ON PREDICTION OF HARMONIZABLE STABLE PROCESSES

by

S. Cambanis and A.G. Miamee

Technical Report No. 110

July 1985

Approved for public release, distribution unlimited
**Report: Prediction of Harmonizable Stable Processes**

S. Cambanis and A.G. Miamee

**Abstract**

Spectral and time domain criteria for a harmonizable stable process to be regular are given, which provide an orthogonal moving average representation. Also criteria for such processes to have linear predictor filters are obtained; these include the positivity of the distance and of the angle between past and future. In the process, the notion of angle between isotropic complex stable random variables is introduced and studied.

**COSATI Codes**

<table>
<thead>
<tr>
<th>FIELD</th>
<th>GROUP</th>
<th>SUB.GR.</th>
</tr>
</thead>
</table>

**Subject Terms**

Harmonizable stable processes, moving average representation, linear predictors, angle, isotropic complex random variables

**Abstract (Continued)**

...
ON PREDICTION OF HARMONIZABLE STABLE PROCESSES*

by

S. Cambanis and A.G. Miamee

Center for Stochastic Processes
Department of Statistics
University of North Carolina
Chapel Hill, N.C. 27514

Abstract. Spectral and time domain criteria for a harmonizable stable process to be regular are given, which provide an orthogonal moving average representation. Also criteria for such processes to have linear predictor filters are obtained; these include the positivity of the distance and of the angle between past and future. In the process, the notion of angle between isotropic complex stable random variables is introduced and studied.


Key words and phrases: Harmonizable stable processes, moving average representations, linear predictors, angle, isotropic complex random variables.

*Research supported by AFORS Grant F49620 82 C0009.

On leave from Isfahan Univ. of Technology, Isfahan, Iran.
1. Introduction.

The prediction theory of second order, and of Gaussian, stationary processes has a vast literature developed over the last several decades and is now standard; see for example Rozanov (1967). On the other hand, the prediction theory of p-th order, $0 < p < 2$, and in particular of stable processes has only recently been the subject of intense investigation.

Here we concentrate on the prediction of stationary stable sequences. The main difficulty compared with the Gaussian case arises from the need to work in Banach, rather than Hilbert, spaces, where orthogonality, projections, and the like have by far weaker properties and are much more unwieldy in their structure. Another source of difficulties is due to the richness of the class of stationary stable processes, which are fully described in Hardin (1982). Stationary stable processes include in particular moving averages of independent stable r.v.'s; harmonizable processes, i.e. Fourier transforms of stable processes with independent increments; sub-Gaussian processes; etc. Surprisingly, all these three classes (and many more) are actually disjoint (see Cambanis and Soltani (1984)), while all stationary Gaussian processes are harmonizable.

At this stage of its development the study of stable processes is frequently proceeding by a comprehensive study of special subclasses, such as moving averages, harmonizable, etc. For instance parameter estimation of autoregressive processes has been developed in Hannan and Kanter (1977), and prediction of autoregressive moving averages (ARMA) has been considered in Cline and Brockwell (1985). Here we concentrate on harmonizable stable processes. Even though they are never ergodic (see LePage (1980) and Cambanis et al. (1984)), their spectral density can be estimated consistently (Masry and Cambanis (1984)).
Prediction theory for harmonizable processes with infinite second moments was initiated by Urbanik (1970). The first results in the stable case were obtained by Hosoya (1978) and (1982) for one step ahead prediction. The general multi-step case was considered by Cambanis and Soltani (1984). The problem of interpolation has been considered by Pourahmadi (1984) and also by Weron (1985) in a more general set-up along with some ergodic properties.

Here we pursue the development of prediction theory of harmonizable stable processes with a view to determine the extent to which the Gaussian (or second order) theory extends to the non-Gaussian stable case. Earlier works mentioned above revealed that the one step ahead predictors are given by the same recipe as in the Gaussian case (with the same spectral density), but when predicting two or more steps ahead the non-Gaussian stable predictors are generally different from their Gaussian counterparts (Cambanis and Soltani (1984)). We show that for stable processes there are three different kinds of predictors one may consider, all of which coincide in the Gaussian case and hence are natural to be considered and studied in this case. One of them is the metric predictor, which minimizes the distance, and which has been considered by the authors mentioned above. Two further predictors, which minimize appropriately defined angles, and which we will call "angle" predictors, are introduced and studied.

Specifically, in Section 3 we present spectral and time domain criteria for regularity (Theorem 1). The spectral criteria are log-integrability of the spectral density (Hosoya (1982) and Cambanis and Soltani (1984)) and a spectral density factorization analogous to the Gaussian case. The time domain criteria are a moving average representation in terms of an orthogonal, but not independent, harmonizable stationary stable sequence, the innovations of the process; and a corresponding orthogonal moving average representation of the one step
ahead metric predictors. Unlike the Gaussian case, here truncation of the orthogonal moving average representation does not generally produce the two or more steps ahead metric predictors. In the non-Gaussian stable case the moving average coefficients have to be changed with each further truncation. However, the truncation of the moving average does in fact produce the \( m \)-step ahead right angle predictors (see Section 4). A corresponding Wold decomposition is described in Theorem 2, and it is shown (Proposition 9) that the moving average and hence also the Wold decomposition obtained here is the best possible, and these stable processes cannot have any of the stronger, and more versatile, Wold decompositions considered in Cambanis et al. (1985).

