NEUVIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 20 MAI 1983

APPLICATION OF OPERATOR THEORY TO SIGNAL MODELING AND PROCESSING

Leon H. Sibul and Guy R. L. Sohie

Applied Research Laboratory and Graduate Program in Acoustics
The Pennsylvania State University, University Park, PA 16802

RESUME

Dans cet article, nous discuterons de la modélisation en termes d'opérateurs, de processus stochastiques. Nous montrerons que le modèle MA (moyenne variable), lequel est le plus largement employé, le modèle AR (autorégressif) et le modèle ARMA (autorégressif à moyenne variable) peuvent s'exprimer sous la forme d'opérateurs linéaires. Un modèle théorique d'opérateur, plus général, permet la généralisation de ces modèles spécifiques. Partant des inégalités de norme de l'opérateur nous definissons des limites supèrieures du carré moyen de l'erreur dans le cas où le véritable modèle du signal est remplacé par un opérateur approximatif mais facilement inversible, qui génère le signal.

Puis nous ultiliserons une approche par opérateur théorique pour déterminer la structure d'un processeur espace temps, optimal afin d'isoler les signaux propagés à travers un milieu de transmission stochastique. On suppose que le signal est noyé dans un bruit coloré, anisotropique et variable dans le temps. La structure du processeur est exprimée à l'aide d'opérateurs qui représentent le milieu de transmission stochastique, le champ de bruit, les diffuseurs aléatoires, les generateurs de faisceaux ect. On applique le théorème de décomposition polaire des opérateurs linéaires bornés pour établir la possible factorisation due processeur espace temps. On utilise les concepts de transformation de contraction stochastique pour étudier la convergence d'algorithmes adaptatifs dont on se sert pour réaliser le processeur espace temps.

SUMMARY

In this paper, we will discuss modeling of stochastic processes in terms of operators. We will show that the widely used moving average (MA), autoregressive (AR), and autoregressive moving average (ARMA) models can be expressed in terms of linear operators. A more general operator-theoretic modeling allows generalization of these specific models. Based on operator norm inequalities, we derive upper bounds for the mean square (signal deconvolution) error for the case where the true signal model is replaced by an approximate but easily invertible operator which generates the signal.

Then we will use an operator-theoretic approach to derive the structure of an optimal space-time processor for extraction of signals that have propagated through a stochastic transmission medium. It is assumed that the signal is immersed in a colored, anisotropic, and time-varying noise. The processor structure is expressed in terms of operators that represent stochastic transmission medium, noise field, random scatterers, beamformers, etc. The polar decomposition theorem for bounded linear operators is applied to establish the factorability of the space-time processor. Stochastic contraction mapping concepts are used to study the covergence of adaptive algorithms that are used to realize the space time processor.

I. INTRODUCTION

It has been pointed out that an operator-theoretic formulation of signal processing and modeling problems provides a unifying framework for the signal detection, extraction, modeling, and data deconvolution problems [1,2,3]. The powerful machinery of functional analysis can be brought to bear on these problems. It is particularly convenient to formulate the signal processing problems as operator problems in a Hilbert space because of the geometrical interpretation of the Hilbert space, and because many answers to the signal processing problems can be recognized as inner products in a Hilbert space [1].

In this paper, we will discuss modeling of stochastic processes in terms of operators. We will show that the widely used moving average (MA), autoregressive (AR), and autoregressive moving average (ARMA) models can be expressed in terms of linear operators. Obviously, general operator-theoretic modeling allows generalization of these models. Next, we will address modeling of stochastic transmission media in terms of stochastic operators. Then we will use an operator-theoretic approach to derive the structure of an optimal space-time processor for extraction of signals that have propagated through a stochastic transmission medium. It is assumed that the signal is immersed in a colored, anisotropic, and time-varying noise. The processor structure is expressed in terms of operators that represent stochastic transmission medium, noise field, random scatterers, beamformers, etc. The polar decomposition theorem for bounded linear operators is applied to establish the factorability of the space-time processor. Stochastic contraction mapping concepts are used to study the convergence of adaptive algorithms that are used to realize the space-time processor.

