INSTITUTUL
DE
MATEMATICA

INSTITUTUL NATIONAL
PENTRU CREATIE
STIINTIFICA SI TEHNICA

ISSN 0250 3638

ON THE NOTION OF COMPLETENESS IN PREDICTION THEORY

by

I.Suciu and D.Timotin

PREPRINT SERIES IN MATHEMATICS
No.93/1981

BUCURESTI

ON THE NOTION OF COMPLETENESS IN PREDICTION THEORY

by

I.Suciu and D.Timotin*)

October 1981

*) Department of Mathematics, National Institute for Scientific and Technical Creation, Bd.Pacii 220, 79622 Bucharest, Romania.

ON THE NOTION OF COMPLETENESS IN PREDICTION THEORY[1]

by

I.Suciu and D.Timotin

The Wiener-Masani approach of prediction theory for finite multivariate stationary processes (Wiener and Masani (1957, 1958)) contains many ideas and constructions which have generated the later developments in prediction theory of infinite variate stationary processes. One of them is the idea of considering the time domain of the processes as a linear space $\mathcal{H}$ on which the C^*-algebra M_q of q x q matrices acts. The present and past of the processes is then defined as the collection of all linear combinations, with coefficients in M_q, of the states of the process up to the moment t=0. The correlations are given by a map Γ from $\mathcal{H} \times \mathcal{H}$ into M_q satisfying some conditions which make Γ an M_q-valued "scalar product" on $\mathcal{H}$. In the Wiener-Masani scheme $\mathcal{H}$ is the linear space of all random norm square integrable vectors $f(\omega) = (f_1(\omega),\dots,f_q(\omega))$ on a probability space $\{\Omega, \mathcal{K}, P\}$. The matrices from M_q act pointwise on $\mathcal{H}$, and the correlation map Γ is given by the Gramian matrix

$$\Gamma[f,g]_{ij} = \int_\Omega f_i(\omega)\overline{g_j(\omega)}\,dP(\omega), \qquad f,g \in \mathcal{H},\ 1 \le i,j \le q$$

This approach is based on the assumption that we are able to identify a finite number q of parameters of the phenomenon under study. The experiences we make in order to get information about it can be described by qxq matrices. (The results of the experiences yield only the correlation matrix $\Gamma[Af,Bg]$; these variables appear after performing the experiences A and B when the phenomenon is in the given states f and g).

But it is generally difficult (and sometimes impossible, e.g. in quantum theory) to identify a finite number of satisfactory parameters. Knowing the time evolution up to a moment t=0 of a finite number of parameters is sometimes not sufficient even to obtain relevant information about their own later evolution. The problem of considering an infinite (or at least not specificate) number of parameters presents therefore not only mathematical interest.

Following the ideas of Wiener-Masani roughly described above, a mathematical model for prediction theory of infinite variate stationary processes was proposed by Suciu and Valuşescu (1978, 1979). In this model the time domain of the process is described by the state space of a correlated action $\{\mathcal{E},\mathcal{K},\Gamma\}$, where the <u>parameter space</u> $\mathcal{E}$ is a (possibly infinite dimensional) Hilbert space, the <u>state space</u> $\mathcal{K}$ is a right module over the C^*-algebra $\mathcal{L}(\mathcal{E})$ of all linear bounded operators on $\mathcal{E}$ and the correlation map Γ takes values in $\mathcal{L}(\mathcal{E})$. In this context it was possible to formulate with a sufficient degree of consistency, the notions and problems of prediction theory. The spectral (or frequency-domain) model which can be attached in a natural way offers the possibility to obtain relevant results in prediction theory by using recent developments of the structure theory for operators on Hilbert spaces.

In this paper we shall discuss the problem of completness of correlated actions. This problem has been simplified at maximum in Suciu and Valuşescu (1978); nevertheless it seems to be important from the theoretical as well as from the practical point of view.