Section 4 deals with the important question of the existence of prediction filters, i.e. of convergent series representations of predictors in terms of the observed values of the process. The main result of the paper, Theorem 3, provides spectral and time domain criteria for the r.v.'s of the process to form a Schauder basis for its linear space. It is remarkable that, in spite of the considerably different geometry of the non-Gaussian stable case, the positivity of angle between past and future, and the positivity of distance between past and future, turn out to characterize again the Schauder basis property. The spectral criteria are likewise analogous to those in the Gaussian case. Under any one of these criteria all predictors can be realized by filters acting on the observed part of the process, and indeed all estimation problems have solutions which can be so realized.

An important related question is to find conditions, stronger than regularity and weaker than those in Theorem 3, which are sufficient for predictor filters to exist. For the second order processes this question has been the subject of study by several authors: Akutowicz (1957), Masani (1960), Mianee and Salehi (1983), Pourahmadi (1984), (1985), and Bloomfield (1984). Some such conditions are given in
Proposition 13, which is inspired by the conditions in Bloomfield (1984). Also the relationship between the existence of a filter for the orthogonal innovations and for the predictors is discussed.

The analysis requires a systematic development of various properties of r.v.'s in the linear space of a harmonizable process. These properties are common to all isotropic linear spaces of complex symmetric α-stable (SαS) r.v.'s (i.e. with radially symmetric distributions) and are presented in Section 2. It turns out that complex symmetric stable linear systems that are isotropic share similar properties with real symmetric stable linear systems, while the same is not true for general (not necessarily isotropic) complex symmetric stable linear systems (see Cambanis (1983)). Section 2 thus deals with isotropic complex stable systems, and characterizes the linearity of conditional expectation (Proposition 6), introduces the concepts of angle and of angle projection and develops their properties (Propositions 4, 5 and 7) and shows that positive angle between subspaces is equivalent to positive distance (Proposition 3).
2. Harmonizable and other isotropic stable systems: Distance and angle.

Harmonizable processes

A harmonizable complex S\&S process \( X = (X_n, n=0, \pm 1, \ldots) \) with spectral measure \( \mu \) on \( \mathbb{R} \) is defined through its finite dimensional characteristic functions

\[
E \exp \left( \int_{-\pi}^{\pi} z \, X_n \, d\mu(z) \right) = \exp \left( -\int_{-\pi}^{\pi} \left| z \right| e^{-in\theta} d\mu(d\theta) \right).
\]

and thus is strictly stationary; or, equivalently, it is defined through its spectral representation

\[
X_n = \int_{-\pi}^{\pi} e^{-in\theta} dZ(\theta)
\]

where \( Z \) is a complex, independently scattered, isotropic S\&S measure on the Borel subsets of \( \mathbb{R} \) with

\[
E \exp \left( i \Re \int_{-\pi}^{\pi} f \, dZ \right) = \exp \left( -\int_{-\pi}^{\pi} \left| f \right|^2 d\mu \right)
\]

for all \( f \in L^2(\cdot) \) (see Hosoya (1982), Cambanis (1983)). The correspondence \( f \rightarrow \int_{-\pi}^{\pi} f \, dZ \)

is an isomorphism between \( L^2(\cdot) \) and the closure in probability \( M^X \) of the linear space of the process \( X = (X_n, -\infty < n < \infty) \), which sends \( e^{-in\theta} \) to \( X_n \). Thus every r.v. \( Y \) in \( M^X \) is of the form \( \int_{-\pi}^{\pi} f \, dZ \) for some \( f \) in \( L^2(\cdot) \), and has an isotropic, i.e. radially symmetric, distribution. The latter is evident from the ch.f. of \( \int f \, dZ \), whence replacing \( f \) by \( (r-\text{is})f \) we have

\[
E \exp \left( i \left( r \Re \int_{-\pi}^{\pi} f \, dZ + s \Im \int_{-\pi}^{\pi} f \, dZ \right) \right) = \exp \left( -(r^2 + s^2)^{1/2} \int_{-\pi}^{\pi} \left| f \right|^2 d\mu \right)
\]

Some further properties of the r.v.'s in \( M^X \) which will be needed in subsequent sections, are generally valid for linear spaces of complex S\&S r.v.'s with radially symmetric distributions, and we therefore develop them now in this set up.
Isotropic complex stable systems.

A complex r.v. $X = X_1 + iX_2$ is called isotropic $S_iS$ if $X_1$ and $X_2$ are jointly $S_iS$ with radially symmetric distribution, i.e. $E \exp \{i(r_1 X_1 + r_2 X_2)\} = \exp \{-c(r_1^2 + r_2^2)/2\}$, or in complex notation with $r = r_1 + ir_2$,

$$E \exp \{i \Re \overline{r}X\} = \exp \{-c |r|^2\}.$$ 

A complex process $X = \{X_t, t \in T\}$ is called isotropic $S_iS$ if every finite complex linear combination $\sum_{n=1}^{N} z_n X_{t_n}$ is a complex isotropic $S_iS$ r.v. Then there exists a measure space $(\mathcal{X}, \mathcal{F}, \mu)$ and complex functions $f_t \in L^i(\mu)$, $t \in T$, such that

$$E \exp \{i \Re \overline{r} \sum_{n=1}^{N} z_n X_{t_n}\} = \exp \{-i |r|^i \sum_{n=1}^{N} f_{t_n}\},$$

