$$= c^i_{\mu j}t^{j-1}$$

for $\nu = j$, j arbitrary, all the terms in the brackets except $(j-1)!c^i_{\mu j}$ cancelling each other. Now, for each i again, we define:

$$x^i_{\mu j}(t) \triangleq [\sum_{\nu=j}^{k_\mu{}^i} \frac{b^i_{\mu\nu}t^{\nu-j}}{(\nu-j)!} e^{t\gamma_\mu{}^i}]1(t)*u(t) \qquad \text{IV.4.11}$$

Differentiating IV.4.11, as given by Thm.A.2.9, in the distributional sense

$$\frac{dx^i_{\mu j}(t)}{dt} = \frac{d}{dt}[\sum_{\nu=j}^{k_\mu{}^i} \frac{b^i_{\mu\nu}}{(\nu-j)!}t^{\nu-j}e^{t\gamma^i{}_\mu}1(t)]*u(t) =$$

$$\sum_{\nu=j}^{k_\mu{}^i}\frac{d}{dt}[\frac{b^i_{\mu\nu}}{(\nu-j)!}t^{\nu-j}e^{t\gamma^i{}_\mu}1(t)]*u(t) \qquad \text{IV.4.12}$$

$$= \sum_{\nu=j+1}^{k_\mu{}^i}[\frac{b^i_{\mu\nu}t^{\nu-j-1}}{(\nu-j-1)!} e^{t\gamma^i{}_\mu}1(t)]*u(t) +$$

$$\gamma_\mu{}^i\sum_{\nu=j}^{k_\mu{}^i}[\frac{b^i_{\mu\nu}t^{\nu-j}}{(\nu-j)!} e^{t\gamma^i{}_\mu}1(t)]*u(t) +$$

$$\sum_{\nu=j}^{k_\mu{}^i}[\frac{b^i_{\mu\nu}t^{\nu-j}}{(\nu-j)!} e^{t\gamma^i{}_\mu}\delta(t)]*u(t) \qquad \text{IV.4.13}$$

In IV.4.13 the first and second summations are easily recognized to be $x^i_{\mu(j+1)}(t)$ and $x^i_{\mu j}(t)$ respectively due to IV.4.11. In the third summation of IV.4.13 all the terms are zero since $t^{\nu-j}e^{t\gamma^i_\mu}\delta(t) = 0$, except for $\nu = j$ in which case we have the term $b^i_{\mu j}e^{t\gamma^i_\mu}\delta(t)*u(t) = b^i_{\mu j}u(t)$. Thus IV.4.13 becomes:

$$\frac{dx^i_{\mu j}(t)}{dt} = x^i_{\mu(j+1)}(t) + \gamma^i_\mu x^i_{\mu j}(t) + b^i_{\mu j}u(t) \text{ where}$$

$$j=1,2,\ldots,k^i_\mu \text{ for } \mu=1,2,\ldots,q_i \qquad \text{IV.4.5}$$

Substituting IV.4.10 into IV.4.7 we obtain:

$$y(t) = \lim_{i\to\infty} \sum_{\mu=1}^{q_i} \left[\sum_{j=1}^{k_\mu^i} \sum_{\nu=j}^{k_\mu^i} \frac{b^i_{\mu\nu}t^{\nu-j}}{(\nu-j)!} e^{t\gamma^i_\mu}1(t)*u(t)\right] +$$

$$\sum_{k=1}^{K} d_k u^{(k)}(t) \qquad \text{IV.4.14}$$

Finally using the definition for $x^i_{\mu j}(t)$, IV.4.11 in IV.4.14 we get IV.4.6.

<u>NOTE IV.4.4</u>: The next theorem, the main result of this section, is the one advertised much earlier. It shows that any object $\mathcal{O}_\mathcal{G}$, can be approximated with objects having a finite dimensional state space. We think this result is important because we are thus given the possibility of approximating closely distributed systems, with lumped RLC networks that have a finite number of

elements. However, what subset of $\mathcal{O}_G$ objects can be approximated by such RLC networks is still an important question that remains open.

DEF.IV.4.3:

Let the objects $\mathcal{O}$ and $\mathcal{O}_i$ for i=1,2,... be given by their I-O list $R_{\hat{I}}$ and $R_{i\hat{I}}$. Then $\mathcal{O}$ is said to be the LIMIT of the objects $\mathcal{O}_i$, $\mathcal{O} = \lim_{i\to\infty} \mathcal{O}_i$, iff:

(i) $\mathcal{O}$ and $\mathcal{O}_i$ for i=1,2,... all have the same input space $U_{\hat{I}}$.

(ii) If $u \in U_{\hat{I}}$ and $(u,y) \in R_{\hat{I}}$, $(u,y_i) \in R_{i\hat{I}}$ for i=1,2,... then $y = \lim_{i\to\infty} y_i$ in $\mathcal{D}'$.

THM.IV.4.3: Every $\mathcal{O}_G$ object is the limit of objects $\mathcal{O}_i$, which have a finite dimensional reduced state description of the form IV.4.1 and IV.4.2.

PROOF: For each i, the expressions IV.4.5 and IV.4.6 can be written in the matrix form:

$$\frac{d}{dt}\begin{bmatrix} x^i_{\mu 1} \\ x^i_{\mu 2} \\ \vdots \\ x^i_{\mu k^i_\mu} \end{bmatrix} = \begin{bmatrix} \gamma^i_\mu & 1 & 0 \cdots\cdots 0 \\ 0 & \gamma^i_\mu & 1 \cdots\cdots 0 \\ \vdots & \vdots & \ddots \quad \vdots \\ \vdots & \vdots & \quad\ddots\; 1 \\ 0 & 0 & 0 \cdots\cdots \gamma^i_\mu \end{bmatrix} \begin{bmatrix} x^i_{\mu 1} \\ x^i_{\mu 2} \\ \vdots \\ x^i_{\mu k^i_\mu} \end{bmatrix} + \begin{bmatrix} b^i_{\mu 1} \\ b^k_{\mu 2} \\ \vdots \\ b_{\mu k^i_\mu} \end{bmatrix} u(t),$$

$$\text{for } \mu = 1,2,\ldots,q_i. \qquad \text{IV.4.15}$$

We write IV.4.15 in the more compact matrix form:

$$\frac{dX^i_\mu}{dt} = J^i_\mu X^i_\mu + B\, u(t) \qquad \text{IV.4.16}$$

where the definitions of the involved entities is self-explanatory. Now combining IV.4.16 for different values of μ, we obtain:

$$\frac{d}{dt}\begin{bmatrix} X^i_1 \\ X^i_2 \\ \vdots \\ X^i_{q_i} \end{bmatrix} = \begin{bmatrix} J^i_1 & 0 \cdots\cdots 0 \\ 0 & J^i_2 \cdots\cdots 0 \\ \vdots & \vdots \quad \ddots \quad \vdots \\ 0 & 0 \cdots\cdots J^i_{q_i} \end{bmatrix} \begin{bmatrix} X^i_1 \\ X^i_2 \\ \vdots \\ X^i_{q_i} \end{bmatrix} + \begin{bmatrix} B^i_1 \\ B^i_2 \\ \vdots \\ B^i_{q_i} \end{bmatrix} u(t), \text{ for each } i. \qquad \text{III.4.17}$$

Equation IV.4.6 can also be written with matrix notation, as:

$$y(t) = \lim_{i\to\infty} [1_{k^i_1}\, 1_{k^i_2} \cdots 1_{k^i_\mu}] \begin{bmatrix} X^i_1 \\ X^i_2 \\ \vdots \\ X^i_{q_i} \end{bmatrix} + \sum_{k=1}^{K} d_k u^{(k)}(t) \qquad \text{IV.4.18}$$

where $1^i_{k_j} \triangleq [1\ 1 \ldots. 1]_{1\times k_j}$ is a row vector.

Now, defining the I-O list $R_{i\hat{I}}$ by IV.4.17 and

$$y_i(t) = [l_{ki_1} l_{ki_2} \cdots l_{ki_\mu}] \begin{bmatrix} X^i_1 \\ X^i_2 \\ \vdots \\ X^i_{q_i} \end{bmatrix} + \sum_{k=1}^{K} d_k u^{(k)}(t) \qquad \text{IV.4.19}$$

we obtain the objects Θ_i, such that $\Theta = \lim_{i\to\infty} \Theta_i$, since $y(t) = \lim_{i\to\infty} y_i(t)$ in $\mathcal{D}'$. Moreover, the state description obtained from the state equations IV.4.17 and IV.4.19 is a reduced description for each Θ_i by Thm.III.4.2, since:

(i) every matrix J^i_μ in IV.4.17 is an elementary Jordan block, and

(ii) the leading entry of each l_{ki_j} ($c_k^{(1)}$ in Thm.III.4.2) is 1, and the last entry of each B_μ ($b_k^{(d_k)}$ in Thm.III.4.2) is $b^i_{\mu k i_\mu} \triangleq (k^i_\mu - 1)! c_{\mu k i_\mu}$, where $c_{\mu k i_\mu}$ is the coefficient of the term with the highest power of t in each simple exponomial and therefore assumed to be non-zero.