In this paper we will consider finite variance complex stochastic processes $x(t,\omega)$, $y(t,\omega)$, $\underline{x}(t,\omega)$, $\underline{y}(t,\omega)$, $t \in T$ and $\omega \in \Omega$ define on the probability space (Ω,F,P). Lower case letters denote scalars, underlined lower case letters denote vectors, capital letters operators, and underlined capital letters matrix operators. Later, for notational convenience, we shall drop arguments ω and t if they are not needed for clarity. In this paper, we will consider stochastic processes and operators defined on complex ℓ_2 or L_2 Hilbert spaces with inner products

$$\langle \underline{x}(\omega), \underline{y}(\omega) \rangle_{\ell_2} = E\left\{\underline{x}^T(\omega)\ \underline{y}^*(\omega)\right\} \tag{1}$$

and

$$\langle x(t,\omega),\ y(t,\omega) \rangle_{L_2} = E\left\{\int_T x(t,\omega) y^*(t,\omega) dt\right\} \tag{2}$$

where T denotes transpose and * denotes complex conjugate. The norm of $x(t,\omega)$, denoted by $\|x(t,\omega)\|$, in the Hilbert space is by definition $\langle x(t,\omega),\ x(t,\omega)\rangle^{1/2}$ and the norm of a bounded linear A operator is $\sup\{\|Ax\|;\ \|x\| = 1\}$.

Since complex separable L_2 and ℓ_2 spaces are isomorphic and isometric [4], we can use either L_2 and ℓ_2 spaces at our convenience; that is, we can carry out analysis in either space and then use the results in the other space. In signal processing terms, the isomorphism between ℓ_2 and L_2 spaces means that discrete time signals are isomorphic with continuous time signals. We can use either integral operators or matrix operators for our analysis.

II. MODELING OF STOCHASTIC PROCESSES

Moving average (MA), autoregressive (AR), and autoregressive moving average (ARMA) models have been widely used to characterize stochastic processes [5,6]. These models have found wide application in linear prediction and system identification. In this section we will show that MA, AR, and ARMA processes can be defined by appropriate linear operators. This approach allows one to use the results from operator theory to study properties of these stochastic processes. In particular, operator theoretic inequalities can be used to determine a bound on the mean square error that occurs when an exact model for a stochastic process is replaced by an approximate but easily invertible model (widely used AR model, for example).

Let $\xi(t)$, $t = 0, \pm 1, \ldots$ be a sequence of uncorrelated random variables such that $E\{\xi(t)\} = 0$ and $E\{\xi^2(t)\} = \sigma^2$, then a finite moving average process is defined by*

$$x(t) = \sum_{j=-m}^{n} a_j \xi(t-j), \quad t = 0, \pm 1, \ldots \tag{3}$$

and if

$$\sum_{j=-\infty}^{\infty} a_j^2 < \infty \tag{4}$$

an infinite moving average process is defined by

$$x(t) = \sum_{j=-\infty}^{\infty} a_j \xi(t-j) . \tag{5}$$

Both MA processes can be defined by

$$x(t) = M\ \xi(t) = \sum_j a_j u^j\ \xi(t) \tag{6}$$

where u^j is a unitary (shift or delay) operator [7]. Since M is a bounded linear operator on a Hilbert space, polar decomposition theorem is applicable [8].

The finite autoregressive process of order q is defined by

$$\sum_{j=0}^{q} b_j\ x(t-j) = \xi(t) \quad t = 0, \pm 1, \ldots \tag{7}$$

and infinite autoregressive process is defined by

$$\sum_{j=0}^{\infty} b_j\ x(t-j) = \xi(t) \quad t = 0, \pm 1, \ldots \tag{8}$$

where $b_0 = 1$, $\sum_{j=0}^{\infty} b_j^2 < \infty$, and $E\{x(s)\xi(t)\} = 0$ for all $s \leq t-1$, and $\xi(t)$ is a white noise process with zero mean and variance σ^2. In the operator notation we have