where $|f|_i$ denotes $||f||_{L^i(\mu)}$ (Hardin (1982)). Equivalently, if $Z$ is a complex, independently scattered, isotropic $S_iS$ measure on $(\mathcal{X}, \mathcal{F}, \mu)$, i.e. for all disjoint sets $I_1, \ldots, I_n \in \mathcal{F}$ of finite $\mu$-measure, $Z(I_1), \ldots, Z(I_n)$ are independent with $E \exp \{i \Re \overline{r} Z(I_k)\} = \exp \{-i |r|^i \mu(I_k)\}$, so that for all $f \in L^i(\mu)$,

$$E \exp \{i \Re \overline{r} \int f dZ\} = \exp \{-i |r|^i \int |f|^i_{L^i(\mu)}\},$$

then the stochastic process $\int f_t(s) dZ(s)$, $t \in T$, is stochastically equivalent to $X$. We then say that $(X_t, t \in T)$ is represented by $f_t$, $t \in T$. If $M^X$ is the closure in probability of the linear span of $(X_t, t \in T)$, then the correspondence $X_t \rightarrow f_t$ extends to an isomorphism between $M^X$ and the subspace $sp f_t$, $t \in T$ of $L^i(\mu)$. $M^X$ is then a complex isotropic $S_iS$ space and every $Y$ in $M^X$ is represented by some $f$ in $L^i(\mu)$. (For general, not necessarily isotropic, complex $S_iS$ processes see Hosoya (1978) and Cambanis (1983)).

The following moment properties will be needed in the sequel. They extend to the complex isotropic case properties known from the real case;
the real analog of (i) is immediate and of (ii) was established in Cambanis et al. (1985).

**Proposition 1.** (i) Assume $0 < \beta < 2$ and let $Y \in M^X$ be represented by $f \in L^\beta(\cdot).$

Then the pair $(\text{Re } Y, \text{Im } Y)$ has the same distribution as the pair $\sqrt{2} \cdot f \cdot R^{1/2}(\mathcal{N}_1, \mathcal{N}_2),$ where $R$ is a positive $\beta/2$ stable r.v. with $E \exp(-uR) = \exp(-u^{\beta/2}), u > 0,$ and is independent of the iid standard normal r.v.'s $\mathcal{N}_1$ and $\mathcal{N}_2$. Moreover, for all $0 < p \leq 1,$

$$E Y_{p}^{1/\beta} = \left(\frac{p^{2p} \pi (p/2) \Gamma((p/\beta))}{\Gamma(-p/\beta)}\right)^{1/p} \cdot \frac{1}{\pi} \cdot \frac{E \left|f\right|^2}{\int \left|f\right|^2 \cdot d\mu},$$

where $\mu$ is a probability measure on $\mathbb{R}.$

(ii) Assume $1 < \beta < 2$ and let $Y_1, Y_2 \in M^X$ be represented by $f_1, f_2 \in L^\beta(\cdot).$

Then for every $1 < p \leq 2,$

$$E Y_1 Y_2^{p-1} = \frac{\int_{\mathbb{R}^2} f_1 f_2^{p-1} \cdot d\mu}{\int \left|f_2\right|^p \cdot d\mu},$$

where for $\alpha > 0$ and complex $z \neq 0,$ $z^{-\alpha} = \left|z\right|^{-\alpha} \text{Re } z.$

**Proof.** (i) Since $Y \in M^X$ we have $Y = \cdot f dZ$ for some $f \in L^\beta(\cdot).$

Putting $r = r_1 + i r_2$ we obtain

$$E \exp i(r_1 Y_1 + r_2 Y_2)) = E \exp i \text{Re } \int \cdot f dZ; = \exp \left\{-(r_1^2 + r_2^2)^{\beta/2} \cdot \left\|f\right\|^2 \right\}.$$ 

The rest follows from (cf. Theorem 7.2 in Masry and Cambanis (1984))

$$E \exp i \cdot f_1 \cdot f_2 r_1 r_2 \cdot \mathbb{R}^{1/2}(\mathcal{N}_1, \mathcal{N}_2)); = E \exp \left\{-\left|r_1^2 + r_2^2\right|^{\beta/2} \cdot \left\|f_1\right\|^2 \right\}.$$ 

(ii) The calculation is similar to that on page 357 of Köthe (1969).

For complex numbers $z \neq 0$ and $w,$ and real $\lambda,$ we have $\left|\lambda + \cdot w \cdot f \right|^p = p \lambda^{p-2} \text{Re } (zw)$ and thus

$$\left. \frac{d}{d\lambda} \left|\lambda + \cdot w \cdot f \right|^p \right|_{\lambda = 0} = p \left|z\right|^{p-2} \text{Re } (zw) - i \text{ Re } (ziw),$$

$$= p \left|z\right|^{p-2} \cdot zw = p \left|z\right|^{p-2} \cdot w.$$
Using part (i) we obtain

\[ pE^\prime Y_2^{p-1} = \frac{d}{d\alpha} \int E|Y_2 + \alpha Y_1|^p - i E|Y_2 + i Y_1|^p \bigg|_{\alpha = 0} \]

\[ = C_{p,\alpha}^{P} d \int \|f_1 + f_1\|_\|P - i \|f_2 + i f_1\|_\|P \bigg|_{\alpha = 0} \]

\[ = C_{p,\alpha}^{P} \int \|f_1 + f_1\|_\|P - i \|f_2 + i f_1\|_\|P \bigg|_{\alpha = 0} \]

\[ = p C_{p,\alpha}^{P} \int \|f_1 + f_1\|_\|P - i \|f_2 + i f_1\|_\|P \bigg|_{\alpha = 0} \]

Coupled with \( E^\prime Y_2^P = C_{P,\alpha}^{P} \|f_2\|_\|P \), this establishes (ii).