<u>NOTE IV.4.5</u>: Similar to the development in section IV.3, it can again be shown that:

$$H'_{t_0}[u_0] \triangleq \{\hat{u}_0 \in U_{(-\infty,t_0]} : \text{for each } i, \int_{-\infty}^{t_0} \tau^p e^{\tau\gamma^i{}_\mu} \hat{u}_0(\tau)d\tau = \int_{-\infty}^{t_0} \tau^p e^{\tau\gamma^i{}_\mu} u_0(\tau)d\tau \text{ where } p = 1,2,\ldots,k^i{}_\mu - 1 \text{ and } \mu = 1,2,\ldots,q_i\} \qquad \text{IV.4.20}$$

or equivalently that:

$$H_{t_0}[u_0] \triangleq \{\hat{u}_0 \in U_{(-\infty,t_0]} : \text{for each } i,$$

$$\sum_{\nu=j}^{k_\mu{}^i} \frac{b^i{}_{\mu\nu}}{(\nu-j)!} \int_{-\infty}^{t_0} (\tau-t_0)^{\nu-j} e^{(\tau-t_0)\gamma^i{}_\mu} u_0(\tau)d\tau =$$

$$\sum_{\nu=j}^{k_\mu{}^i} \frac{b^i{}_{\mu\nu}}{(\nu-j)!} \int_{-\infty}^{t_0} (\tau-t_0)^{\nu-j} e^{(\tau-t_0)\gamma^i{}_\mu} \hat{u}_0(\tau)d\tau$$

$$\text{for } j=1,2,\ldots,k^i{}_\mu \text{ and } \mu = 1,2,\ldots,q_i\} \qquad \text{IV.4.21}$$

and $\mathcal{H}'_{t_0} = \{H'_{t_0}[u_0] : u_0 \in U_{(-\infty,t_0]}\}$ constitutes a half reduced partitioning. The equivalence of IV.4.20 to IV.4.21 follows from $c_{\mu_{k^i{}_\mu}} \neq 0$ and parts of the proof of Thm.III.4.2. Now defining successively:

$$x^i{}_{\mu j}(t_0) \triangleq \sum_{\nu=j}^{k_\mu{}^i} \frac{b^i{}_{\mu\nu}}{(\nu-j)!} \int_{-\infty}^{t_0} (t_0-\tau)^{\nu-j} e^{(t_0-\tau)\gamma^i{}_\mu} u(\tau)d\tau$$

i,μ,j as in IV.4.21, then

$X^i_\mu(t_0) \triangleq [x^i_{\mu_1}(t_0)\ x^i_{\mu_2}(t_0) \ldots x^i_{\mu_{k^i_\mu}}(t_0)]^T$ for i,μ as in IV.4.21, then $X^i(t_0) \triangleq [X^i_1(t_0)\ X^i_2(t_0) \ldots x^i_{q_i}(t_0)]^T$ for $i=1,2,\ldots$ and finally the set $\Sigma'_G(t_0) \triangleq$

$$\{X(t_0) : X(t_0) = [X^1(t_0), X^2(t_0), \ldots]^T\}, \qquad \text{IV.4.22}$$

we can show the existence of a one to one, onto mapping C'_{t_0} between $\Sigma'_L(t_0)$ and $\mathcal{H}'_{t_0}$ due to IV.4.21. $\Sigma'_G(t_0)$ and

$$y(t) \triangleq \bar{A}_{G(t_0,\infty)}(\sigma_0,u) =$$

$$\lim_{i\to\infty} \sum_{\mu=1}^{q_i} \sum_{j=1}^{k_\mu{}^i} \Big[\sum_{\nu=j}^{k_\mu{}^i} \frac{(t-t_0)^{\nu-j}}{(\nu-j)!} e^{(t-t_0)\gamma^i{}_\mu} x^i{}_{\mu\nu}(t_1)\Big] +$$

$$w(t)*u(t), \text{ for } t > t_0 \qquad \text{IV.4.23}$$

constitute a half reduced state description of the object $\mathcal{O}_G$ (IV.4.23 is easily derivable from IV.4.11 and IV.4.14). An important point of the state description $(\Sigma'_G, \bar{A}_G)$ is the countable dimension of its state vectors, Def.IV.4.22.

CHAPTER V

CONCLUSIONS

It is our hope that with the discussion in Chapter II, the state axioms have reached their final form. The main contributions of this chapter have been this final form of the state axioms and the establishment of the strong connection between equivalence classes of inputs and reduced state descriptions. We have shown that the reduced state description of a causal object is almost unique. This chapter has also provided us with means of constructing state descriptions that are half reduced.

Using the concepts and the results of Chapter II in Chapter III, we proved that linear and/or time-invariant objects can always be provided with linear and/or time-invariant state descriptions; a result of rather academic value which shows that properties of objects need not, and it is our belief that they should not, be given in terms of their state descriptions.

One of the contributions of Chapter IV has been to obtain a half reduced state description for a large class of distributed objects based on the construction in Chapter II, and to generalize concepts such as "Fundamental Matrix" used in lumped systems. The other main result of this chapter was to approximate very general distributed systems by lumped objects possessing reduced state descriptions using results from Chapter IV.

Many open questions that constitute a rich basis for further research arose during the development of the thesis. A few important ones, starting with the obvious question about the state description of non-linear and/or time varying objects, are:

1. The reflection in the state description of properties other than linearity and time-invariance, such as continuity, of the system.
2. The forms the state descriptions will take after interconnections of different objects necessitating a study of the equivalence classes of inputs from the individual classes of each system.
3. Studies about the stability of the system using the Hilbert matrix representation obtained in Chapter IV and spectral theory.
4. The approximation of distributed systems by stable and lumped (or lumped RLC) objects by placing restrictions on their convolutional representation.

Finally, the synthesis procedures obtained in [DA], for the state description in [RE5] (given in Chapter I), constitute another solid justification and application of the State Space Theory. It is strongly possible that some syntehsis procedures can also be derived from the state descriptions in Chapter IV of this thesis.

LIST OF REFERENCES

[AR] Arsac, J. Fourier Transforms and the Theory of Distributions. Englewood Cliffs, N.J.: Prentice-Hall, 1966.

[BA1] Balakrishnan, A. V. "Linear Systems with Infinite Dimensional State Space." Symposium on System Theory, Polytechnic Institute of Brooklyn, April 20, 21, 22, 1965.

[BA2] _____. "On the Problem of Deducing States and States Relations from Input-Output Relations for Linear Time Varying Systems." SIAM Journal on Control, Vol. 5, No. 3 (1967), pp. 309-325.

[BA3] _____. "On the State Space Theory of Linear Systems." Journal of Mathematical Analysis and Applications, Vol. 14 (1966), pp. 371-391.

[BA4] _____. "Foundations of the State Space Theory of Continuous Systems 1." Journal of Computer and System Sciences, Vol. 1 (1967) pp. 91-116.

[BA] Bashkow, T. "The A Matrix, New Network Description." IRE Transaction on Circuit Theory, Vol. CT-4 (September, 1957), pp. 117-119.

[BEL] Beltrami, E. J. and Wholers, M. R. Distributions and Boundary Values of Analytic Functions. New York: Academic Press, 1966.

[BE] Berberian, S. K. Introduction to Hilbert Space. New York: Oxford University Press, 1961.

[BRE] Bremermann, H. J. "Some Remarks on Analytic Representations and Products of Distributions." J. SIAM Appl. Math., Vol. 15 (1967), pp. 929-943.

[BR] Bryant, P. R. "Order of Complexity of Electrical Networks." Ins. Elec. Engrs. (London). Monograph No. 335 E, June, 1959.

[COL] Coddington, E. A. and Levinson, N. Theory of Ordinary Differential Equations. New York: McGraw-Hill, 1955.

[CO] Cooke, R. G. Infinite Matrices and Sequence Spaces. London: Macmillan, 1950.

[DA] Daryanani, G. T. and Resh, T. A. "Foster Distributed-Lumpted Network Synthesis." IEEE Transactions on Circuit Theory, and the 1968-IEEE International Symposium on Circuit Theory, Miami Beach, Florida, December, 1968.

[DO1] Dolezal, V. Dynamic of Linear Systems. Prague: Publishing House of the Czechoslovak Academy of Sciences, 1964.

[DO2] ______. "On Linear Passive n-parts with Time-Varying Elements." J. SIAM Appl. Math, Vol. 15 (1967), pp. 1018-1029.

[GU] Guttinger, W. "Generalized Functions in Elementary Particle Physics and Passive System Theory: Recent Trends and Problems." J. SIAM Appl. Math., Vol. 15 (1967), pp. 964-1000.

[HE] Hewitt, E. and Stromberg, K. "Real and Abstract Analysis." Springer-Verlag, New York, Inc., 1965.

[HO] Hoffmann, K and Kunze, R. Linear Algebra. Englewood Cliffs, N.J.: Prentice Hall, 1961.

[HOR] Horwáth, J. Topological Vector Spaces and Distributions. Reading, Mass.: Addison-Wesley, 1966.

[KA] Kalman, R. E. "Mathematical Description of Linear Dynamical Systems." J. SIAM Control, Ser. A, Vol. 1, No. 2 (1963), pp. 152-192.

[HU] Huffman, D. "The Synthesis of Sequential Switching Circuits." J. Franklin Inst., Vol. 257 (April, 1954), pp. 161-190.

[KE] Kestelman, H. Modern Theories of Integration. Oxford at Clarendon Press, 1937.

[KI] Kinarawala, B. "Analysis of Time Varying Networks." 1961 IRE International Convention Record, pt. 4, pp. 268-276.

[LI] Liverman, T. P. G. "Physically Motivated Definitions of Distributions." J. SIAM Appl. Math., Vol. 15, (1967), pp. 1048-1076.

(MA] Manning, A. An Introduction to Animal Behavior. A series of student texts in Contemporary Biology. London: Addison-Wesley, 1967.

[MO] Moore, E. F. "Gedanken Experiments on Sequential Machines." Automata Studies. Princeton, N. J.: Princeton University Press, 1956, pp. 129-153.

[NE1] Newcomb, R. W. "The Foundations of Network Theory." The Inst. of Engr., Elec. and Mech. Trans., Vol. EM-6 (1964), pp. 7-12, Australia.

[NE2] ______. Linear Multiport Synthesis. New York: McGraw-Hill, 1966.

[PA] Papoulis, A. The Fourier Integral and its Applications, Electronic Science Series. New York: McGraw-Hill, 1962.

[PO] Porter, W. A. Modern Foundations of Systems Engineering. New York: The Macmillan Company, 1966.

[RE1] Resh, J. A. "An Improvement in the State Axioms." Report No. R-282, Coordinated Science Laboratory, University of Illinois, Urbana, Illinois, March, 1966.

[RE2] ______. "Improvements in the State Axioms." Conference Record, Tenth Midwest Symposium on Circuit Theory, Purdue University, Lafayette, Indiana, May, 1967.