1. Best estimation in the state space.

Let $\mathcal{E}$ be a complex Hilbert space and denote by $\mathcal{L}(\mathcal{E})$ the C^*-algebra of all linear bounded operators on $\mathcal{E}$. A unital right $\mathcal{L}(\mathcal{E})$-module is a vector space $\mathcal{H}$ together with an action $(A,h)\mapsto Ah\ (h\in\mathcal{H},\ A\in\mathcal{L}(\mathcal{E}))$ of $\mathcal{L}(\mathcal{E})$ on $\mathcal{H}$ which satisfies

A_1. $Ih=h$, $h\in\mathcal{H}$, I=the identity operator on $\mathcal{E}$.

A_2. $A(h+g)=Ah+Ag$, $h,g\in\mathcal{H}$, $A\in\mathcal{L}(\mathcal{E})$

A_3. $(A+B)h=Ah+Bh$, $h\in\mathcal{H}$, $A,B\in\mathcal{L}(\mathcal{E})$

A_4. $(AB)h=B(Ah)$, $h\in\mathcal{H}$, $A,B\in\mathcal{L}(\mathcal{E})$

An $\mathcal{L}(\mathcal{E})$-correlation on $\mathcal{H}$ is a map Γ from $\mathcal{H}\times\mathcal{H}$ into $\mathcal{L}(\mathcal{E})$ satisfying

Γ_1. $\Gamma[h,h]\geq 0$, $\Gamma[h,h]=0 \Rightarrow h=0$, $h\in\mathcal{H}$

Γ_2. $\Gamma[h,g]=\Gamma[g,h]^*$, $h,g\in\mathcal{H}$

Γ_3. $\Gamma\left[\sum_{j=1}^{n} A_jh_j,\ \sum_{k=1}^{p} B_kg_h\right]=\sum_{j=1}^{n}\sum_{k=1}^{p} A_j^*\Gamma[h_j,g_k]B_k$

A correlated action is a triple $\{\mathcal{E},\mathcal{H},\Gamma\}$ where $\mathcal{E}$ -the parameter space is a Hilbert space, $\mathcal{H}$ -the state space - is a unital right $\mathcal{L}(\mathcal{E})$-module and Γ -the correlation - is an $\mathcal{L}(\mathcal{E})$-valued correlation on $\mathcal{H}$.

For any correlated action $\{\mathcal{E},\mathcal{H},\Gamma\}$ we can construct the Hilbert space $\mathcal{K}$ uniquely determined (up to adequate isomorphism) by the following property:

There exists an algebraic embedding $h \mapsto V_h$ of the right $\mathcal{L}(\mathcal{E})$-module $\mathcal{H}$ into the right $\mathcal{L}(\mathcal{E})$-module $\mathcal{L}(\mathcal{E},K)$ satisfying

M_1. $\Gamma[h,g]=V_h^* V_g$, $h,g \in \mathcal{H}$

M_2. The elements $\gamma_{a,h}=V_h a$ span a dense subspace in $\mathcal{K}$ when a runs over $\mathcal{E}$ and h runs over $\mathcal{H}$.

The construction of $\mathcal{K}$ follows the construction of the Aronszajn reproducing kernel Hilbert space, starting from the operatorial kernel Γ . Recall only that the scalar product on the generators of $\mathcal{K}$ has the form

$$(2.1) \qquad (\gamma_{a,h}, \gamma_{b,g})_{\mathcal{K}} = (\Gamma[g,h]a,b)_{\mathcal{E}}$$

We shall call $\mathcal{K}$ the measuring space of the correlated action $\{\mathcal{E},\mathcal{H},\Gamma\}$.

The basic problem in all prediction or filtering theory is to estimate a certain, desired, behaviour of the phenomenon under study using the information already obtained about it. We may wish to know the behaviour of the phenomenon at the next moment from the knowledge of its behaviour up to the present moment (prediction), or to prescribe to the phenomenon a given behaviour (filtering). And, of course, we look for the best estimation. We shall try to define in our context the notions of known behaviour, estimating and best estimation.

First we shall accept some extramathematical conventions. We obtain informations about the behaviour of the phenomenon by some received message. We accept that any received message describes a <u>possible state</u> of the phenomenon. So the received message describes a certain state $h \in \mathcal{H}$ which we call <u>prepared state</u>.