In the Gaussian case \( \epsilon = 2 \) the moment expression in (i) holds for all \( p > 0 \) and in (ii) for \( p = 2 \) as well.

Now putting \( Y_1 = f_1 \), we have that \( \|Y_1\|_\| \) defines a norm when \( 1 < \epsilon \leq 2 \) and a quasi-norm when \( 0 < \epsilon < 1 \) on \( M^X \), which metrizes convergence in probability (Cambanis (1983)) and which, by (i) of Proposition 1, is equivalent to convergence in \( L^p(\|) \), 0-\( p \)-\( \| \).

When \( 1 < \epsilon \leq 2 \) and \( Y_1, Y_2 \in M^X \) are represented by \( f_1, f_2 \in L^\epsilon(\|) \), the covariance of \( Y_1 \) with \( Y_2 \) is defined by

\[ [Y_1, Y_2] = \int f_1 f_2 \| -1 \|d. \]

and by (ii) of Proposition 1 we have for all \( 1 < p \) (provided \( Y_2 \neq 0 \))

\[ \frac{[Y_1, Y_2]}{\|Y_2\|_\|} = \frac{EY_1 Y_2^{p-1}}{E Y_2^P}. \]
By H"older's inequality we have $\lbrack Y_1,Y_2 \rbrack = f_1 \cdot f_2 \cdot \phi^{-1} = Y_1 \cdot Y_2 \cdot \phi^{-1}$ with equality if and only if $Y_1 = zY_2$ for some complex $z$. The covariation of a harmonizable process is

$$[X_n,X_m] = \int e^{-i(n-m)d} \phi(\cdot) d\cdot,$$

the familiar form of the covariance of a stationary process. In the Gaussian case $\phi = 2$ the covariation reduces to one-half the covariance.

We say that the r.v.'s $Y_1$ and $Y_2$ in $M^X$ are mutually orthogonal, or plain orthogonal, if $[Y_1,Y_2] = 0$ and $[Y_2,Y_1] = 0$. When $[Y_1,Y_2] = 0$ we say that $Y_2$ is orthogonal to $Y_1$, $Y_2 : Y_1$, which is thus a nonsymmetric notion and coincides, in view of Proposition 1 (ii), with $Y_2$ being James-orthogonal to $Y_1$ as elements in any $L^p(\cdot)$, $1 \leq p \leq \infty$ (see Cambanis et al. (1985)) for a discussion in the real case). While independence and orthogonality are equivalent in the Gaussian case $\phi = 2$, when $1 < \phi < 2$ independence implies mutual orthogonality but the converse is not generally true. This is because when $0 < \phi < 2$, $Y_1$ and $Y_2$ are independent if and only if their representing functions $f_1$ and $f_2$ have disjoint supports, i.e. $f_1 \cdot f_2 = 0$ a.e. $\cdot$ (Cambanis (1983)), while mutual orthogonality merely means that $\int f_1 f_2 \cdot \phi^{-1} d\cdot = 0 = \int f_2 f_1 \cdot \phi^{-1} d\cdot$. 


We now show that just as in the real case regressions on one r.v. are linear.

**Proposition 2.** If $1 < t < 2$ and $Y_1, Y_2 \in M^X$, then

$$E(Y_2 | Y_1) = \frac{[Y_2, Y_1]}{[Y_1, Y_1]} Y_1.$$  

**Proof.** For any two jointly S. S. complex r.v.'s $Y_1, Y_2$, it is shown in Cambanis (1983) that $E(Y_2 | Y_1) = cy_1$ iff $[Y_2 - cy_1, \text{Re}(\bar{z}Y_1)]_	au = 0$ for all complex $z$ in which case $c = [Y_2, Y_1]/[Y_1, Y_1]_	au$, i.e. iff

$$[Y_1, Y_1][Y_2, \text{Re}(\bar{z}Y_1)]_	au = [Y_2, Y_1][Y_1, \text{Re}(\bar{z}Y_1)]_	au.$$  

Now let $Y_1, Y_2 \in M^X$ be represented by $f_1, f_2 : L'(\omega)$. Then the necessary and sufficient condition becomes

$$\int f_1 \text{d}u \cdot \int f_2(\text{Re} z f_1)^{<\tau, -1>} \text{d}u = \int f_2(\text{Re} z f_1)^{<\tau, -1>} \text{d}u \cdot \int f_1(\text{Re} \bar{z} f_1)^{<\tau, -1>} \text{d}u.$$  

Now from

$$E \exp \{\text{Re} (\bar{z}_1 Y_1 + \bar{z}_2 Y_2)\} = \exp \left\{ \int \left| \bar{z}_1 f_1 + \bar{z}_2 f_2 \right| \text{d}u \right\}$$

$$= \exp \left\{ \int \left| \bar{z}_1 f_1 + \bar{z}_2 f_2 e^{-i \arg f_1} \right| \text{d}u \right\}$$

it follows that $(Y_1, Y_2)$ is also represented by $(f_1, f_2 e^{-i \arg f_1})$. Thus without loss of generality we may take $f_1$ to be real, whence $\text{Re} (\bar{z} f_1) = (\text{Re} z) f_1$ and the necessary and sufficient condition for linear regression is clearly satisfied.