[RE3] ______. "On the Construction of State Spaces." Proceedings Fifth Annual Allerton Conference on Circuit and System Theory, University of Illinois, Urbana, Illinois, October, 1967.

[RE4] ______. "On Canonical State Equations for Distributed Systems." Conference Record, Eleventh Midwest Symposium on Circuit Theory, University of Notre Dame, Notre Dame, Indiana, May, 1968.

[RE5] ______, and Göknar, I. C. "Derivation of Canonical State Equations for A Class of Distributed Systems." Proceedings Sixth Annual Allerton Conference on Circuit and System Theory, University of Illinois, Urbana, Illinois, October, 1968.

[RI] Riesz, F. and Nagy, G. S. Z. Leçons d'Analyse Fonctionelle. Budapest: Academie des Sciences de Hongrie, 1953.

[SC1] Schwartz, L. Theorie des Distributions. Vols. I and II. Actualités Scientifiques et Industrielles. Paris: Hermann et Cie., 1950-1951.

[SC2] ______. Methodes Mathématiques pour les Sciences Physiques. Paris: Hermann et Cie., 1961.

[SH] Shannon, C. E. "A Mathematical Theory of Communication." Bell Sys. Tech. J., Vol. 27 (1948), pp. 379-423, 623-656.

[TR] Treves, F. Topological Vector Spaces, Distributions and Kernels. New York: Academic Press, 1967.

[WE] Weiss, L. and Kalman, R. E. "Contributions to Linear System Theory." RIAS Tech. Report, No. 64-69 (April, 1964).

[WO] Wohlers, M. R. and Beltrami, E. J. "Distribution Theory as the Basis of Generalized Passive Network Analysis." IEEE Trans. on Circuit Theory, Vol. CT-12 (1965), pp. 164-170.

[ZA1] Zadeh, L. A. "From Circuit Theory to System Theory." Proceedings of the IRE, Vol. 50, No. 5 (May, 1962), pp. 856-865.

[ZA2] ______, and Dosoer, C. A. Linear System Theory. New York: McGraw-Hill, 1963.

[ZA3] ______. "The Concept of State in System Theory." Views on General System Theory. Edited by M. O. Mesarovic. New York: John Wiley and Sons, 1963

[ZE1] Zemanian, A. H. Distribution Theory and Transform Analysis. New York: McGraw-Hill, 1965.

[ZE2] ______. "Orthonormal Series Expansions of Certain Distributions and Distributional Transform Calculus." Journal of Math. Anal. Appl., Vol. 14, No. 2 (May, 1966), pp. 263-275.

[ZE3] Zemanian, A. H. "An Introduction to Generalized Functions and the Generalized Laplace and Legendre Transformations." SIAM Review, Vol. 10, No. 1 (January, 1968).

[ZE4] ______. "Applications of Generalized Functions to Network Theory." Summer School on Circuit Theory, Prague, Czechoslovakia, June-July, 1968.

APPENDIX A

A.1--About Distribution Theory

It has been some twenty years since Schwartz introduced and developed his theory of distributions, a theory that owes its birth to physicists, who have used the delta function since the nineteenth century [ZE1 preface].

Mathematicians have plunged into it and a large body of mathematical literature has been published in areas such as ordinary and differential equations, operational calculus, transformation theory and functional analysis. This impetus mathematics has gained from physics did not prevent the more and more abstractization of distribution theory, which is now going the entangled paths of topology and topological vector spaces [TR, HOR].

In mathematical sciences, the most notable application of distribution theory has been to quantum field theory [ZE4, p. 1]. In network and system theory it has been extensively used in the axiomatic foundation of system theory [ZE1, 4], in the time-domain theory of linear n-ports, in obtaining a frequency-domain criterion for the causality of active networks [ZE4, secs. 4, 5, 6], in the theory of generalized Bode equations and in the characterization of various broad classes of systems by their real frequency behavior, [AR], [BEL], [WO], [GU], etc. Distributions have also been used in an essential way in the analysis and synthesis of time varying networks, see e.g. [NE1-2], [DO1-2]. In other subjects, various classical problems which had been solved in terms of classical mathematics, become open problems once again when reformulated in terms of distribution theory [ZE4, pp. 1].

There exists a serious drawback to distribution theory and this is its uselessness in the theory of nonlinear systems, which is due to the fact that the product of two distributions cannot be defined in general, but only when one of the distributions is a special one. However, efforts are being made to generalize the product of distributions which are used in quantum field theory [BRE]. This may render possible the application of distributions, at least to special classes of nonlinear systems.

Despite this extensive use of distributions, some applied scientists are reluctant to accept the description of physical quantities by a concept that is not an ordinary point function, but is something of functional nature [PA]. That this objection of philosophical nature is not justifiable, can be shown as follows. First of all we can take the attitude of Newcomb when, he assumes physical variables are infinitely differentiable, and justifies it with: "since no physical measurement can prove otherwise" [NE2, pp. 6]. We can equally well say "all physical quantities are distributions since no physical measurement can prove otherwise." However we shall try to do more, since such reasonings can prove anything (i.e. nothing).

To begin with, the assumption that a physical variable F can be characterized with an ordinary function f(t) is a convenient idealization [PA]. Why should we not characterize F with a distribution, if that is also a convenient idealization (which we think it is)? A propos, Zemanian writes: "It is impossible to observe the instantaneous values f(t) of F. Any measuring instrument would merely record the effect that F produces on it over some nonvanishing interval of time" [ZE1]. Now, it may be that this uncertainty about F being representable by a function, is due to the imperfection of our measuring instruments. The fact is that the imperfection will always be there since indicators will always be subject to parasitic effects such as mass, and we will always justify our theories with measurements using such instruments. Thus, it is a more realistic assumption, as we can infer that much from the physical measurements, to characterize F by a distribution.

Finally, Liverman in a recent Article [LI], gives a physical motivated definition of distributions, by showing that one obtains the same space of distributions when one confines himself to the testing functions that are probability densities, i.e. $\rho(x) \geqslant 0$, $\rho(x)$ is infinitely smooth and $\int \rho(x)dx = 1$, instead of considering the space $\mathfrak{D}$ of all testing functions. As physical background Liverman roughly says: if $f(x,t)$ is the characterization of the physical variable F, where x and t are the space and time variables, a measurement of F yields a quantity $f(x,t)+e_n(x,t)$, where e_k's are error functions and a particular one e_n is in effect during the experiment. Furthermore, a measurement of the location (x,t), actually occurs in $(\xi,\xi+d\xi)x(\tau,\tau+d\tau)$ with probability $\rho(\xi,\tau)d\xi.d\tau$. Then the expected value of a measurement of F, intended to be at (x,t) is the weighted average: $\langle f+e_n,\rho\rangle = \langle f,\rho\rangle + \langle e_n,\rho\rangle$. The functions e_k are random and we assume the expected value of $\langle e_k,\rho\rangle$ over various k, to be zero. Thus:

$$\langle f+e_n,\rho\rangle = \langle f,\rho\rangle = \int f(\xi,\tau)\rho(\xi,\tau)d\xi.d\tau.$$

Finally Liverman points that, to say $\lim_{\nu} f_\nu(x,t)$ exists pointwise or uniformly becomes a physically non-verifiable, mathematical assertion. The statement $\lim_{\nu} \langle f_\nu,\rho\rangle$ exists for every probability density is operationally much more relevant, and consistency requires that we include into our list all functionals f such that $\langle f,\rho\rangle = \lim_{\nu} \langle f_\nu,\rho\rangle$ for all probability densities ρ. This leads us to generalized functions, which turns out to be a better pencil and paper depiction of physical phenomena, in the presence of errors in the experimental determination of physical variables.

In the following sections of the Appendix we introduce the necessary definitions and already proven results in A.2., and prove some new results in A.3. mainly about the orthonormal series expansions of distributions, that are needed in Chapter IV.

A.2--A Brief Review; Some Definitions and Results in Distribution Theory

NOTE A.2.1: The definitions and notations used are consistent with those used in [ZE1,2]. Known results are given without proof, where a reference to the proof is made with the page and theorem number of the corresponding theorem in the literature.

DEF.A.2.1:

A function is INFINITELY SMOOTH ON A SET iff it has continuous derivatives of all orders on that set.

The space of all complex valued functions $\rho(t)$ that are infinitely smooth and zero outside some finite interval is called THE SPACE OF TESTING FUNCTIONS, and is denoted by $\mathfrak{D}$.

DEF.A.2.2:

A sequence of testing functions $\{\rho_\gamma(t)\}_{\gamma=1}^{\infty}$ CONVERGES IN $\mathfrak{D}$ iff the $\rho_\gamma(t)$ are all in $\mathfrak{D}$, are all zero outside some

THM.A.2.5 [ZE1, pp. 115, Thm. 5.2.1]

The direct product $f(t)xg(\tau)$ of two distributions $f(t)$ and $g(\tau)$ is a distribution in $\mathcal{D}'_{t,\tau}$.

DEF.A.2.5:

The CONVOLUTION of two distributions f and g over $\mathbb{R}$ is given by the expression

$$\langle f*g,\rho\rangle \triangleq \langle f(t)xg(\tau),\rho(t+\tau)\rangle \triangleq \langle f(t),\langle g(\tau),\rho(t+\tau)\rangle\rangle \qquad \text{A.2.3}$$

NOTE A.2.5: A problem arises in the definition of the convolution. In A.2.2, $\rho(t,\tau)$ and thus $\langle g(\tau),\rho(t,\tau)\rangle$ had bounded support, but in A.2.3 $\rho(t+\tau)$ is infinitely smooth without having bounded support and therefore it is not a testing function. However a meaning can be attached to A.2.3 if either the supports of f and g are suitably restricted or some conditions are placed on the behavior of the distributions as their arguments approach infinity (we will not give the theorems related to this last situation because definitions of new testing function and distribution spaces are required; they may be found in [SC1, vol. II]). The following theorem illustrates when the convolution process can be given a meaning. In section A.3 we investigate another case that is not given in the literature where the convolution can be defined.

fixe

inte

for

FACT

of e

are

the

DEF

to a

such

<f

α

NO

di

ev

co

b

e

fixed finite interval I and for every fixed nonegative interger k, the sequence $\{\rho_\gamma^{(k)}(t)\}_{\gamma=1}^{\infty}$ converges uniformly for $-\infty<t<\infty$.