$$B(x(t)) = \sum_{j=0}^{\infty} b_j u^j (x(t)) . \tag{9}$$

The autoregressive moving average process (ARMA) is defined by

$$\sum_{j=0}^{q} d_j x(t-j) = \sum_{k=0}^{p} c_k\ \xi(t-k) \tag{10}$$

or in operater notation

$$D(x(t)) = C)\xi(t)) \tag{11}$$

where

$$D(\ \ ) = \sum_{j=0}^{q} d_j u^j (\ \ ) \tag{12}$$

and

$$C(\ \ ) = \sum_{k=0}^{p} c_k u^k (\ \ ) . \tag{13}$$

We can express the ARMA process by

$$x_1(t) = D^{-1}\ C(\xi(t)) , \tag{14}$$

the AR process by

$$x_2(t) = B^{-1}(\xi(t)) , \tag{15}$$

and the MA process by

$$x_3(t) = M(\xi(t)) . \tag{16}$$

This demonstrates that all three processes can be expressed in terms of linear operators which operate

---

* For notational simplicity, we do not use ω in this section to denote stochastic processes.

on the uncorrelated noise process. The essential properties of these processes can be determined from the study of their defining operators. MA, AR, and ARMA models can be generalized by using more general operators in their definitions. This seems to be appropriate for the signal processing and system identification problems.

III. MODELING ERRORS

In many system modeling, identification and deconvolution application of an autoregressive or all-pole model is used. The reason for this is two-fold: first, the inverse filter for the all-pole model is a simple all-zero filter which can be easily implemented by adaptive linear prediction filters [6]; the second reason is that adaptive identification of an all-pole model by a simple all-zero inverse filter is equivalent to the maximum entropy spectrum analysis [9]. It is important to determine an upper boundary on the mean square error when a more general process or model is replaced by an all-pole (AR) model. For this end we derive an operator inequality that has been used to bound the error incurred when an arbitrary kernel of a Fredholm integral equation is replaced by a degenerate kernel, i.e., an exact operator equation is replaced by an operator equation with a known inverse [10]. A similar approach is applicable to the approximate modeling of stochastic processes. Let us consider a pair of operator equations

$$B\, x_o(t,\omega) = \xi(t,\omega) \tag{17}$$

$$A\, x(t,\omega) = \xi(t,\omega)\ , \tag{18}$$

where B is the AR operator defined by Eq. (9), $\xi(t,\omega)$ is the uncorrelated noise process defined in the previous section, and A is the actual operator with unknown inverse. We note that if operator B is given in terms of the poles of the filter, its inverse is given by the zeroes that are at the same location as the poles of the B. We would like to determine an upper bound for the norm

$$\| x(t,\omega) - x_o(t,\omega) \| = \| A^{-1}\, \xi(t,\omega) - B^{-1}\, \xi(t,\omega) \| \tag{19}$$

From Eq. (17) and (18) we have*

$$\begin{aligned} x - x_o &= (A^{-1} - B^{-1})\, \xi \\ &= B^{-1}(B - A)\, A^{-1}\, \xi \\ &= B^{-1} \Delta\, A^{-1}\, \xi\ , \end{aligned} \tag{20}$$

where Δ is the difference between the approximate model and exact operator, i.e.,

$$\Delta = (B - A)\ , \tag{21}$$

hence

$$A = B + \Delta = B(I + B^{-1}\,\Delta) \tag{22}$$

$$A^{-1} = (I + B^{-1}\,\Delta)^{-1}\, B^{-1} \tag{23}$$

where I is the identity operator. If $p = \| B^{-1}\,\Delta \| < 1$, then [11]

$$(I + B^{-1}\Delta)^{-1} = I - B^{-1}\Delta + (B^{-1}\Delta)^2 - (B^{-1}\Delta)^3 \ldots \tag{24}$$