We assume also that for any finite system $h_1,\dots,h_n$ of prepared states in $\mathcal{H}$ and $A_1,\dots,A_n \in \mathcal{L}(\mathcal{E})$, the state $h=\sum_{k=1}^{n} A_k h_k$ is also prepared. This assumption has not only mathematical but also technical support; it means that we can perform a finite number of experiences on the received states in order to prepare a new possible received state.

Consider the information we can extract from a submodule M of prepared states from the mathematical point of view. Suppose we have a received state h; what we can measure are the numerical values $(\Gamma[h,g]a,b)_{\mathcal{E}}$ for any possible received state $g \in \mathcal{H}$ and a,b in $\mathcal{E}$. These values determine the vector $V_h a = \gamma_{a,h}$ for any $a \in \mathcal{E}$. So together with the submodule M in $\mathcal{H}$ we have determined the subspace $\mathcal{K}_M = \bigvee_{h \in M} V_h \mathcal{E}$ of $\mathcal{K}$. We shall denote by P_M the orthogonal projection from $\mathcal{K}$ onto $\mathcal{K}_M$. Now let f be an arbitrary state in $\mathcal{H}$ and h a known state in M. Since all the measured values which appear together with the possible received state of $\mathcal{H}$ have the form

$$(\Gamma[f,g]a,b)_{\mathcal{E}} = (\gamma_{a,f}, \gamma_{b,g})_{\mathcal{K}}$$

it follows that if we use h instead of f the error we expect to appear in the measuring process is

$$|(\gamma_{a,f}, \gamma_{b,g}) - (\gamma_{a,h}, \gamma_{b,g})| = |(\gamma_{a,f} - \gamma_{a,h}, \gamma_{b,g})| \leq$$

$$\|\gamma_{a,f} - \gamma_{a,h}\| \cdot \|\gamma_{b,g}\|$$

We may take $\|\gamma_{a,f} - \gamma_{a,h}\|$ as a numerical measure of the error. But

$$\|\gamma_{a,f} - \gamma_{a,h}\| \geq \|\gamma_{a,f} - P_M \gamma_{a,f}\| = \|(I-P_M)\gamma_{a,f}\|$$

If we denote

(*) $$(\Delta^2_{f,M} a,a) = \|(I-P_M)\gamma_{a,f}\|^2$$

then

$$(\Delta^2_{f,M} a,a) = \|(I-P_M)V_f a\|^2_{\mathcal{K}} = (V_f^*(I-P_M)V_f a,a)_{\mathcal{E}}$$

It follows that (*) defines a positive operator $\Delta^2_{f,M}$ on .
We shall call $\Delta_{f,M}$ the _error operator_ of estimation of f by elements from M.

Proposition. _We have_

$$\Delta^2_{f,M} = \inf_{h \in M} \Gamma[f-h, f-h]$$

where the infimum is taken in the partially ordered set of positive operators on $\mathcal{L}(\mathcal{E})$.

Proof. From (*) it follows that for any $a \in \mathcal{E}$ we have

$$(\Delta^2_f a,a) = \|(I-P_M)\gamma_{a,f}\|^2 = \inf_{k \in \mathcal{K}_M} \|\gamma_{a,f} - k\|^2 =$$

$$= \inf_{\substack{h_1,\dots,h_n \in M \\ a_1,\dots,a_n \in \mathcal{E}}} \|\gamma_{a,f} - \sum_{k=1}^{n} \gamma_{a_k,h_k}\|^2 = \inf_{\substack{h_0=f,\ h_1,\dots,h_n \in M \\ a_0=a,\ a_1,\dots,a_n \in \mathcal{E}}} \|\sum_{k=0}^{n} \gamma_{a_k,h_k}\|^2$$

Hence for any $h \in M$ and $a \in \mathcal{E}$ we have

$$(\Gamma[f-h,f-h]a,a) = \|\gamma_{a,f-h}\|^2 \geq (\Delta^2_{f,M} a,a)$$

Let Q be a positive operator in $\mathcal{L}(\mathcal{E})$ such that $Q \leq \Gamma[f-h,f-h]$ for any $h \in M$. Then

$$(Qa,a) \leq \inf_{h\in M} (\Gamma[f-h,f-h]a,a) = \inf_{h\in M} \| \gamma_{a,f-h} \|^2 =$$

$$= \inf_{h\in M} \| \gamma_{a,f} - \gamma_{a,h} \|^2 = \| (I-P_M)\gamma_{a,f} \|^2 = (\Delta^2_{f,M}a,a) .$$

The following definition is now natural.