FACTS A.2. [ZE1, pp.5]

1: $\mathfrak{D}$ is closed under convergence in $\mathfrak{D}$ i.e. the limit of every sequence that converges in $\mathfrak{D}$, is also in $\mathfrak{D}$.

2: $\{\rho_\gamma(t)\}_{\gamma=1}^{\infty}$ converges in $\mathfrak{D}$, to ρ iff all the ρ_γ are in $\mathfrak{D}$ and are zero outside a fixed finite interval and the sequence $\{\rho_\gamma - \rho\}_{\gamma=1}^{\infty}$ converges to zero in $\mathfrak{D}$.

DEF.A.2.3:

Denoting the functional by f, the number it assigns to any $\rho \in \mathfrak{D}$ by $<f,\rho>$ a DISTRIBUTION is a functional on $\mathfrak{D}$ such that:

$$<f,\rho_1+\alpha\rho_2> = <f,\rho_1> + \alpha<f,\rho_2> \text{ for } \rho_1,\ \rho_2 \in \mathfrak{D} \text{ and } \alpha \in \mathbb{C}$$

if $\{\rho_\gamma(t)\}_{\gamma=1}^{\infty}$ converges to 0 in $\mathfrak{D}$ then the numbers $<f,\rho_\gamma>$ converge to 0. The space of all distributions on $\mathfrak{D}$, denoted by $\mathfrak{D}'$, is called the DUAL SPACE OF $\mathfrak{D}$.

NOTE A.2.2: In most of our discussions we will deal with distributions that are defined over the real line $\mathbb{R}$. However for some theorems, especially the ones about the convolution of distributions, we will have to use distributions over n-dimensional spaces. Thus we have to expand our definitions to multi-dimensional cases. For this let

tha

par

Den

whe

(p_γ

out

ver

fu

p_γ

DE

<t

o

z

S

w

e

let $x \triangleq (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. The TESTING FUNCTIONS are those that vanish outside a compact set in $\mathbb{R}^n$ and for which all partial derivatives exist and are continuous for all x. Denoting the partial derivative by

$$D^k \rho(x) \triangleq \frac{\partial^{k_1+k_2+\cdots+k_n}}{\partial x_1^{k_1} \partial x_2^{k_2} \cdots \partial x_n^{k_n}} \rho(x_1, x_2, \ldots, x_n)$$

where $k \triangleq k_1 + k_2 + \cdots + k_n$ a sequence of testing functions $\{\rho_\gamma(x)\}_{\gamma=1}^{\infty}$ CONVERGES IN $\mathfrak{D}$ TO ZERO iff all $\rho_\gamma(x)$ are zero outside a fixed compact subset of $\mathbb{R}^n$ and $\{D^k \rho_\gamma(x)\}_{\gamma=1}^{\infty}$ converges to zero for any choice of k.

Again, a DISTRIBUTION ON $\mathbb{R}^n$ is a linear, continuous functional on $\mathfrak{D}$ defined over $\mathbb{R}^n$ (continuous in the sense $\rho_\gamma \to 0$ in $\Longrightarrow \langle f, \rho_\gamma \rangle \to 0$ in $\mathbb{C}$).

<u>DEF.A.2.4</u>:

Two distributions f and g are said to be EQUAL iff $\langle f, \rho \rangle = \langle g, \rho \rangle, \ \forall \rho \in \mathfrak{D}$.

The SUPPORT of a testing function $\rho \in \mathfrak{D}$ is the closure of the set of all points where $\rho(t)$ is different than zero, and is denoted by supp $\rho(t)$.

Two distributions f and g are EQUAL OVER THE OPEN SET Ω iff $\langle f, \rho \rangle = \langle g, \rho \rangle$ for every testing function ρ, with supp $\rho(t) \subset \Omega$.

The complement of the union of all open sets, over each of which a distribution f equals zero, is called the

SUPPORT of f, denoted supp f(t). If a set Θ contains the support of a distribution, that distribution is said to be CONCENTRATED ON Θ.

<u>THM.A.2.1:</u> [ZE1, pp. 30, Thm. 1.8.1.]

If a distribution is equal to zero on every set of a collection of open sets, then it is equal to zero on the union of these sets.

<u>THM.A.2.2</u>: [TR, pp. 266, Thm. 24.6]

The distributions in $\mathbb{R}$ which are concentrated on a point, are the finite linear combinations of the δ-functional and its derivatives.

<u>DEF.A.2.5</u>:

A sequence of distributions $\{f_\gamma\}_{\gamma=1}^{\infty}$ CONVERGES IN $\mathfrak{D}'$ iff for every $\rho\in\mathfrak{D}$ the sequence of numbers $\{\langle f_\gamma,\rho\rangle\}_{\gamma=1}^{\infty}$ converges. The LIMIT $\langle f,\rho\rangle$ of $\{\langle f_\gamma,\rho\rangle\}_{\gamma=1}^{\infty}$ defines a functional on $\mathfrak{D}$, and the next theorem proves that f is a distribution.

A series $\sum_{\gamma=1}^{\infty} f_\gamma$ of distributions CONVERGES in $\mathfrak{D}'$ iff the sequence $h_m \triangleq \sum_{\gamma=1}^{m} f_\gamma$ of partial sums converges in $\mathfrak{D}'$.

<u>THM.A.2.2</u>: [ZE1, pp. 37, Thm. 2.2.1]

If a sequence of distributions $\{f_\gamma\}_{\gamma=1}^{\infty}$ converges in $\mathfrak{D}'$ to the functional f, then f is also a distribution i.e. the space $\mathfrak{D}'$ is closed.

NOTE A.2.3: One way of generating an important class of distributions is to imbed locally summable functions into $\mathcal{D}'$ through the convergent integral [ZE2, pp. 264]

$$<T_f,\rho> \triangleq \int_{-\infty}^{\infty} f(t)\cdot\overline{\rho(t)}\cdot dt \quad \forall\rho\in\mathcal{D} \qquad \text{A.2.1}$$

More precisely, the distribution T_f, because of A.2.1, represents the equivalence class of functions that equal f almost everywhere. It is also worthwhile to note that if $T_f = T_g$ in $\mathcal{D}'$ then $f \overset{a.e.}{=} g$ is also true. Thus we shall denote T_f by f, any function in the equivalence class that T_f represents, and call such distributions REGULAR DISTRIBUTIONS.

A.2.1 is not the only way to generate distributions from functions. Another standard procedure that leads to the concept of PSEUDOFUNCTION is given in [ZE1], [TR]. However there also are functions such as $e^{1/t}$ which do not define distributions no matter what procedure one tries on them [TR, pp. 226].

The above simple discussion is useful since we consider regular distributions frequently in our work and is necessary for the next theorem that is of importance in section IV.4.

THM.A.2.3: [TR, pp. 304, Thm. 38.3]

Let Ω be an open subset of $\mathbb{R}^n$. Any distribution in Ω is the limit of a sequence of polynomial functions in $\mathcal{D}'$.

NOTE A.2.4: Now we concentrate on the convolution of distributions which is a very general process. Various types of differential equations, difference equations and integral equations are all special cases of convolution equations [ZE1, pp. 114]. The convolution is also a very general way of characterizing linear, time-invariant and continuous systems that we use in our developments of Chapter IV.

THM.A.2.4: [ZE1, pp. 74, Cor. 2.7.2a]

Let x be an n-dimensional real variable and y an m-dimensional real variable. Also, let $\rho(x,y)$ be a testing function in $\mathcal{D}$ defined over $\mathbb{R}^{n+m}$. If $f(x)$ is a distribution defined over $\mathbb{R}^n$, then $\Theta(y) \triangleq \langle f(x),\rho(x,y)\rangle$ is a testing function of y in $\mathcal{D}$ and an arbitrary partial derivative $D_y^k\Theta(y)$ with respect to the components of y is given by: $D^k\Theta(y) = \langle f(x), D_y^k\rho(x,y)\rangle$.

DEF.A.2.6:

Let $\rho(t,\tau)$ be a testing function in $\mathcal{D}_{t,\tau}$, defined over $\mathbb{R}^2$, and let $f(t)\in\mathcal{D}'_t$ $g(\tau)\in\mathcal{D}'_\tau$ be distributions over $\mathbb{R}^1$. Then by THM.A.2.4. $\langle g(\tau),\rho(t,\tau)\rangle$ is a testing function in $\mathcal{D}_t$ and the DIRECT PRODUCT $f(t)xg(\tau)$ is defined by

$$\langle f(t)xg(\tau),\rho(t,\tau)\rangle \triangleq \langle f(t), \langle g(\tau), \rho(t,\tau)\rangle\rangle \qquad \text{A.2.2}$$

<u>THM.A.2.5</u> [ZE1, pp. 115, Thm. 5.2.1]

The direct product $f(t)xg(\tau)$ of two distributions $f(t)$ and $g(\tau)$ is a distribution in $\mathcal{D}'_{t,\tau}$.

<u>DEF.A.2.5</u>:

The CONVOLUTION of two distributions f and g over $\mathbb{R}$ is given by the expression

$$\langle f*g,\rho\rangle \triangleq \langle f(t)xg(\tau),\rho(t+\tau)\rangle \triangleq \langle f(t),\langle g(\tau),\rho(t+\tau)\rangle\rangle \qquad \text{A.2.3}$$

<u>NOTE A.2.5</u>: A problem arises in the definition of the convolution. In A.2.2, $\rho(t,\tau)$ and thus $\langle g(\tau),\rho(t,\tau)\rangle$ had bounded support, but in A.2.3 $\rho(t+\tau)$ is infinitely smooth without having bounded support and therefore it is not a testing function. However a meaning can be attached to A.2.3 if either the supports of f and g are suitably restricted or some conditions are placed on the behavior of the distributions as their arguments approach infinity (we will not give the theorems related to this last situation because definitions of new testing function and distribution spaces are required; they may be found in [SC1, vol. II]). The following theorem illustrates when the convolution process can be given a meaning. In section A.3 we investigate another case that is not given in the literature where the convolution can be defined.