Hence,

$$\begin{aligned} \| x - x_o \| &\leqslant \| B^{-1}\Delta \|\ \| (I + B^{-1}\Delta)^{-1} \|\ \| B^{-1} \|\ \| \xi \| \\ &= \| B^{-1}\Delta \|\ \| I - B^{-1}\Delta + (B^{-1}\Delta)^2 \ldots \|\ \| B^{-1} \|\ \| \xi \| \\ &< \rho(1 + \rho + \rho^2 \ldots)\ \| B^{-1} \|\ \| \xi \| \\ &= \frac{\rho}{1-\rho}\ \| B^{-1} \|\ \| \xi \| \end{aligned} \tag{26}$$

or

$$\| x - x_o \|^2 < \frac{\rho^2}{(1-\rho)^2}\ \| B^{-1} \|^2\ \| \xi \|^2\ , \tag{27}$$

which is the desired result that allows us to evaluate the signal modeling errors.

IV. MODELING OF STOCHASTIC TRANSMISSION MEDIA

Wave propagation in stochastic medium is determined by a linear differential equation which has stochastic coefficients and stochastic boundary conditions. The solution to this problem can be written in terms of an integral operator [2, 12]. The kernel of the integral operator is "random Green's function," or the impulse response of the linear stochastic (randomly time-varying) transmission channel.

There are three possible ways one can define the impulse response of a linear time-variant channel [13]. We use the definition that has useful interpretations for use in random scattering and that has convenient Fourier transform relations [12, 14, 15]. Let $h(t,\tau,\omega)$ be the response of a randomly time-varying channel at time t to a unit impulse applied at time τ. τ can be interpreted as the round trip propagation delay to the backscattering site. The output of the randomly time-varying linear channel is

$$y(t) = \int_0^\infty h(t,\tau,\omega)\ x(t-\tau)\,d\tau \tag{28}$$

The causality condition requires that $h(t,\tau,\omega) = 0$ for $\tau < 0$.

In the operator notation, we have

$$y(t) = L\ x(t) \tag{29}$$

where

$$L(\quad) = \int_0^\infty d\tau\ h(t,\tau,\omega)\ u^\tau(\quad)\ . \tag{30}$$

u^τ is the previously used delay operator. A simple example of the single channel propagation is the special case of distinct multipath propagation. In such a case we have

$$\begin{aligned} y(t) &= \int_0^\infty \sum_{i=1}^{N} a_i(t,\omega)\ \delta(\tau-\tau_i)\ x(t-\tau)\,d\tau \\ &= \sum_{i=1}^{N} a_i(t,\omega)\ u^{\tau_i}\ x(t) \\ &= [\underline{a(t,\omega)\ u^\tau}]^T\ x(t) \end{aligned} \tag{31}$$

where the elements of vector $[\underline{a(t,\omega)u^\tau}]$ are $a_i(t,\omega)u^{\tau_i}$. Hence,

$$\begin{aligned} L(\quad) &= \sum_{i=1}^{N} a_i(t,\omega)\ u^{\tau_i}(\quad) \\ &= [\underline{a(t,\omega)\ u^\tau}]^T(\quad)\ . \end{aligned} \tag{32}$$

Wave propagation in a more general linear stochastic transmission medium involves the application of stochastic integral operators. The output at time t and at a field point $\bar{r}$ is given by

---

* To simplify the notation we have dropped t and ω from the arguments.

$$y(t,\bar{r},\omega) = \int_V \int_0^\infty h(\bar{r},\bar{r}',t,\tau,\omega)\; x(t-\tau,\bar{r}-\bar{r}')\; dv\, d\tau, \qquad \omega \in \Omega \tag{33}$$

where V indicates volume integration and $h(\bar{r},\bar{r}',t,\tau,\omega)$ is a "random Green's function" [12]. This is in operator notation

$$L(\ ) = \int_V \int_0^\infty dv\, d\tau\; h(\bar{r},\bar{r}',t,\tau,\omega)\; u^{\tau} u^{\bar{r}'}(\ ) \;. \tag{34}$$