<u>Definition</u>. We say that the state $\hat{f}$ in $\mathcal{K}$ is <u>the best estimation</u> by states in M of the state f in $\mathcal{K}$ if

1) $\gamma_{a,\hat{f}} \in \mathcal{K}_M$, $a \in \mathcal{E}$

2) $\Gamma[f-\hat{f},f-\hat{f}] = \Delta^2_{f,M}$

In this case relation (*) implies $V_{\hat{f}} = P_M V_f$; so the best estimation, if it exists, is unique. But the existence of the best estimation is not always assured. Moreover, even in the case when it exists it may not belong to M. That is, it can not be prepared performing a finite number of experiences on the received states.

Now it becomes obvious that we need a notion of completeness of the corelated action $\{\mathcal{E},\mathcal{K},\Gamma\}$ which should assure the existence of the best estimation. Moreover, it should provide an approximating procedure (preferable recursive) to approach f by elements in M, in the sense that the errors which appear at each step should tend, in some sense, to the minimum error operator $\Delta_{f,M}$.

2. Some ideas for the axiom of completeness

Before stating some possible forms of completeness we have to be more precise concerning what we have to ask for such an axiom; that is, what are the main properties and results concerning prediction theory that we should like to obtain as corrolaries of the proposed axiom of completeness. We have already stated such a requirement at the end of the preceding paragraph, concerning the existence of the best estimation. As a support for further requirements we shall quote some remarks made by N.Wiener: "Important as is the method of prediction given in this paper it has strict limitations in practice (this is, in fact, true of any method of prediction) and should never be used to determine a curve which may be determined in a stricly geometrical manner. Statistical prediction is essentially a method of refining a prediction which would be perfect by itself in an idealised case but which is corrupted by statistical errors, either in the observed quantity itself or in the observation. Geometrical facts must be predicted geometrically and analytical facts analytically, leaving only statistical facts to be predicted statistically" (Cf. Wiener (1950), pages 70-71).

We may derive the fact that Wiener's methods in prediction and filtering are based on the assumption that the phenomenon and the measuring system contain sufficiently many random "corruption"; by comparing different results measured in sufficiently many different experiences we must be able to sharpen the form of "noises". Then the possible obtainable information is what remains from the measured data after removing these noises. So the notion of completeness of a correlated action must be related also to the abundance of "noises" in data.

To be more precise we recall some fundamental facts from the prediction theory of discrete stationary processes. A stationary process in the correlated action $\{\mathcal{E}, \mathcal{K}, \Gamma\}$ is a doubly infinite sequence $\{f_n\}_{n=-\infty}^{\infty}$, $f_n \in \mathcal{K}$, such that $\Gamma[f_n, f_{n+k}]$ depends only on k and not on n. The function $k \mapsto \Gamma_f[k] = \Gamma[f_n, f_{n+k}]$ is then a positive definite function on the group $\mathbb{Z}$ of the integers with values in $\mathcal{L}(\mathcal{E})$ - the so called <u>autocorrelation function</u> of the process. For the process $\{f_n\}$ we shall adopt the following notations.

$\mathcal{K}_f^n = \{h \in \mathcal{K}, h = \sum_k A_k f_k$, where $k \leq n$ and $f_k \neq 0$ only for a finite number of values of k $\}$

$$\mathcal{K}_f^n = \bigvee_{k=-\infty}^{n} V_{f_n} \mathcal{E}$$

$\Delta_f^n = \Delta_{f_{n+1}, \mathcal{K}_n}$ — the error estimation of f_{n+1} by elements from $\mathcal{K}_n$.

From (*) it follows that Δ_f^n does not depend on n; we put $\Delta_f^n = \Delta_f$ and call it the <u>prediction error operator</u> of the process $\{f_n\}$.