THM.A.2.6 [ZE1, pp. 124, Thm. 5.4.1]

Let f and g be two distributions over $\mathbb{R}$. Then f*g exists as a distribution over $\mathbb{R}$, under any one of the following conditions:

(i) Either f or g has a bounded support

(ii) Both f and g have supports bounded on the left (or on the right).

THM.A.2.7 [ZE1, pp. 124, Ex. 5.4.1]

If f and g are locally summable functions whose supports satisfy one of the conditions stated in THM.A.2.6, then their distributional convolution $h(t) = f(t)*g(t)$ is given almost everywhere by the regular distribution corresponding to the locally integrable function

$$h(t) = \int_{-\infty}^{\infty} f(\tau).g(t-\tau)d\tau.$$

THM.A.2.8: [ZE1, pp. 127, Ex. 5.4.3]

The convolution of $\delta^{(m)}(t-a)$ with any distribution in $\mathcal{D}'$, is given by: $\delta^{(m)}(t-a)*f(t) = f^{(m)}(t-a)$ $m=1,2,\ldots$

THM.A.2.9: [ZE1, pp. 132]

A convolution may be differentiated, by differentiating either one of the distributions in it, i.e.

$$(f(t)*g(t))^{(m)} = f^{(m)}(t)*g(t) = f(t)*g^{(m)}(t)$$

<u>THM.A.2.10</u>: [ZE1, pp. 136, Thm. 5.6.1]

Let the sequence of distributions $\{f_\gamma\}_{\gamma=1}^{\infty}$ converge in $\mathcal{D}'$ to f. Then $\{f_\gamma * g\}_{\gamma=1}^{\infty}$ converges in $\mathcal{D}'$ to f*g if $\{f_\gamma\}_{\gamma=1}^{\infty}$, f and g all have supports bounded from the left.

<u>NOTE A.2.6</u>: The remaining part of this section is devoted to the orthonormal series expansion of certain distributions as given by Zemanian [ZE2] and which constitutes the main tool in obtaining an infinite dimensional state description of a large class of systems in section IV.3. We first give the necessary notation, then state the theorems that we use later.

<u>NOT.A.2.1</u>: [ZE2, pp. 262-265]

I = (a,b) denotes an open interval on the real line and the case $a = -\infty$, $b = \infty$ is not excluded. L_I^2 is the space of square summable functions on I with the usual inner product $\langle f,g\rangle = \int_a^b f(t)\overline{g(t)}dt$ for $f,g \in L_I^2$.

$\mathcal{D}_I$ denotes the space of all testing functions in $\mathcal{D}$, whose supports are contained in I. $\mathcal{D}'_I$ is the space of distributions defined on $\mathcal{D}_I$.

With $\theta_k(t) \neq 0$ and infinitely smooth on I, η denotes the linear differentiation operator: $\eta \triangleq \theta_0 D^{n_1} \theta_1 D^{n_2} \ldots D^{n_\nu} \theta_\nu$, where the n_k are nonnegative integers and $D^k = \frac{d^k}{dt^k}$. The θ_k and n_k are so chosen that $\eta = \overline{\theta}_\nu(-D)^{n_\nu} \ldots (-D)^{n_2} \overline{\theta}_\nu(-D)^{n_1} \overline{\theta}_0$. Moreover it is assumed that η

poss

Ψ_n,

ple

no

of

are

du

di

fo

(

DE

f

a

(

c

e

possesses real eigenvalues λ_n and normalized eigenfunctions Ψ_n, $n=1,2,\ldots$ with the properties that $\{\Psi_n\}_{n=1}^{\infty}$ is a complete orthonormal sequence in L_I^2 and the λ_n are real, have no finite point of accumulation (this also means no value of λ_n is assumed more than a finite number of times) and are so numbered that $|\lambda_1|\leqslant|\lambda_2|$ ...

The use of the symbol $<.,.>$ to denote the inner product in L_I^2 conforms with its use to denote the number, a distribution makes correspond to a testing function, since for a regular distribution f, $<f,\rho> = \int_{-\infty}^{\infty} f(t)\overline{\rho}(t)dt$ (Note A.2.3).

DEF.A.2.8:

The set of all infinitely smooth, complex valued functions on I, such that $\gamma_k(\rho)\underline{\underline{\Delta}}[\int_a^b |\eta^k\rho(t)|^2 dt]^{\frac{1}{2}}<\infty, k=0,1,2,\ldots$ and $<\eta^k\rho,\Psi_n> = <\rho,\eta^k\Psi_n>$ for each n and k, is the space $\mathfrak{A}$ OF THE TESTING FUNCTIONS with γ_k's taken as seminorms of $\mathfrak{A}$.

A sequence $\{\rho_\gamma\}_{\gamma=1}^{\infty}$ is a CAUCHY SEQUENCE IN $\mathfrak{A}$ if each ρ_γ is in $\mathfrak{A}$ and for each k, $\gamma_k(\rho_\gamma-\rho_i)\to 0$ as γ and i tend independently to infinity. The corresponding convergence is referred to as CONVERGENCE IN $\mathfrak{A}$.

FACT.A.2: [ZE2, pp. 265]

3: $\mathfrak{D}_I$ is contained in $\mathfrak{A}$ and convergence in $\mathfrak{D}_I$ implies convergence in $\mathfrak{A}$.

4: Each Ψ_n is in $\mathfrak{A}$.

THM.A.

COR.A

conv

comp

THM.

ser

THM

Th

c

THM.A.2.11: [ZE2, pp. 265, Thm. 1]

$\mathfrak{A}$ is a sequentially complete space.

COR.A.2.1: [ZE2, pp. 268]

If $\{\rho_n\}_{n=1}^{\infty}$ is a sequence of testing functions that converges in $\mathfrak{A}$, then $\{\rho_n\}_{n=1}^{\infty}$ converges uniformly on every compact subset of I.

THM.A.2.12: [ZE2, pp. 267, Lem. 1]

If ρ is in $\mathfrak{A}$ then $\rho(t) = \sum_{n=1}^{\infty} \langle\rho,\Psi_n\rangle\ \Psi_n(t)$ where the series converges in $\mathfrak{A}$.

THM.A.2.13: [ZE2, pp. 268, Lem.2]

Let $\{a_n\}_{n=1}^{\infty}$ denote a sequence of complex numbers.

Then, $\sum_{n=1}^{\infty} a_n \Psi_n$ converges in $\mathfrak{A}$ iff $\sum_{n=1}^{\infty} |\lambda_n|^{2k}|a_n|^2$ converges for every k.

DEF.A.2.9:

The set of all linear, continuous functionals on $\mathfrak{A}$ is the SPACE OF DISTRIBUTIONS $\mathfrak{A}'$, and the number that $f \in \mathfrak{A}'$ assigns to any $\rho \in \mathfrak{A}$ is denoted by $\langle f,\rho\rangle$. (By a continuous functional on $\mathfrak{A}$ we again mean if $\rho_\gamma \to 0$ in $\mathfrak{A}$ then the numbers $\langle f,\rho_\gamma\rangle \to 0$).

A sequence of distributions $\{f_\gamma\}_{\gamma=1}^{\infty}$ CONVERGES IN $\mathfrak{A}'$ iff for every $\rho \in \mathfrak{A}$ the sequence of numbers $\{\langle f_\gamma,\rho\rangle\}_{\gamma=1}^{\infty}$ converges i.e. $\mathfrak{A}'$ has the weak topology generated by the seminorms $\eta_\phi(f) = |\langle f,\rho\rangle|$.

THM.A.2.14: [ZE2, pp. 269, Thm. 2]

$\mathfrak{A}'$ is a sequentially complete space.

FACT A.2 [ZE2, pp. 269]

5: By FACT A.2.3 the restriction of $f \in \mathfrak{A}'$ to $\mathfrak{D}_I$ is in $\mathfrak{D}'_I$, and convergence in $\mathfrak{A}'$ implies convergence in $\mathfrak{D}'_I$.

6: By the above fact, L^2_I and therefore $\mathfrak{A}$ is imbedded into $\mathfrak{A}'$ by defining the number $f \in L^2_I$ assigns to $\rho \in \mathfrak{A}$ as

$$\langle f,\rho\rangle \triangleq \int_b^a f(t)\overline{\rho}(t)dt.$$

NOTE A.2.7: Another subspace of $\mathfrak{A}'$ is the space of all distributions with compact support in I. This with FACT A.2.6 give us an idea about the size of $\mathfrak{A}'$. FACT A.2.6 also confirms us of the consistency to use the symbol $\langle .,. \rangle$ in DEF.A.2.9.

The next theorem is the result which required all this preparation.

THM.A.2.15 [ZE2, pp. 270, Thm. 3]

If $f \in \mathfrak{A}'$ then $f = \sum_{n=1}^{\infty} \langle f,\psi_n\rangle \psi_n(t)$ where the series converge in $\mathfrak{A}'$.

THM.A.2.16 [ZE2, pp. 270, Thm.5]

Let b_n denote complex numbers. Then $\sum_{n=1}^{\infty} b_n \psi_n(t)$ converges in $\mathfrak{A}'$ iff there exists an integer $q \geqslant 0$ such that $\sum_{\lambda_n \neq 0} |\lambda_n|^{-2q} |b_n|^2$ converges. Moreover, if f denotes the sum $\sum_{n=1}^{\infty} b_n \psi_n(t)$ in $\mathfrak{A}'$ then $b_n = \langle f, \psi_n \rangle$.

A.3--Some New Results

NOTE A.3.1:

In this section we start by stating some facts, which are already well known, then we continue with some lemmas and theorems that are necessary to define the convolution of a distribution in $\mathfrak{A}'$, with inputs from the input space $U_{(-\infty,\infty)}$ of section IV.3. The interval of interest is $\hat{I} = (-\infty,\infty)$, Not.A.2.1 and the Θ_k's in the definition of the differential operator η are assumed to be bounded on (T_k,∞) for some T_k, k=0,1,....