Frequently, in spatial signal processing applications, we are interested in a vector of signals

$$\underline{y(t,\omega)} = \begin{bmatrix} y(\bar{r}_1,t,\omega) \\ \vdots \\ y(\bar{r}_i,t,\omega) \\ \vdots \\ y(\bar{r}_n,t,\omega) \end{bmatrix} = \int_V \int_0^\infty \underline{h(\bar{r},\bar{r}',t,\tau)u^{\tau}u^{\bar{r}'}}\, x(t,r)dvdt$$

$$= \underline{L}\; x(t,\bar{r}) \;. \tag{35}$$

In this case $\underline{L}$ is an n x 1 matrix. A special case of interest is the case when the source is in far field. Then

$$y(t,\omega) = \int_V \int_0^\infty h(R,t)\; u^{\tau} \begin{bmatrix} u^{\tau_1} \\ \vdots \\ u^{\tau_i} \\ \vdots \\ u^{\tau} \end{bmatrix} x(t)\; d\tau \tag{36}$$

where R is the distance to the center of the array, u^{τ} is the delay operator that corresponds to the propagation delay to the center of the array, and u^{τ_i} is the delay operator corresponding to the differential delays to each space sample. The differential delay vector obviously contains information from which direction of the source can be determined. The concept of beamsteering makes use of this information. A uniformly shaded beamformer is simply

$$L(\ ) = \left[u^{-\tau_1}\ldots u^{-\tau_i}\ldots u^{-\tau_N}\right](\ )\;; \tag{37}$$

that is a row vector of shift operators that will compensate for differential propagation delays so that a signal from desired look direction is brought into time-coincidence. An operator for an M output beamformer is an M x N matrix of shift operators, where N is the number of sensors.

All the linear operators, scattering operator $\underline{L}_S$, propagation operator $\underline{L}_P$ and beamformer $\underline{L}_B$, can be cascaded so that

$$\underline{L}(\ ) = \underline{L}_B\, \underline{L}_P\, \underline{L}_S(\ ) \;. \tag{38}$$

In the next section, we will assume that the linear operator L is a cascade of operators that are pertinent to a specific problem. We note that the beamformer operator is a deterministic operator whereas the propagation and scattering operators are stochastic operators [2,12].

V. SIGNAL EXTRACTION AND DETECTION

Let the received signal vector $\underline{r(t)}$ be

$$\underline{r(t)} = \underline{L}\; \underline{s(t)} + \underline{n(t)} \tag{39}$$

where $\underline{L}$ is a linear operator, $\underline{s(t)}$ is the signal of interest, and $\underline{n(t)}$ is the interfering noise with the covariance matrix $\underline{R}_n(t,u)$. In general, $\underline{n(t)}$ can be a nonstationary, colored, anisotropic noise field. To illustrate the essential operator-theoretic ideas, we assume that $\underline{r(t)}$ is a Gaussian process. Signal and noise are assumed to be complex, zero mean, mutually uncorrelated processes. Three cases are of interest

1. $\underline{L}$ is a known linear operator,
2. $\underline{L}$ is an unknown deterministic operator,
3. $\underline{L}$ is an unknown linear stochastic operator with known or determinable covariance matrix.

In this paper we will emphasize the third case. It is of interest to estimate $\underline{L}\ \underline{s}$. It is well known that the best estimate in sense of quadratic cost of $\underline{y}$, given data vector $\underline{r}$, is the conditional mean $E\{\underline{y}/\underline{r}\}$ [16]. For zero mean, complex Gaussian distributions, the conditional expectation is given by [16]

$$E\{\underline{y}/\underline{r}\} = \underline{C}_{12}\, \underline{C}_{22}^{-1}\, \underline{r} \tag{40}$$

where $\underline{C}_{12}$ and $\underline{C}_{22}$ are elements of the joint covariance matrix of vectors $\underline{y}$ and $\underline{r}$; that is,

$$\begin{bmatrix} \underline{C}_{11} & \underline{C}_{12} \\ \underline{C}_{21} & \underline{C}_{22} \end{bmatrix}$$