A natural way of defining an angle between r.v.'s $Y_1$ and $Y_2$ in $M^X$ when $1 < 2$ is as follows. We define a complex valued cosine of the angle of $Y_1$.
with $Y_2 \neq 0$ by

$$\cos_{\alpha}(Y_1, Y_2) = \frac{Y_1 \cdot Y_2}{||Y_1||_{\alpha} \cdot ||Y_2||_{\alpha}} = \frac{[Y_1 \cdot Y_2]_{\alpha}}{||Y_1||_{\alpha} \cdot ||Y_2||_{\alpha}}.$$ 

By Proposition 1(ii), it can also be written as

$$\cos_{\alpha}(Y_1, Y_2) = \frac{E (Y_1 Y_2)^{p-1}}{(E ||Y_1||^p ||Y_2||^p)^{1/p} (E ||Y_1||^p ||Y_2||^p)^{1/p}} = \frac{Y_1 \cdot Y_2}{||Y_1||_p \cdot ||Y_2||_p} = \cos_{p}(Y_1, Y_2)$$

for all $1 < p < \infty$. Thus the cosine defined through covariation agrees with that defined through $L^p(.)$ for all $1 < p < \infty$. Henceforth we will simply write $\cos(Y_1, Y_2)$, instead of $\cos_{\alpha}(Y_1, Y_2)$ or $\cos_{p}(Y_1, Y_2)$. When either $Y_1 = 0$ or $Y_2 = 0$, we define $\cos(Y_1, Y_2) = 0$. Clearly $|\cos(Y_1, Y_2)| \leq 1$ with equality only when $Y_1 = z Y_2$ for some complex $z$. The cosine of the angle of a subspace $N_1$ with another subspace $N_2$ of $M^X$ is defined by

$$r(N_1, N_2) = \sup \{ |\cos(Y_1, Y_2)| : Y_1 \in N_1, \; Y_2 \in N_2 \}.$$

and thus $(N_1, N_2) \leq 1$. Extending an idea of Helson and Szégo (1960), we say that $N_1$ and $N_2$ are at positive angle if $r(N_1, N_2) < 1$ or equivalently (as we will see in the next proposition) $r(N_2, N_1) < 1$.

The distance between two subspaces $N_1$ and $N_2$ of $M^X$ is denoted by

$$d_1(N_1, N_2) = \inf \|Y_1 - Y_2\|_{\alpha}^1 : Y_1 \in N_1, \; Y_2 \in N_2,$$

or by $d_p(N_1, N_2) = \inf (E \|Y_1 - Y_2\|_p) 1/\alpha^1 : Y_1 \in N_1, \; Y_2 \in N_2, \; \alpha = 1 = \|Y_1\| \cdot \|Y_2\|$.

In view of Proposition 1(i) we have $d_p(N_1, N_2) = \text{Const}(p, \alpha) d_1(N_1, N_2)$. We say that $N_1$ and $N_2$ are at positive distance and write $d(N_1, N_2) > 0$ if $d_1(N_1, N_2) > 0$, or equivalently if $d_p(N_1, N_2) > 0$ for some $0 < p < \infty$. We now show that when $1 < 2$ two subspaces of $M^X$ are at positive angle if and only if they are at positive distance. This is a crucial property needed for the development in Section 4.
Proposition 3. Let $1 \leq 2$ and $N_1, N_2$ be two subspaces of $M^X$. The following are equivalent.

(i) $d(N_1, N_2) > 0$.

(ii) $\varphi(N_1, N_2) < 1$.

(iii) $\varphi(N_2, N_1) < 1$.

Proof. The proof of (i) $\Rightarrow$ (ii) (in the real case) was shown to us by Jan Rosinski. Since (i) is symmetric in $N_1, N_2$, it suffices to show (i) $\Rightarrow$ (ii).

Now to show (ii) $\Rightarrow$ (i), it suffices to show that $d(N_1, N_2) = 0$ implies $(N_1, N_2) = 1$. Assume $d(N_1, N_2) = 0$. Then there exist $Y_n, Z_n, N_1, N_2$, such that $\|Y_n - Z_n\|_{\infty} \leq \|Y_n - Z_n\|_{\infty} = 0$. By Hölder's inequality

$$\|Y_n - Z_n\|_{\infty} \leq \|Y_n - Z_n\|_{\infty} \|Z_n\|_{\infty}.$$ 

It follows that $[Y_n, Z_n] \to 0$ and thus $[Y_n, Z_n] \to 1$, i.e. $\cos(Y_n, Z_n) \to 1$ and $\varphi(N_1, N_2) = 1$.

We now show (i) $\Rightarrow$ (ii). Assume (i) and put $\varepsilon = d(N_1, N_2) > 0$. Let $M_1, M_2$ be the subspaces of $L^X$ which represent $N_1, N_2$. Then (i) implies that for all $f_1, f_2 \in M_1$ with $\|f_1\| = 1 = \|f_2\|$, we have $\|f_1 - f_2\| > \varepsilon > 0$. By the uniform convexity of $L^X$ (cf. Köthe (1969), § 26.7) there is $\lambda = \varepsilon > 0$ such that $\|f_1 + f_2\|_1 \geq 2(1-\lambda)$. Thus for every $0 \leq 1$ we have

$$\frac{1}{\lambda'} \|f_2 + \lambda f_1\|_{1,1} = \frac{1}{\lambda'} \|f_2 + (1-\lambda) f_1 + f_2\|_{1,1} = \|f_2\|_{1,1}$$

$$= \frac{1}{\lambda'} \|f_2\|_{1,1} \|1 + 1 - 2f_2\|_{1,1} = \|f_1 + f_2\|_{1,1} - 1 \leq 2(1-\lambda).$$