FACT A.3.1:

$u(t) \in U_{D_{(-\infty,\infty)}} \Longrightarrow u(t)$ is locally summable. $U_{D_{(-\infty,\infty)}}$ is as in DEF.IV.3.1.

FACT A.3.2:

$\{\psi_n(t)\}_{n=1}^{\infty}$ is a complete orthonormal sequence for $L^2_{(-\infty,\infty)} \Longleftrightarrow \{\psi_n(\tau-t)\}_{n=1}^{\infty}$ is one for any finite τ.

FACT A.3.3:

$$\int_{-\infty}^{\infty} |f(t)|^2 dt < \infty \quad \text{and } f(t) \text{ is infinitely smooth} \Rightarrow f(t)$$

is bounded everywhere and $\lim\limits_{|t| \to \infty} f(t) = 0$.

FACT A.3.4:

$$\int_{-\infty}^{\infty} |f(t)|^2 dt < \infty \quad \text{and } f(t) \text{ is infinitely smooth} \implies$$

$$\int_{-\infty}^{\infty} \left| \frac{d^k f(t)}{dt^k} \right|^2 dt < \infty \text{ for } k = 1,2,\ldots.$$

NOTE A.3.2: The following two lemmas are necessary for the result of the convolution, to be later defined, to be in $\mathcal{D}$. The first one exhibits a special kind of testing function in $\mathcal{A}$. The second one provides us with a certain convergence, both to be used in the definition of the convolution. $U_{D_{(-\infty,\infty)}}$ we use in these lemmas is the one given by Def. IV.3.1.

LEMMA A.3.1:

Let $u(t) \in U_{\hat{D_I}}$, $\rho(t) \in \mathcal{D}$ and let $\Theta(t)$ be infinitely smooth, bounded with supp $\Theta \subset (b,\infty)$ for some finite b. Then

$$h(t) \triangleq \int_{\infty}^{\infty} \Theta(t).u(\tau).\ \rho(t+\tau) d\tau \qquad \text{A.3.1}$$

is a testing function in $\mathfrak{A}$ and

$$\int_{-\infty}^{\infty} |h(t)|^2 dt \leq K \int_{-\infty}^{\infty} |\rho(t)|^2 dt \qquad \text{A.3.2}$$

K a constant.

PROOF:

First we note that h(t) is well defined for each t since $\rho(t+\tau)$ has compact support for each t and $u(\tau)$ is locally integrable by fact A.3.1.

We have three things to be shown for h(t) to be in $\mathfrak{A}$.

(i) That h(t) is infinitely smooth; which is true since $\Theta(t)$ and $\rho(t+\tau)$ are infinitely smooth.

(ii) That $\int_{-\infty}^{\infty} |\eta^k h(t)|^2 dt < \infty$; which will be shown as follows. First we claim:

$\int_{\infty}^{\infty} |h(t)|^2 dt < \infty$. In fact:

$$\int_{-\infty}^{\infty} |h(t)|^2 dt = \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} \Theta(t).u(\tau)\rho(t+\tau)d\tau \right|^2 dt$$

$$= \int_{-\infty}^{\infty} |\Theta(t)|^2 \left| \int_{-\infty}^{\infty} u(z-t)\rho(z)dz \right|^2 dt \qquad \text{A.3.3}$$

by letting $z = t + \tau$.

As ρ has compact support, supp $\rho \subset [\alpha,\beta]$ with α,β finite

A.3.3 becomes: $\int_{-\infty}^{\infty}|h(t)|^2 dt = \int_{-\infty}^{\infty}|\Theta(t)|^2 \left|\int_{\alpha}^{\beta} u(z-t)\rho(z)dz\right|^2 dt.$

$u(z-t)$ is certainly square summable on $[\alpha,\beta]$ for each t.

Applying CBS inequality to $\left|\int_{\alpha}^{\beta} u(z-t)\rho(z)dz\right|^2$, we obtain

$$\int_{-\infty}^{\infty}|h(t)|^2 dt \leqslant \int_{-\infty}^{\infty}|\Theta(t)|^2 \left[\int_{\alpha}^{\beta}|u(z-t)|^2 dz\right]\left[\int_{\alpha}^{\beta}|\rho(z)|^2 dz\right]dt$$

$$\leqslant \int_{\alpha}^{\beta}|\rho(z)|^2 dz\left[\int_{-\infty}^{\infty}|\Theta(t)|^2\int_{\alpha}^{\beta}|u(z-t)|^2 dz\, dt\right] \qquad \text{A.3.4}$$

All we have left to be shown is the convergence of

$\left[\int_{-\infty}^{\infty}|\Theta(t)|^2 \int_{\alpha}^{\beta}|u(z-t)|^2 dz\, dt\right]$ in A.3.4 For this we consider

$$\int_{\alpha}^{\beta}\left[\int_{\infty}^{\infty}|\Theta(t)u(z-t)|^2 dt\right]dz = \int_{\alpha}^{\beta}\left[\int_{\infty}^{\infty}|\Theta(z-t)u(t)|^2 dt\right]dz \qquad \text{A.3.5}$$

By hypothesis $\Theta(z-t) = 0$ for $z-t<b$ i.e. for $z-b<t$. A.3.5 then gives:

$$\int_{\alpha}^{\beta}\left[\int_{\infty}^{\infty}|\Theta(t)u(z-t)|^2 dt\right]dz = \int_{\alpha}^{\beta}\left[\int_{\infty}^{z-b}|\Theta(z-t)u(t)|^2 dt\right]dz$$

$$\leq \int_{\alpha}^{\beta} \Big[\int_{-\infty}^{\beta-b} |\Theta(z-t)u(t)|^2 dt \Big] dz \qquad \text{since } z \leq b$$

$$\leq \int_{\alpha}^{\beta} \Big[\int_{-\infty}^{\beta} M|u(t)|^2 dt \Big] dz \qquad \text{since } \Theta \text{ is bounded everywhere}$$

$$\leq \int_{\alpha}^{\beta} C.dz = K < \infty \qquad \text{since } u(t) \in L^2_{(-\infty,b)} \text{ for any finite b.}$$

By [KE, pp. 206, Thm. 280],

$$\int_{\infty}^{\infty} |\Theta(t)|^2 \Big[\int_{\alpha}^{\beta} |u(z-t)|^2 dz \Big] dt = \int_{\alpha}^{\beta} \Big[\int_{\infty}^{\infty} |\Theta(t)u(z-t)|^2 dt \Big] dz < K \qquad \text{A.3.6}$$

and A.3.6 combined with A.3.4 gives A.3.2.

To show:

$$\int_{\infty}^{\infty} |\eta^k h(t)|^2 dt = \int_{\infty}^{\infty} |\eta^k \{\Theta(t) \int_{\infty}^{\infty} u(\tau)\rho(t+\tau)d\tau\}|^2 dt < \infty \qquad \text{A.3.7}$$

note that $\eta^k h(t)$ is a finite sum of terms of the form $\gamma(t) \int_{\infty}^{\infty} u(\tau)\phi(t+\tau)d\tau$, where $\gamma(t)$ has as factor either $\Theta(t)$ or one of its derivatives, multiplied by some Θ_k or its derivative of some order. Therefore $\gamma(t)$ is infinitely smooth, bounded due to Note. A.3.1 with supp $\gamma(t) \subset (b,\infty)$, since supp $\Theta(t) \subset (b,\infty)$.

$\phi(t)$ is a testing function in since $\phi(t) = \rho^{(i)}(t)$ for some integer $i \geq 0$. Thus $\gamma(t) \int_{\infty}^{\infty} u(\tau)\phi(t+\tau)d\tau$ satisfy the hypothesis of the lemma and the proof we used to show $\int_{-\infty}^{\infty}|h(t)|^2 dt < \infty$ can be applied to each term of $\eta^k h(t)$ to obtain finally $\int_{-\infty}^{\infty}|\eta^k h(t)|^2 dt < \infty$, since a finite sum of square summable terms is square summable.

(iii) Finally that $\langle \eta^k h, \psi_n \rangle = \langle h, \eta^k \psi_n \rangle$ is shown as follows:

$$\langle h, \eta\psi_n \rangle \triangleq \int_{-\infty}^{\infty} h(t)\, \overline{\eta\ \psi_n(t)}.dt$$

$$= \int_{-\infty}^{\infty} h(t)\ \overline{[\Theta_o(t) D^{n_1}\Theta_1(t) D^{n_2} \ldots D^{n_\nu}\Theta_\nu(t)\psi_n(t)]}dt \qquad \text{A.3.8}$$

Since the Θ_k's and ψ_η are functions of the real variable t, A.3.8 can be written as:

$$\langle h, \eta\ \psi_n \rangle = \int_{-\infty}^{\infty} h(t).\ \overline{\Theta_o(t)}\ D[D^{n_1-1}\overline{\Theta}_1 D^{n_2}\Theta_2 \ldots D^{n_\nu}\overline{\Theta_\nu}.\overline{\psi_n}]dt$$

A.3.9

To integrate by parts we let $v(t) = h(t)\overline{\Theta_o(t)}$ and $du = D[D^{n_1-1} \ldots D^{n_\nu}\ \overline{\Theta_\nu\ \psi_n}]$ dt in A.3.9. To get

$$<h,\eta\psi_n> = h(t)\overline{\Theta_o(t)}[D^{n_1-1}\overline{\Theta_1}D^{n_2}\ldots D^{n_\nu}\overline{\Theta_\nu(t)}\psi_n(t)]\Big|_{-\infty}^{\infty}$$

$$- \int_{-\infty}^{\infty} D[\overline{\Theta_o(t)}h(t)]D^{n_1-1}[\overline{\Theta_1}D^{n_2}\ldots D^{n_\nu}\overline{\Theta_\nu(t)\psi_n(t)}]dt \qquad \text{A.3.10}$$