with $\underline{C}_{11} = \{E\ \underline{y}\ \underline{y}^H\}$, $\underline{C}_{12} = \{E\ \underline{y}\ \underline{r}^H\} = \underline{C}_{21}^H$, and $\underline{C}_{22} = E\{\underline{r}\ \underline{r}^H\}$. Hence,

$$\underline{y} = (\widehat{\underline{Ls}}) = E\ \underline{Ls}/\underline{r}$$

$$= E\{\underline{L}\ \underline{P}\ \underline{L}^H\}[E\{\underline{L}\ \underline{P}\ \underline{L}^H\} + \underline{R}_n]^{-1}\ \underline{r} \tag{41}$$

where P is $E\{\underline{s}\ \underline{s}^H\}$ for stochastic signal or simply $\underline{s}\ \underline{s}^H$ for deterministic signal. It can be shown that the log likelihood function is given by

$$2\ell = \left[E\{\underline{L}\ \underline{P}\ \underline{L}^H\}[E\{\underline{L}\ \underline{P}\ \underline{L}^H\} + \underline{R}_n]^{-1}\ \underline{r}\right]^H \underline{R}_n^{-1} = (\widehat{\underline{Ls}})^H \underline{R}_n^{-1}\ \underline{r}. \tag{42}$$

This equation defines the signal processor structure shown in Figure (1). The log likelihood ratio detector computes the inner product of two vectors $(\widehat{\underline{L}\ \underline{s}})$, the mean square estimate of the signal

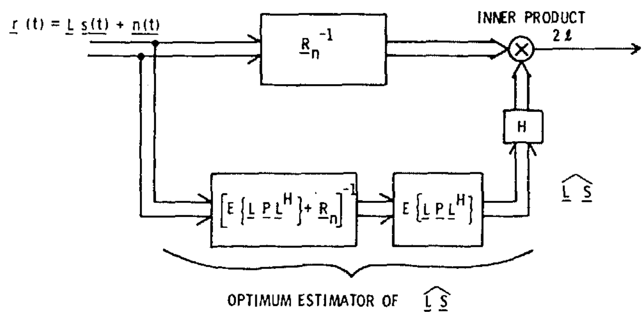

Figure 1. Maximum likelihood detector for stochastic signals.

and medium transformation, and $\underline{R}_n^{-1}\ \underline{r}$, adaptively whitened received vector. We note that $\{\underline{L}\ \underline{s}\}$ can be factored into two operators $[E\{\underline{L}\ \underline{P}\ \underline{L}^H\} + \underline{R}_n]^{-1}$ and $E\{\underline{L}\ \underline{P}\ \underline{L}^H\}$. The first operator can be determined by usual adaptive techniques from the received data given hypothesis H_1. The second operator $E\{\underline{L}\ \underline{P}\ \underline{L}^H\}$ can be determined from the a priori statistics of the stochastic channel and signal. Dr. Ricker has pointed out that sequential estimation techniques can be used to estimate $E\{\underline{L}\ \underline{P}\ \underline{L}^H\}$. This would lead to sequential detector/estimator structures. It should be pointed out that for likelihood detection one needs to estimate $\{\underline{L}\ \underline{s}\}$ not $\hat{\underline{L}}$ or $\hat{\underline{s}}$; that is, the deconvolution of $\hat{\underline{L}}$ $\hat{\underline{s}}$ is not necessary. It can be shown that the operator $E\{\underline{L}\ \underline{P}\ \underline{L}^H\}$ can be expressed in terms of generalized scattering functions [15, 21] and stochastic Green's functions [2, 12].