On the other hand, as in the proof of Proposition 1.(ii), we have

$$\frac{d}{d\lambda} \|f_2 + \lambda f_1\|_{1,1} \big|_{\lambda = 0} = \text{Re} \int f_1 f_2^* \lambda > 0.$$
It follows that
\[ \text{Re} \left( \int f_1 f_2^* \, d\mu \right) \leq 1 - 2\kappa. \]

Since this is true for every \( f_1 \in M_1 \) with \( \|f_1\|_\alpha = 1 \), replacing \( f_1 \) by \( f_1 \exp -i \arg \left( \int f_2 f_2^* \, d\mu \right) \) we obtain
\[ \left| \int f_2 f_2^* \, d\mu \right| \leq 1 - 2\kappa. \]

It follows that for all \( Y_1 \in N_1, Y_2 \in N_2 \) with \( \|Y_1\|_\alpha = \|Y_2\|_\alpha \), we have
\[ [Y_1, Y_2] \leq 1 - 2\kappa \] and hence \( \kappa(N_1, N_2) \leq 1 - 2\kappa < 1 \).

Now fix a subspace \( N \) of \( M^X \) and an element \( Y \in M^X \setminus N \). In the Gaussian case \( \kappa = 2 \), the projection \( P(Y \mid N) \) of \( Y \) onto \( N \) is the element of \( N \) which is characterized by the orthogonality of \( Y - P(Y \mid N) \) and \( N \), which in this case (\( \kappa = 2 \)) is equivalent to either of the following:

\[
\begin{align*}
[n, Y - P(Y \mid N)]_\alpha &= 0, \quad \text{for all } n \in N, \\
[n, Y]_\alpha &= [n, P(Y \mid N)]_\alpha, \quad \text{for all } n \in N, \\
[Y, n]_\alpha &= [P(Y \mid N), n]_\alpha, \quad \text{for all } n \in N, \\
[Y - P(Y \mid N), n]_\alpha &= 0, \quad \text{for all } n \in N.
\end{align*}
\]

When \( \kappa < 2 \) however, since the covariation is neither symmetric nor linear in its second argument, only the last two conditions are generally equivalent, and in general the first three conditions are distinct. Thus when \( \kappa < 2 \) there are three possible ways of defining projection, via the first three conditions above. The first condition leads to the metric projection \( m(Y \mid N) \), which is the unique element in \( N \) minimizing the distance to \( Y \) from \( N \), in any of the applicable metrics discussed above, for instance
\[
\|Y - m(Y \mid N)\|_\alpha = \inf_{n \in N} \|Y - n\|_\alpha.
\]
and is uniquely determined by
\[ [n, Y-m(Y|N)]_t = 0 \text{ for all } n \in \mathbb{N}, \]
(cf. Singer (1970a)). Using the second and third conditions we define the left, resp. right, angle projection of \( Y \) onto \( N \) as an element \( a_L(Y|N) \), resp. \( a_R(Y|N) \), of \( N \) which satisfies
\[ [n,Y] = [n, a_L(Y|N)]_t \text{ for all } n \in \mathbb{N}, \]
resp.
\[ [Y,n] = [a_R(Y|N),n]_t \text{ for all } n \in \mathbb{N}. \]

The following properties further justify the terminology used.

**Proposition 4.** If \( a_L(Y|N) \), resp. \( a_R(Y|N) \), exists then \( \|a_L(Y|N)\|_t \leq \|Y\|_t \), resp. \( \|a_R(Y|N)\|_t \leq \|Y\|_t \), and if moreover \( a_L(Y|N) \neq 0 \), resp. \( a_R(Y|N) \neq 0 \), then the left, resp. right, angle projection direction minimizes the left, resp. right, angle, i.e.
\[
\sup_{n \in \mathbb{N}} \cos(n,Y) = \sup_{n \in \mathbb{N}} \cos(n, a_L(Y|N)) = \frac{\|a_L(Y|N)\|_t}{\|Y\|_t},
\]
resp.
\[
\sup_{n \in \mathbb{N}} \cos(Y,n) = \sup_{n \in \mathbb{N}} \cos(Y, a_R(Y|N)) = \frac{\|a_R(Y|N)\|_t}{\|Y\|_t}.
\]

**Proof.** If \( n \in \mathbb{N}, \|n\|_t = 1 \), and \( a_L(Y|N) \neq 0 \) we have
\[
\cos(n,Y) = \frac{[n,Y]}{\|Y\|_t} = \frac{[n,a_L(Y|N)]}{\|Y\|_t} = \frac{\|a_L(Y|N)\|_t}{\|Y\|_t} \cos(n,a_L(Y|N)),
\]
and the result follows from Hölder's inequality. Likewise for \( a_R(Y|N) \).

We now show that left angle projections always exist uniquely and we
characterize their direction.