The shift operator of the process is the unitary operator U_f defined on $\mathcal{K}_f = \bigvee_{n=-\infty}^{\infty} V_{f_n} \mathcal{E}$ by

$$U_f\left(\sum_n V_{f_n} a_n\right) = \sum_n V_{f_{n+1}} a_n .$$

It is clear then that

$$V_{f_k} = U_f V_{f_{k-1}}$$

and, denoting $V_f = V_{f_o}$, we have

$$V_{f_n} = U_f^n V_f$$

$$\mathcal{K}_f = \bigvee_{n=-\infty}^{\infty} U_f^n V_f \mathcal{E}$$

Also from (*) it follows that, for any n,

$$\Delta_f = (V_f^* U_f^{*n} (I-P_h) U_f^n V_f)^{1/2}$$

where P_n denotes the orthogonal projection from $\mathcal{K}_f$ onto $\mathcal{K}_f^{n-1}$. Clearly

$$0 \le \Delta_f^2 \le V_f^* (I-P_n) V_f \le V_f^* V_f = \Gamma[f_o, f_o]$$

The process $\{f_n\}$ is called <u>deterministic</u> if $\Delta_f = 0$. In this case $\mathcal{K}_f = \mathcal{K}_f^n$ for any $n \in Z$. (Note that this is stronger than the definition given in Suciu and Valuşescu (1978); see the corollary to the proposition below.)

We say that $\{f_n\}$ is a <u>white noise</u> if $\Delta_f = \Gamma[f_o, f_o]$. It is easy to check that $\{f_n\}$ is a white noise process if and only if $\Gamma[f_n, f_m] = \delta_{nm} \Gamma[f_o, f_o]$.

We say that the processes $\{f_n\}$ and $\{g_n\}$ are <u>stationary cross-corelated</u> if the function $\Gamma_{f,g}[k] = \Gamma[f_n, g_{n+k}]$ depends only on k and not on n. This is equivalent to the existence on the space $\mathcal{K}_{f,g} = \mathcal{K}_f \vee \mathcal{K}_g$ of a unitary operator $U_{f,g}$ such that $U_{f,g}|_{\mathcal{K}_f} = U_f$, $U_{f,g}|_{\mathcal{K}_g} = U_g$. We say that the white noise $\{g_n\}$ is <u>contained</u> in the process $\{f_n\}$ provided

1) $\{f_n\}$, $\{g_n\}$ are stationary cross-correlated.

2) $\Gamma[f_n, g_m] = 0$ for $m > n$

3) $\mathrm{Re}\, \Gamma[f_n - g_n, g_n] \ge 0$ for $n \in Z$

4) $V_g \mathcal{E} \subset \mathcal{K}_f^o$

Proposition. If $\{f_n\}$ contains $\{g_n\}$ then $\Delta_g \leq \Delta_f$.

Proof. We have to prove that, for any $h \in \mathcal{H}_f^o$

$$\Gamma[g_1, g_1] \leq \Gamma[f_1 - h, f_1 - h]$$

But,

$$\Gamma[f_1 - h, f_1 - h] = \Gamma[f_1 - (h+g_1) + g_1 \ , \ f_1 - (h+g_1) + g_1] =$$

$$= \Gamma[f_1 - (h+g_1), f_1 - (h_1 g_1)] + \Gamma[g_1, g_1] + 2\mathrm{Re}\ \Gamma[f_1 - (h+g_1), g_1]$$

But, since $\Gamma[h, g_1] = 0$ for any $h \in \mathcal{H}_f^o$, we have $\mathrm{Re}\ \Gamma[f_1 - (h+g_1), g_1] =$ $= \mathrm{Re}\ \Gamma[f_1 - g_1, g_1] \geq 0$, so our relation follows.

Then

$$\Delta_g^2 = \Gamma[g_1, g_1] \leq \Delta_f^2$$

Corollary. If the process $\{f_n\}$ is deterministic then it does not contain any (nonzero) white noise.

The converse of the last assertion is not, in general, true, and it is natural to take it as a first requirement concerning an axiom of completeness:

C_1. If the process $\{f_n\}$ does not contain any (non zero) white noise then it is deterministic.

This requirement is in accordance with the passage from Wiener quoted above. The prediction problem for the process $\{f_n\}$ is not trivial only in the case $\{f_n\}$ contains a white noise.