VI. APPLICATION OF POLAR DECOMPOSITION THEOREM

The polar decomposition theorem in operator theory states that every bounded linear operator T on a Hilbert space can be expressed as [8]:

$$T = S\ U \tag{43}$$

or

$$T = U'S' \tag{44}$$

where S (resp. S') is self-adjoint and U (resp. U') is a partial isometry.*

This theorem has an interesting interpretation in the study of space-time processors. For this purpose, consider a simple beamformer for plane waves with uniform shading and over an infinite spatial/temporal observation interval, i.e.:

$$(Ux)(t,\underline{r}) \triangleq \int_{\mathbb{R}^2}\int_{\mathbb{R}} \delta(t-u+\langle s\cdot r\rangle/c)\ x(u,s)\ ds\ du \tag{45}$$

The purpose of U is to bring all plane waves [that is, functions of the form $f(t-\langle \underline{r}\cdot\underline{l}\rangle/c)$] coming from a direction given by the unit-length vector $l \in \mathbb{R}^3$ into "time-coincidence" at the output (⟨⟩ denotes inner product in $\mathbb{R}^3$; c denotes the velocity of propagation). Equation (45) is the operator representation of a continuous aperture beamformer, where the Eq. (37) is the vector representation of a beamformer for a discrete aperture. It is straightforward to show that Equation (45) corresponds to a partial isometry since

$$\|Ux\|^2_{L_2} = \|x\|^2 \quad \forall x \in H^{\dagger}$$

$$= 0 \quad .\ x = \theta \tag{46}$$

so the subspace $M = H$, $M\underline{\perp} = \{\theta\}$ (more strongly, U turns out to be a unitary operator).

The case of nonuniform shading can be included by multiplying Equation (45) by a "spatial window"

$$(Bx)(t,\underline{r}) \triangleq w(r)(Ux)(t,\underline{r}) \tag{47}$$

where $w(\cdot)$ is some real-valued function of $\underline{r}$. The "multiplicative" operator given by:

$$(Sx)(t,\underline{r}) = w(\underline{r})x(t,\underline{r}) \tag{48}$$

is obviously a self-adjoint operator since:

$$\langle Sx,y\rangle = \langle x,Sy\rangle_H \tag{49}$$

$$\forall x,y \in H$$

Thus, our far field beamformer can be written as:

$$B = S\ U, \tag{50}$$

a special case of the polar decomposition theorem.

VII. ADAPTIVE PROCESSORS AS STOCHASTIC OPERATORS

In our context, an adaptive processor can be defined as a stochastic operator on the space of received "signals" or random processes. A recursive adaptive algorithm corresponds to a recursive definition of a sequence of random mappings which converge, in some sense, to the "desired" processor. Thus, in the framework of adaptive implementation, we are concerned with the convergence of a sequence of stochastic operators. While deterministic fixed-point theorems such as the classic Banach contraction mapping theorem play a major role in deterministic convergence problems [17], probabilistic versions of certain well-known fixed-point theorems can be important in establishing convergence conditions in a probabilistic sense. Although fixed-point theorems in probabilistic analysis have been extensively studied [18, 19], most versions seem to use an "almost everywhere" convergence condition. Due to our definition of "distance" (or norm) in the stochastic space and since the physical interpretation of average power leads to the mean square convergence criterion as a more widely accepted concept, we have used a mean square version of a stochastic contraction mapping theorem. Details have been reported in previous papers [3, 20] and a forthcoming dissertation [22].

VIII. CONCLUSIONS

In this paper, we have given a number of examples which show that operator theoretic approach provides a unifying framework for signal processing and modeling problems. These problems include modeling of stochastic processes and transmission channels, derivation of structures for optimum space-time processors, factorability of space-time processors, and use stochastic contraction mapping concepts to study the convergence of adaptive algorithms. We expect the operator-theoretic approach to signal processing will be a rich area of research.

ACKNOWLEDGMENT

This work has been done at the Applied Research Laboratory and supported by NAVSEA under Contract N00024-79-C-6043.

REFERENCES

[1] E. Parzen, in discussion of the paper "Interrelated Problems in Estimation and Detection" by H. Cox, NATO Advanced Study Institute on Signal Processing, Vol. II, pp. 23-66, Enschede, Netherlands, 1968.

[2] G. Adomian, "Random Operator Equations in Mathematical Physics I," J. of Math. Phys., Vol. 11, No. 3, pp. 1069-1084, 1970.