Proposition 5. Let $N$ be a subspace of $M^X$ and $Y \in M^X 
\cap N$. The left angle projection $a_{\alpha}(Y, N)$ exists and is unique, and if it is not zero, the left angle projection direction $\delta_{\alpha}(Y, N) = a_{\alpha}(Y, N) \frac{1}{\|a_{\alpha}(Y, N)\|}$ is characterized as the element of $N$ which satisfies

\[ [n - [n, \delta_{\alpha}(Y, N)], \delta_{\alpha}(Y, N), Y] = 0 \text{ for all } n \in N. \]

Proof. If for all $n \in N$, $[n, Y] = 0$ then it follows that $a_{\alpha}(Y, N) = 0$ uniquely. We therefore assume that $[n, Y]$ does not vanish for all $n \in N$. We will show that there is a unique left angle projection direction $\delta_{\alpha}(Y, N)$ in $N$ with unit norm (written for simplicity $\alpha$):

\[ \sup_{n \in N} |\cos (n, Y)| = |\cos (\alpha, Y)|, \]

(\text{where we may in fact delete the absolute value on the right hand side}) and that it is characterized as stated. It will then follow immediately from the characterization of $\alpha$ that

\[ a_{\alpha}(Y, N) = [\alpha, Y]^{-1} \alpha, \]

satisfies for all $n \in N$, $[n, a_{\alpha}(Y, N)] = [\alpha, Y][n, \alpha] = [n, Y]$, hence it is a left angle projection of $Y$ onto $N$, and its uniqueness follows from Proposition 4 and the uniqueness of $\alpha$.

Let $f \in L^*(\cdot)$ and the subspace $M$ of $L^*(\cdot)$ represent $Y$ and $N$, and let $g$ represent $n \in N$ with $\|n\| = 1$. We have $\cos(n, Y) = \langle g f^{-1} \cdot d \cdot \|f\|^{-1} \cdot n, \alpha \rangle$, and thus to show the existence and uniqueness of $\alpha$ it is equivalent to show that there is a unique $g_{\alpha} : M$ with $\|g_{\alpha}\| = 1$ such that
\[ S = \sup \{ \int_{\Omega} |g|^\alpha - 1 \, d\mu : g \in M, \|g\| = 1 \} = \int_{\Omega} |g|^\alpha - 1 \, d\mu. \]

There exists a sequence \( g_n \in M, \|g_n\| = 1 \), such that \( \|g_n f^{\alpha - 1} \|_{\alpha} \rightarrow S. \) Since the unit sphere in \( L^\alpha(\Omega) \) is weakly compact, there is a subsequence \( \{g_{n_k}\} \) converging weakly to some \( g_0 : \int_{\Omega} |g_0 f^{\alpha - 1} \, d\mu \rightarrow 0 \), so that \( S = \|g_0 f^{\alpha - 1} \|_{\alpha}. \) Also, since weak limits from \( M \) belong to \( M \), as \( M \) is a subspace of the reflexive Banach space \( L^\alpha(\Omega) \), we have \( g_0 \in M. \) As it is clear that \( \|g_0\| = 1 \), existence of \( g_0 \) is established. To show its uniqueness, assume there are two distinct directions \( \text{sp} \{g_1\} \) and \( \text{sp} \{g_2\} \) with \( g_1, g_2 \in M \) and \( g_1 \| = 1 = \|g_2\| \), such that \( \|g_1 f^{\alpha - 1} \|_{\alpha} = \|g_2 f^{\alpha - 1} \|_{\alpha} = S \). Then \( h_i = g_i \exp \{-i \arg \langle g_i f^{\alpha - 1} \rangle \}, i = 1, 2, \) belong to \( M \), have unit norms and satisfy \( h_i f^{\alpha - 1} \) belongs to \( M \), so that

\[ \frac{1}{2} \int_{\Omega} (h_1 + h_2) f^{\alpha - 1} \, d\mu = S. \]

Since \( \text{sp} \{g_1\} \neq \text{sp} \{g_2\} \), we have \( h_1 \neq h_2 \) and putting \( \|h_1 - h_2\| = \epsilon \geq 0 \), by the strong convexity of \( L^\alpha(\Omega) \), there is \( \lambda = \lambda(\epsilon) \geq 0 \) such that \( \|h_1 + h_2\|_{\alpha} \leq 2(1 - \epsilon). \) It follows that \( h = (h_1 + h_2) h_1 + h_2 \|_{\alpha}^{-1} \) belongs to \( M \), has unit norm, and satisfies

\[ \int_{\Omega} (h_1 + h_2) f^{\alpha - 1} \, d\mu = S, \]

contradicting the definition of the supremum \( S \). Hence the uniqueness of \( g_0 \) is established.

The unique maximizing element \( g \) must satisfy \( (d/d\alpha) F(q_\alpha + \epsilon g)/\epsilon = 0 \) for all \( q \in M \), where

\[ F(q) = \int_{\Omega} gf^{\alpha - 1} \, d\mu + \epsilon \langle q f^{\alpha - 1} \rangle. \]

(see Luenberger (1969), pp. 188-189) i.e. \( \int_{\Omega} gf^{\alpha - 1} \, d\mu + \epsilon \langle q f^{\alpha - 1} \rangle = 0 \), \( q \in M \). Putting \( g = g_0 \), we find \( \epsilon = \langle g_0 f^{\alpha - 1} \rangle = 0 \), and thus the condition becomes
Expressing it in terms of the space $M^X$, this condition becomes

$$[n,Y]-[n,\xi]\xi[n,Y]_\xi=0 \text{ for all } n,\eta.$$ 

Hence the proof is complete.

In studying right angle projections we will use the following characterization of linearity of regression, which is a complex version of a result in Cambanis et al. (1985).