If we accept that the possible relevant information about the process is obtained by removing the white noises contained in it, it is important to know if there exists a maximal white noise contained in the process (We say that the white noise $g=\{g_n\}$ is the maximal white noise contained in $f=\{f_n\}$ if, whenever $g'=\{g'_n\}$ is another white noise contained in f, g' is contained in g). So, another str onger requirement whould be:

C_2. Any stationary process $f=\{f_n\}$ contains a maximal white noise $g=\{g_n\}$, and $\Delta_g=\Delta_f$.

This is a laticial property of the set of white noises in $\mathcal{H}$ endowed with the partial order relation given by inclusion of white noises. We do not insist here on the laticial formulation of this property.

For a stationary process $\{f_n\}$ in consider now $G_n^f=\mathcal{K}_n^f-\mathcal{K}_{n-1}^f$. Clearly $G_n^f=U^nG_o^f$ and we call $G^f=G_o^f$ the innovation subspace of f_n. We can then formulate requirement

C'_2: If $\{f_n\}$ is a stationary process and G^f is the innovation subspace of $\{f_n\}$ then there exists g_n in $\mathcal{H}$ such that $V_{g_n}=U^nP_{G^f}V_f$.

Proposition. C_2 is equivalent to C'_2.

Proof. Suppose $\{g_n\}$ is a white noise contained in $\{f_n\}$. It is obvious that $V_{g_o}\mathcal{E}\subset G^f$. Let $a\in\mathcal{E}$. Then

$$(\Gamma[f_o,g_o]a,a)=(V_{g_o}a,V_{f_o}a)=(V_{g_o}a,P_{G^f}V_{f_o}a)$$

This already implies that, if C'_2 is satisfied, it gives us a maximal white noise. Since

$$\|P_G V_{f_o}a\|^2=((I-P_n)V_{f_o}a,(I-P_n)V_{f_o}a)=(\Delta_f^2a,a)$$

the second requirement of C_2 is also fulfilled. Return now to an arbitrary white noise contained in $\{f_n\}$. We have

$$\|V_{g_o} a\|^2 = (\Gamma[g_o, g_o] a, a) = (\Delta_g^2 a, a)$$

Condition 3) from the definition of a white noise contained in f says that

$$\mathrm{Re}(\Gamma[f_o, g_o] a, a) \geq (\Gamma[g_o, g_o] a, a)$$

therefore

$$(**) \qquad \mathrm{Re}(V_{g_o} a, P_{G^f} V_{f_o} a) \geq \|V_{g_o} a\|^2$$

If C_2 is satisfied, then, taking in the last relation for g the maximal white noise yielded by C_2, we see that, since $\Delta_f = \Delta_g$ implies $\|V_{g_o} a\| = \|P_G f V_{f_o} a\|$, the Schwarz inequality applied to (**) shows that we must actually have $V_{g_o} a = P_{G^f} V_{f_o} a$, so C_2' follows.

Let us remark that C_2' insures the existence of $\hat{f}_n = f_n - g_n$ which is easily seen to be the best estimation of f_n by elements from $\mathcal{H}_f^{n-1}$.

Finally, we may ask that the maximal white noise produced by requirement C_2 should allow us to separate completely the "deterministic part" of the process from the part "corrupted by noises".

C_3. If $\{f_n\}$ is a stationary process, and $G_\infty^f = \bigvee_{n=-\infty}^{\infty} G_n^f$ there exists v_n in $\mathcal{H}$ such that $V_{v_n} = P_{G_\infty^f} V_{f_n}$.

In this case one can show easily that $\{v_n\}$ is a stationary process cross-corelated with $\{f_n\}$; the maximal white noise $\{g_n\}$ contained in $\{v_n\}$ has the property that

$\mathcal{K}_g = \mathcal{K}_v$. Moreover, if $u_n = f_n - v_n$, then $\{u_n\}$ is deterministic. All these facts form what can be called a "Wold decomposition" for stationary processes; in operator theory language they correspond to the usual Wold-von Neumann decomposition of the isometric operator $U_f^*|_{\mathcal{K}_f^o}$.