---

* A linear operator U is a partial isometry if:

$$\exists\ M \subset H:\ \|Ux\| = \begin{cases} \|x\| & x \in M \\ 0 & x \in M\underline{\perp} \end{cases}$$

where $\underline{\perp}$ denotes orthogonal complement.

† H denotes deterministic Hilbert space and $H\Omega$ denotes Hilbert space of stochastic processes.

[3] G. R. L. Sohie and L. H. Sibul, "Application of Hilbert Space Theory to Optimal and Adaptive Space-Time Processing," Proceedings of 16th Princeton Conference on Information Sciences and Systems, Princeton University, Princeton, 1982.

[4] L. A. Lusternik and V. J. Sobolev, Elements of Functional Analysis, Hindustan Publishing Corp., Delhi, India, 1961, p. 134.

[5] L. H. Koopmans, The Spectral Analysis of Time Series, Academic Press, New York, 1974.

[6] B. Friedlander, "System Identification Techniques for Adaptive Signal Processing," Circuits, Systems and Signal Processing, Vol. 1, No. 1, pp. 1-41, 1982.

[7] R. B. Holmes, "Mathematical Foundations of Signal Processing," SIAM Review, Vol. 21, No. 3, pp. 361-387, July 1979.

[8] P. A. Fuhrmann, Linear Systems and Operators in Hilbert Space, McGraw-Hill, New York, 1981.

[9] A. Van Den Bos, "Alternate Interpretation of Maximum Entropy Spectral Analysis," IEEE Trans. Inform. Theory, Vol. IT-17, pp. 493-494, 1971.

[10] L. V. Kantrovich and V. I. Krylov, Approximate Methods of Higher Analysis, translated from Russian by C. D. Benster, Interscience, New York, 1958.

[11] A. N. Kolmogorov and S. V. Fomin, Elements of the Theory of Functions and Functional Analysis, Vol. 1 Translated from the first Russian edition by L. F. Boron, Graylock Press, Rochester, New York, p. 102, 1957.

[12] L. H. Sibul, "Optimum Detection and Estimation in Stochastic Transmission Media," in Applied Stochastic Processes, G. Adomian, editor, pp. 247-267, Academic Press, New York, 1980.

[13] T. Kailath, "Channel Characterization: Time-Variant Dispersive Channels," Chapter 6 in Lectures on Communication Theory, E. J. Baghdady, editor, McGraw-Hill, New York, 1961.

[14] L. J. Ziomek, "A Scattering Function Approach to Underwater Acoustic Detection and Signal Design," Ph.D. dissertation in Acoustics, The Pennsylvania State University, 1981.

[15] L. J. Ziomek and L. H. Sibul, "Broadband and Narrowband Signal-to-Interference Ratio Expressions for Doubly Spread Target," J. Acoust. Soc. Am., Vol. 72, No. 3, pp. 804-819, 1982.

[16] H. Cox, "Interrelated Problems in Estimation and Detection I," NATO Advanced Study Institute on Signal Processing, Vol. II, pp. 23-1 - 23-32, Enschede, Netherlands, 1968.

[17] D. G. Luenberger, Optimization by Vector Space Methods, J. Wiley and Sons, New York, 1969.

[18] A. T. Bharucha-Reid, Random Integral Equations, Academic Press, New York, 1972.

[19] O. Hans, "Random Operator Equations," Proc. Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. II, pt. 1, pp. 185-202, University of California press, Berkeley, 1961.

[20] G. R. L. Sohie and L. H. Sibul, "Stochastic convergence properties of the Adaptive Gradient Lattice," Proceedings of International Conf. on Acoustics Speech and Signal Processing, Boston, MA, April 1983.

[21] R. Laval, "Time-Frequency-Space Generalized Coherence and Scattering Functios," in Aspects of Signal Processing, edited by G. Tacconi, D. Reidel Publishing Company, pp. 69-88, Dordrecht-Holland, 1977.

[22] G. R. L. Sohie, "Application of Hilbert Space theory to Optimal and Adaptive Array Processing," Ph.D. Thesis in Acoustics, The Pennsylvania State University, 1983.