But $h(\infty) = 0$ by FACT A.3.3 since $h(t)$ is square summable and infinitely smooth, $h(-\infty) = 0$, since $h(t)$ has support bounded at left. So the first expression on the right side of A.3.10 is zero and:

$$<h,\eta\psi_n> = \int_{-\infty}^{\infty}(-D)\ [\overline{\Theta_o(t)}h(t)]D^{n_1-1}[\overline{\Theta_1}D^{n_2}\ldots D^{n_\nu}\overline{\Theta_\nu(t)\psi_n(t)}]dt$$

A.3.11

As in (ii) $-D[\overline{\Theta_o(t)}h(t)]$ is composed of two terms each of which, satisfying the hypothesis of the present lemma, is square summable making $-D[\overline{\Theta_o(t)}h(t)]$ square summable. Thus integration by parts can be used for A.3.11 again, with the same reasoning as for A.3.9, to yield:

$$<h,\eta\psi_n> = \int_{-\infty}^{\infty}(-D)^2[\overline{\Theta_o(t)}h(t)]D^{n_1-2}[\overline{\Theta_1}D^{n_2}\ldots D^{n_\nu}\overline{\Theta_\nu(t)\psi_n(t)}]dt$$

With exactly the same arguments, repeating this process $n_1 + n_2 + \ldots + n_\nu$ times we will end up with:

$$<h,\eta\psi_n> = \int_{-\infty}^{\infty} \psi_n(t)[\overline{\Theta_\nu}(-D)^{n_\nu}\ldots(-D)^{n_2}\overline{\Theta_\nu}(-D)^{n_1}\overline{\Theta_o(t)}h(t)]dt$$

$$= <\eta h,\ \psi_n>$$

due to the assumed form for η, Not. A.2.1. In order to show $<h,\eta^k\psi_n> = <\eta^k h,\psi_n>$ we note that the operator η^k has the same form as η i.e. $\eta^k = [\Theta_o D^{n_1}\ldots D^{n_\nu}\Theta_\nu]\ldots [\Theta_o D^{n_1}\ldots D^{n_\nu}\Theta_\nu]$ where the bracketed term occurred k times in succession. The integration by parts can be repeated as many times as we want yielding $<\eta^k h,\psi_n> = <h,\eta^k\psi_n>$ for finite k.

LEMMA A.3.2

Let $u(t)\in U_{D(-\infty,\infty)}$ and $\Theta(t)$ be as in Lemma A.3.1, and let $\{\rho_\gamma(t)\}_{\gamma=1}^{\infty}\subset\mathfrak{D}$ converge to zero in $\mathfrak{D}$. Then

$$\{h_\gamma(t) \triangleq \int_{-\infty}^{\infty}\Theta(t)u(\tau)\rho_\gamma(t+\tau)d\tau\}_{\gamma=1}^{\infty} \quad \text{converges to zero in } \mathfrak{A}.$$

PROOF:

From Lemma A.3.1 we have that:

$\int_{-\infty}^{\infty} |h_\gamma(t)|^2 dt \leqslant K \int_{-\infty}^{\infty}|\rho_\gamma(t)|^2 dt \quad \gamma=1,2,\ldots.$ We note that K is independent of γ, due to expressions A.3.4 and A.3.6, and due to the definition of convergence in $\mathfrak{D}$ (Def. A.2.2) which requires supp $\rho_\gamma\subset[\alpha,\beta]$ for $=1,2,\ldots.$ Thus

$$\lim_{\gamma\to\infty}\int_{\alpha}^{\beta}|\rho_\gamma(t)|^2 dt = 0, \text{ implying } \lim_{\gamma\to\infty}\int_{-\infty}^{\infty}|h_\gamma(t)|^2 dt = 0.$$

Again, $\eta^k h(t)$ is the sum of a finite number of terms each of the form $\gamma(t)\int_{-\infty}^{\infty} u(\tau)\phi_\gamma(t+\tau)d\tau$ with $\gamma(t)$ and $\phi_\gamma(t)$ as in Lemma A.3.1 for $\gamma=1,2,\ldots.$ Moreover

$$\int_{-\infty}^{\infty}|\gamma(t)|^2[\int_{-\infty}^{\infty}|u(\tau)\phi_\gamma(t+\tau)|^2d\tau]dt \leqslant C\int_{\alpha}^{\beta}|\rho_\gamma^{(i)}(t)|^2dt \text{ for}$$

some i and for $\gamma=1,2,\ldots.$ As the convergence of $\rho_\gamma(t)$'s is in $\mathcal{D}$, $\rho_\gamma^{(i)}$ converges to zero for any i. So

$$\lim_{\gamma\to\infty}\int_{\alpha}^{\beta}|\rho_\gamma^{(i)}(t)|^2dt = 0 \Longrightarrow$$

$$\lim_{\gamma\to\infty}\int_{-\infty}^{\infty}|\gamma(t)|^2[\int_{-\infty}^{\infty}|u(\tau)\phi_\gamma(t+\tau)|^2d\tau]dt = 0.$$ Using Minkowski's

inequality, as we have a finite number of terms we conclude

$$\lim_{\gamma\to\infty}\int_{-\infty}^{\infty}|\eta^k\{\Theta(t)\int_{-\infty}^{\infty}u(\tau)\rho_\gamma(t+\tau)d\tau\}|^2dt = 0 \quad \text{for each } k.$$

<u>NOTE A.3.3</u>: Now we define the convolution of a distribution in $\mathcal{O}'$ having support bounded from the left with $u(t)\in U_{D(-\infty,\infty)}$ of Def. IV.3.1. We need the two previous lemmas to prove the outcome of the convolution to be in $\mathcal{D}'$. This definition coincides with the usual definition of convolution if supp $u(t)$ is bounded from the left.

<u>DEF.A.3.1</u>:

Let $w(t)\in\mathcal{O}'$ be such that supp $w(t)\subset[b,\infty]$, b finite, and let $u(t)\in U_{(-\infty,\infty)}$. Choose an infinitely smooth $\Theta(t)$

such that it equals one over some neighborhood of supp w(t) and zero outside this neighborhood. Finally let $\rho(t) \in \mathcal{D}$ be arbitrary. The CONVOLUTION of w(t) with u(t), denoted $w(t) * u(t)$ is defined by:

$$<w(t)*u(t),\ \rho(t)> \triangleq <w(t), \Theta(t) \int_{-\infty}^{\infty} u(\tau)\rho(t+\tau)d\tau> \qquad \text{A.3.12}$$

<u>THM A.3.1</u>:

Let w(t), u(t), Θ(t), ρ(t) be as in Def. A.3.1. Then $w(t)*u(t)$ as given by A.3.12 is well defined and is a distribution in $\mathcal{D}'$.

<u>PROOF</u>:

First we note that as Θ(t) is infinitely smooth $w(t)\Theta(t)$ is well defined and $w(t)\Theta(t) = w(t)$. Then by Lem. A.3.1 $\Theta(t)\int_{-\infty}^{\infty} u(\tau)\rho(t+\tau)d\tau$ is a testing function in $\mathfrak{A}$ and as $w(t) \in \mathfrak{A}'$, $<w(t), \Theta(t) \int_{-\infty}^{\infty} u(\tau)\rho(t+\tau)d\tau>$ is well defined. Moreover:

$$<w(t)*u(t), \rho(t)> \triangleq <w(t), \Theta(t) \int_{-\infty}^{\infty} u(\tau)\rho(t+\tau)dt>$$

$$= <w(t)\Theta(t), \int_{-\infty}^{\infty} u(\tau)\rho(t+\tau)d\tau>$$

$$= <w(t),\ <u(\tau), \rho(t+\tau)>> \qquad \text{A.3.13}$$

Since u(t) is locally summable it is imbedded in $\mathcal{D}'$. The expression A.3.13 verifies that A.3.12 is indeed a convolution where Θ(t) is necessary in making h(t) a function in $\mathfrak{A}$.

Since $\rho(t)\in\mathcal{D}$ was arbitrary, we will be done if we can show $w(t)*u(t)$ is linear and continuous on $\mathcal{D}$.

$w(t)*u(t)$ is linear, since for ρ_1 and $\rho_2\in\mathcal{D}$ and $a\in\mathbb{R}$ we have:

$$<w(t)*u(t),\ a\rho_1(t)+\rho_2(t)> \triangleq <w(t),\Theta(t)\int_{-\infty}^{\infty}u(\tau)[a\rho_1(t+\tau)+\rho_2(t+\tau)]d\tau>$$

$$= <w(t),\Theta(t)\int_{-\infty}^{\infty}u(\tau)a\rho_1(t+\tau)d\tau>$$

$$+ <w(t),\Theta(t)\int_{-\infty}^{\infty}u(\tau)\rho_2(t+\tau)d\tau>$$

$$= a <w(t)*u(t),\rho_1(t)> + <w(t)*u(t),\rho_2(t)>$$

$w(t)*u(t)$ is continuous on $\mathcal{D}$. If $\{\rho_\gamma(t)\}_{\gamma=1}^{\infty}$ is a zero convergent sequence in $\mathcal{D}$ then:

$<w(t)*u(t),\rho_\gamma(t)> = <w(t),\Theta(t)\int_{-\infty}^{\infty}u(\tau)\rho_\gamma(t+\tau)d\tau>$ converges to zero as $\gamma\to\infty$ since $\Theta(t)\int_{-\infty}^{\infty}u(\tau)\rho_\gamma(t+\tau)d\tau$ converges to zero in $\mathfrak{A}$ by Lemma A.3.2 and $w(t)\in\mathfrak{A}'$.

<u>NOTE A.3.4</u>: The next lemma, may be one that does not require a proof. Although it is not explicitly mentioned in [ZE2], it must be true for $\mathfrak{A}'$ to be a distribution space. Since we use it in our proofs we felt to prove it briefly would be adequate.

LEM.A.3.3:

$$\psi(t)\in\mathfrak{A} \Rightarrow \psi'(t)\in\mathfrak{A}$$

PROOF:

(i) $\psi'(t)$ is infinitely smooth.