All requirements C_1-C_3 are satisfied in the case treated by Suciu and Valuşescu (1978); that is, if we ask simply that $\mathcal{H} = \mathcal{L}(\mathcal{E}, \mathcal{K})$. This seems, however, to be too strong a supposition, and is not sufficiently motivated by extramathematical arguments. As we shall see below, there are natural examples which satisfy C_1-C_3 , but do not verify $\mathcal{H} = \mathcal{L}(\mathcal{E}, \mathcal{K})$.

In search of other possible axioms, recall first that $\mathcal{H}$ is embedded in a natural way in $\mathcal{L}(\mathcal{E}, \mathcal{K})$ where $\mathcal{K}$ is the measuring space. We may therefore consider completeness requirements given by usual uniform structures on $\mathcal{L}(\mathcal{E}, \mathcal{K})$.

1. $\mathcal{H}$ is complete with respect to the operatorial norm in $\mathcal{L}(\mathcal{E}, \mathcal{K})$.

This axiom can be easily formulated without any reference to $\mathcal{K}$, since $\|V_h\|^2 = \|V_h^* V_h\| = \|\Gamma[h,h]\|$. Therefore, $\|h\|_\Gamma = \|\Gamma[h,h]\|^{1/2}$ is a norm on $\mathcal{H}$.

But it does not assure neither one of requirements C_1-C_3 , nor the existence of the best estimation. It is used in the theory of Hilbert modules (Cf. Dupré and Fillmore (1980), Kasparov (1980)).

2. $\mathcal{H}$ is complete with respect to the strong operatorial topology in $\mathcal{L}(\mathcal{E}, \mathcal{K})$.

This axiom can also be formulated without any reference to $\mathcal{K}$ since $\|V_h a\|^2 = (V_h a, V_h a) = (V_h^* V_h a, a) = (\Gamma[h,h]a, a)$. But it is unfortunately too strong for our purposes. Indeed, it can be shown that such a requirement for $\mathcal{H}$ already implies

<u>Proposition</u>. If $\mathcal{H}$ is complete with respect to the strong operatorial topology, then $\mathcal{H}=\mathcal{L}(\mathcal{E},\mathcal{K})$.

<u>Proof</u>. Let $T\in\mathcal{L}(\mathcal{E},\mathcal{K})$. We will show that for any finite dimensional subspace $\mathcal{E}_o\in\mathcal{E}$ and any $\varepsilon>0$, there is $h_o\in\mathcal{H}$, such that $\| V_{h_o}|_{\mathcal{E}_o}-T|_{\mathcal{E}_o}\|<\varepsilon$.

Suppose $\mathcal{E}_o\subset\mathcal{E}$ is a finite dimensional subspace. Since the linear span of $V_h a$, $h\in\mathcal{H}$, $a\in\mathcal{E}$ is dense in $\mathcal{K}$, we may find $T'\in\mathcal{L}(\mathcal{E},\mathcal{K})$, $\|T'|_{\mathcal{E}_o}-T|_{\mathcal{E}_o}\|<\varepsilon$ and $T'(\mathcal{E}_o)$ consists only of vectors of the form $V_h a$. Now, let $\{e_k\}_{k=1}^n$ be an orthonormal basis in $\mathcal{E}_o$.

Choose $h_i^{(k)}\in\mathcal{H}$, $f_i^{(k)}\in\mathcal{E}$, $i=1,\ldots,n_k$, such that $T'(e_k)=\sum_{i=1}^{n_k} V_{h_i^{(k)}} f_i^{(k)}$ and define $A_i^{(k)}\in\mathcal{L}(\mathcal{E})$ by $A_i^{(k)}(e_k)=f_i^{(k)}$ and $A_i^{(k)}=0$ on the orthogonal of e_k.

Then

$$V_{\sum_{k=1}^{n}\sum_{i=1}^{n_k} A_i^{(k)} h_i^{(k)}}\, e_m=\sum_{k=1}^{n}\sum_{i=1}^{n_k} V_{h_i^{(k)}} A_i^{(k)} e_m=$$

$$=\sum_{i=1}^{n_m} V_{h_i^{(m)}} f_i^{(m)}=T'(e_m)$$

so, if $h_o=\sum_{k=1}^{n}\sum_{i=1}^{n_k} A_i^{(k)} h_i^{(k)}$ then $V_{h_o}\big|_{\mathcal{E}_o}=T'\big|_{\mathcal{E}_o}$, and the proof is finished.

Moreover, any submodule M of closed in this topology is of the form $\mathcal{L}(\mathcal{E},\mathcal{K}_o)$ where $\mathcal{K}_o$ is a closed subspace of .

What is more important, one can exhibit natural examples which do not satisfy axioms 1 and 2, but which behave perfectly well from the point of view of prediction theory, that is, they have all properties C_1-C_3. These examples are the Schatten - von Neumann classes $\mathcal{C}_p$ of compact operators in $\mathcal{L}(\mathcal{E},\mathcal{K})$, equipped with the natural correlated action induced by $\mathcal{L}(\mathcal{E},\mathcal{K})$. If we take all the compact operators they satisfy axiom 1, but not axiom 2; any other ideal does not satisfy neither 1 nor 2. But we can easily prove the existence of best estimators, the existence of a maximal white noise contained in any stationary process, as well as a usual Wold decomposition theorem. In fact, we may perform all the usual constructions of Suciu and Valuşescu (1978) and check that they do not lead us outside the class of operators under consideration ; the last assertion is a consequence of the ideal property. This fact suggests that a completeness axiom for a correlated action should be more specifically connected with the prediction theory of stationary processes. It should allow us to obtain properties C_1-C_3 and it should cover a larger class of examples than just $\mathcal{L}(\mathcal{E},\mathcal{K})$.

We shall end by proposing such an axiom, which has the disadvantage that it is not directly formulated in terms of $\mathcal{H}$ but uses the measuring space $\mathcal{K}$. It is suggested by the proofs of C_1-C_3 in complete correlated actions in the sense of Suciu and Valuşescu (1978), and it is sufficient to make these proofs work in the general case.

3. If $h \in \mathcal{H}$, and P is an orthogonal projection in $\mathcal{L}(\mathcal{K})$, then there exists $h' \in \mathcal{H}$, such that $V_{h'} = PV_h$.

This axiom is satisfied in the case of operator ideals we have discussed; as we have already remarked, this fact is a consequence of the ideal property.

Concerning the construction of an approximation procedure, let us remark that approximation in the strong operator topology is always possible and that it can be given by a constructive procedure (see Timotin (1981)). Suciu and Valuşescu (1978) show that in a special case we may get a strong convergent series that gives the best estimation. However, it is not clear what are the conditions that could allow us to obtain some stronger form of convergence, maybe more related to the specific case of prediction theory. This problem is closely connected that of finding an appropiate completeness axiom; they both point to finding some sort of uniform structure on $\mathcal{H}$ that would be best suited for prediction purposes.

REFERENCES

[1] DUPRE, M.I. and FILLMORE, P.A. (1980. Triviality theorems for Hilbert modules, Topics in Modern Operator Theory, 5th Int.Conf. on Op.Th., Timişoara and Herculane (Romania), June 2-12, 1980, Birkhäuser Verlag, 71-79.

[2] KASPAROV, G.G. (1980). Hilbert C^*-modules: theorems of Stinespring and Voiculescu, J.of Op.Theory, 4, 133-150.

[3] SUCIU, I. and VALUSESCU, I. (1978). Factorization theorems and prediction theory, Rev.Roum.Math.Pures et Appl. 23, 1393-1423.

[4] SUCIU, I. and VALUSESCU, I. (1979). A linear filtering problem in complete correlated actions, J.of Multivariate Analysis, 9, 599-613.

[5] TIMOTIN, D. (1981). The Levinson algorithm in linear prediction, submitted to the Proceedings of the 6th Int. Conf. on Op.Th. Timişoara and Herculane.

[6] WIENER, N. (1950). The extrapolation, interpolation and smoothing of stationary time series, New York, 1950.

[7] WIENER, N. and MASANI, P. (1957). The prediction theory of multivariate stochastic processes, I., Acta Math.98, 111-150.

[8] WIENER, N. and MASANI, P. (1958). The prediction theory of multivariate stochastic processes, II, Acta Math. 99. 93-139.