(ii) $\int_{-\infty}^{\infty}|\psi'(t)|^2dt < \infty$ by Fact A.3.4. As in Lemma A.3.1 $\eta^k\psi'(t)$ is composed of a finite number of terms $\Theta(t)\phi(t)$ where $\Theta(t)$ is infinitely smooth with

$\phi(t) = [\psi'(t)]^{(n)}$ $n = 0,1,2,\ldots.$ Thus by Fact A.3.4 $\phi(t)$ and hence $\Theta(t)\phi(t)$ are square summable for every finite k. Again using Minkowski's inequality we can obtain $\int_{-\infty}^{\infty}|\eta\ \psi'(t)|^2dt < \infty$ $k=0,1,2,\ldots.$

(iii) To prove $<\eta^k\psi',\psi_n> = <\psi',\eta^k\psi_n>$ we can proceed exactly as we did in Lem. A.3.1, i.e. using integration by parts.

THM.A.3.2:

Let $w(t)$ be in $\mathfrak{A}'$ with supp $w(t)\subset[b,\infty]$ and let $u(t)\in U_{(-\infty,\infty)}$. Suppose $w(t) = \lim_{\gamma\to\infty} w_\gamma(t)$ in $\mathfrak{A}'$ also with supp $w_\gamma(t)\subset[b,\infty]$ for $\gamma = 1,2,\ldots,b$ finite. Then:

$$w(t)*u(t) = \lim_{\gamma\to\infty}[w_\gamma(t)*u(t)] \text{ in } \mathcal{D}'.$$

PROOF:

$w(t)*u(t)$ is well defined by Thm. A.3.1 and so is $w_\gamma(t)*u(t)$ for each γ. Then for $\rho(t)\in\mathcal{D}$ and an infinitely smooth $\Theta(t)$ which equals one over a neighborhood of supp $w(t)$ and zero outside we have:

$$<w(t)*u(t),\rho(t)> \triangleq <w(t),\Theta(t)\int_{-\infty}^{\infty}u(\tau)\rho(t+\tau)d\tau> \quad \text{by Def.A.3.1.}$$

$$= <\lim_{\gamma\to\infty} w_\gamma(t),\Theta(t)\int_{-\infty}^{\infty}u(\tau)\rho(t+\tau)d\tau>$$

$$= \lim_{\gamma\to\infty} <w_\gamma(t),\Theta(t)\int_{-\infty}^{\infty}u(\tau)\rho(t+\tau)d\tau> \quad \text{A.3.14}$$

Since $\Theta(t)\int_{-\infty}^{\infty}u(\tau)\rho(t+\tau)d\tau$ and by definition of convergence in $\mathcal{A}'$ i.e. $\lim_{\gamma\to\infty} w_\gamma(t) = w(t)$ in $\mathcal{A}'$ iff $\lim_{\gamma\to\infty} <w_\gamma(t),\phi> = <w(t),\phi(t)>$ $\forall\phi\in\mathcal{A}$. Thus A.3.14 gives:

$$<w(t)*u(t),\rho(t)> = \lim_{\gamma\to\infty} <w_\gamma(t)*u(t),\rho(t)> \qquad \forall\rho\in\mathcal{D}$$

$$= <\lim_{\gamma\to\infty} [w_\gamma(t)*u(t)],\rho(t)> \quad \forall\rho\in\mathcal{D}$$

THM.A.3.3:

Let $w(t)\in\mathcal{A}'$ with supp $w(t)\subset[b,\infty]$. Then $w(t)$ can be written as:

$$w(t) = \sum_{n=0}^{\infty} <w,\psi_n> \psi_n(t)1(t-b) + \sum_{k=0}^{K} d_k\delta^{(k)}(t-b) \quad \text{A.3.15}$$

where K is finite.

PROOF:

Since $w(t) \in \mathfrak{O}'$ we can write $w(t) = \sum_{n=0}^{\infty} \langle w, \psi_n \rangle \, \psi_n(t)$ by Thm. A.2.11. Then we define

$$f(t) \triangleq w(t) - \sum_{n=0}^{\infty} \langle w, \psi_n \rangle \psi_n(t) 1(t-b).$$ In order to prove the theorem all we have to show is: $f(t)$ can at most have its support concentrated at the point b.

Let $\rho(t)$ be any testing function with supp $\rho \subset (-\infty, b)$. We can easily write

$$\langle f(t), \rho(t) \rangle = \langle w(t), \rho(t) \rangle - \langle \sum_{n=0}^{\infty} \langle w, \psi_n \rangle \, \psi_n(t) 1(t-b), \rho(t) \rangle = 0$$

since both terms defining f have their support in $[b, \infty)$.

Let $f(t)$ have its support in (b, ∞). Then:

$$\langle f(t), \rho(t) \rangle = \langle w(t), \rho(t) \rangle - \langle \sum_{n=0}^{\infty} \langle w, \psi_n \rangle \psi_n(t) 1(t-b), \rho(t) \rangle$$

$$= \langle w(t), \rho(t) \rangle - \langle \sum_{n=0} \langle w, \psi_n \rangle \psi_n(t), \rho(t) \rangle = 0$$

Since $1(t) = 1$ and $1(t-b)$ is infinitely smooth on (b, ∞).

Thus $f(t) = 0$ on $(-\infty, b)$ and on (b, ∞) hence it is zero on the union of these open sets Thm. A.2.1. Therefore $f(t)$ has support concentrated to the origin. As only finite linear combinations of the delta functional and its derivatives are concentrated at a point, Thm. A.2.2,

$f(t) = \sum_{k=0}^{\infty} d_k . \delta^{(k)}(t)$ and A.3.15 follows.

<u>NOTE.A.3.5</u>: The last two theorems are of importance in section IV.3 when obtaining the state description of a large class of objects. Another result, exactly similar to Thm. A.3.3 is useful in section IV.4 and is stated in this note.

Let a sequence of infinitely smooth functions $\Theta_\gamma(t)$ converge in $\mathcal{D}'$ to the distribution $w(t) \in \mathcal{D}'$, $\mathcal{D}'$ defined on $(-\infty,\infty)$, with supp $w(t) \subset [b,\infty]$. Then

$$w(t) = \lim_{\gamma\to\infty} \Theta_\gamma(t)1(t) + \sum_{k=0}^{K} d_k \, \delta_k(t), \text{ K finite.} \qquad \text{A.3.16}$$

The proof is exactly in the lines of the proof of Thm. A.3.3.

APPENDIX B

HILBERT MATRICES

NOTE B.1: Although they are a natural extension of finite matrices, infinite matrices, i.e. matrices with infinite rows and columns, do not occupy much place in today's literature, possibly because they are preempted by the theory of abstract transformations and operators. A good book available on the subject is Cook's Infinite Matrices and Sequence Spaces [CO] written in 1950 from where stems the following short discussion.

As the theorems will show a Hilbert Matrix, a name that seems to be abandoned in general but for the matrix $[(p+q)^{-1}]_{pq}$, is nothing but a bounded operator on a sequence space. Bounded operators are important and much is known about them. Moreover as every linear operator on the Hilbert space ℓ^2 of square summable sequences can be written as an infinite matrix [PO, pp. 412, Ex.1], that makes the Hilbert Matrices important, especially if we encounter them in technical characterizations as we did in section IV.3.

DEF.B.1:

Let a double series $\sum_{m,n} C_{m,n}$ be given. We form the sequence $S_{p,q}$ of partial sums by finite rectangles, i.e. $S_{p,q}$ is obtained by adding all terms whose first index is $\leqslant p$ and whose second index $\leqslant q$. Then $\sum_{m,n} C_{m,n}$ is said to be PRINGSHEIM-CONVERGENT iff for every $\varepsilon>0$ there is a number s, independent of ε, and two numbers $P(\varepsilon)$ and $Q(\varepsilon)$ such that $p \geqslant P(\varepsilon)$, $q \geqslant Q(\varepsilon)$ implies $|S_{p,q}-s| \leqslant \varepsilon$. The number S is called the INNER (or PRINGSHEIM) LIMIT of the double sequence $S_{p,q}$.

DEF.B.2:

For an infinite matrix $A = [a_{mn}]$ a BILINEAR FORM is defined as $x^T A y \triangleq \sum_{m,n=1}^{\infty} x_m a_{mn} y_n$ where $x^T=(x_1,x_2,\ldots)$, $y^T = (y_1,y_2,\ldots)$ and the convergence is Pringsheim.

DEF.B.3:

Let E denote the unit hypersphere, i.e.

$$E \triangleq \{x = (x_1,x_2,\ldots): ||x|| \triangleq [\sum_{n=1}^{\infty} |x_n|^2]^{\frac{1}{2}} \leqslant 1\}.$$

An infinite matrix $A = [a_{mn}]$ is called a HILBERT MATRIX iff $x^T Ay$ is Pringsheim convergent on E.

THM.B.1: [CO, pp. 253, Cor. 2]

A necessary and sufficient condition that A should be a Hilbert matrix is that:

$\sum_{n=1}^{\infty} | \sum_{m=1}^{\infty} x_m \, a_{m,n} |^2 \leqslant M$ for every $x \in E$ (or that

$\sum_{m=1}^{\infty} | \sum_{n=1}^{\infty} a_{m,n} \, y_n |^2 \leqslant N$).

<u>THM.B.2</u>: [CO, pp. 260, Thm. 9.5.V]

A = $[a_{mn}]$ is a Hilbert matrix if $\sum_{m=1}^{\infty} |a_{mn}| \leqslant M$ independent of n and if $\sum_{n=1}^{\infty} |a_{mn}| \leqslant N$ independent of m.

<u>NOTE B.2</u>: Thm.B.1 shows that a Hilbert Matrix is a bounded operator on ℓ^2 and THM.B.2 is the one we use to show that the infinite matrix A in IV.3.1 is a Hilbert Matrix.

That Hilbert matrices are not compact operators (bounded linear operators that map bounded sets into relatively compact sets) is easily seen since the identity matrix I is a Hilbert matrix but not a compact operator since it maps the unit hypersphere, whose closure is not compact, into itself.