**Proposition 6.** Let $N$ be a subspace of $M^X$ and $Y: M^X \rightarrow N$. Then $E(Y|N)$ if and only if there exists $\tilde{Y} \in N$ such that $[Y-\tilde{Y},Z]=0$ for all $Z \in N$ and then $\tilde{Y}=E(Y|N)$.

**Proof.** Let $Y = Y_1 + iY_2$, $Z = Z_1 + iZ_2$ be represented by $f,g$ respectively and put

$$:(r_1,r_2) = E \exp(i \text{ Re}(rY + Z)) = E \exp(i(r_1Y_1 + r_2Y_2 + Z))$$

Then

$$E \cdot e^{i \text{Re}(Z)\text{Y}} = E \cdot e^{i \text{Re}(Z)\text{Y}} = E \cdot e^{i \text{Re}(Z)\text{Y}} = E \cdot e^{i \text{Re}(Z)\text{Y}}$$

(cf. proof of (ii) of Proposition 1). It follows that

$$E \cdot e^{i \text{Re}(Z)[E(Y|N)-\tilde{Y}]} = E \cdot e^{i \text{Re}(Z)(Y-\tilde{Y})}$$

$$= i \cdot \exp(-\frac{1}{2}[Y,\text{Y}]) \cdot [Y,\text{Y}].$$
Now $E(Y;N). N$ iff for some $\bar{Y}. N$ we have $E(Y;N) = \bar{Y}$ or equivalently $\text{LHS} = 0$ for all $Z. N$, i.e. $\text{RHS} = 0$ for all $Z. N$.

We now show that right angle projection does not always exist, that it is unique whenever it exists, and that it coincides with conditional expectation whenever the latter is linear. Recall that in the Gaussian case $\lambda = 2$, conditional expectations are always linear and coincide with the metric and both angle projections.

**Proposition 7.** Let $N$ be a subspace of $M^X$ and $Y. M^X. N$, and $1 < \lambda < 2$.

(i) The right angle projection of $Y$ onto $N$ may not exist in general.

(ii) If the right angle projection exists, then it is unique.

(iii) If the conditional expectation $E(Y;N)$ is linear, then the right angle projection exists and they are equal: $a_r(Y;N) = E(Y;N)$.

**Proof.** (i) The right angle projection fails to exist even in the real case.

Here is an example. Take $I = [0,1]$, $\mu = \text{Lebesgue}$, $f_1 = 1_{[0,2/3]}$, $f_2 = 1_{[1/3,1]}$, $f = 1_{[0,1]}$, $Y = \int f dZ$, $Y_1 = \int f_1 dZ$, $i = 1, 2$, $N = \text{span} Y_1, Y_2$. If $a_r$ exists it must be of the form $a_r = aY_1 + bY_2 = \int (af_1 + bf_2) dZ$ for some $a, b$; and it must satisfy $[Y, n] = [a_r, n]$ for all $n. N$, i.e. for all $n = xY_1 + yY_2 = \int (xf_1 + yf_2) dZ$. Thus $a, b$ must satisfy

$$\int_0^1 (af_1 + bf_2)(xf_1 + yf_2) = \int_0^1 (xf_1 + yf_2)^\lambda$$

for all $x, y$.

Putting $x = 0$ and $y = 0$ gives $a = b = 2/3$, and then putting $x = y$ gives the contradiction $2^\lambda = 4$. A genuinely complex isotropic example can be provided by taking

$I = (-\infty, \infty)$, $\mu = \text{Lebesgue}$, $Y_k = \int e^{-ik\cdot z} dZ$, $k = 0, 1, 2$, $N = \text{span} Y_0, Y_1$.

$Y = Y_0 + Y_1 + Y_2$, and reaching a contradiction likewise (using a property shown in Example 4.5 in Cambanis et al. (1985)).
(ii) Suppose there are $a_1, a_2 \in \mathbb{N}$ such that for all $n \in \mathbb{N}$: $[Y,n]_t = [a_1,n]_t$ and $[Y,n]_t = [a_2,n]_t$. Then $[a_1,n]_t = [a_2,n]_t$ and $[a_1-a_2,n]_t = 0$ for all $n \in \mathbb{N}$. Taking $n = a_1 - a_2$ gives $||a_1-a_2||_t = 0$ and thus $a_1 = a_2$.

(iii) If $E(Y^N) \in \mathbb{N}$, by Proposition 6 we have for all $n \in \mathbb{N}$, $[Y-E(Y^N),n] = 0$, i.e. $[Y,n]_t = [E(Y^N),n]_t$, and thus by (ii), $a_r(Y^N) = E(Y^N)$.

The following examples show that in the non-Gaussian stable case $1 < 2$, the metric projection, the left angle projection, and the right angle projection may all be distinct, and even have distinct directions.

**Example 1:** Where the metric, left angle, and right angle projections have the same direction but are distinct. Take $I = [0,1]$, $\nu = \text{Lebesgue}, 1 < 2$, $Y = \int (0,2/3) dZ, W = \int (0,1) dZ, N = \text{sp} W$. It is easily seen that

$$a_r(Y,W) = E(Y, W) = \frac{2}{3} W, a_e(Y,W) = (\frac{2}{3})^{\frac{1}{1-1}} W, m(Y,W) = \frac{1}{1+2^{\frac{1}{1-1}}} W,$$

and hence they are all distinct.

**Example 2:** Where the metric, left and right projections have distinct directions. Take $I = [0,1]$, $\nu = \text{Lebesgue}, 1 < 2$, $Y_1 = \int (0,1/2) dZ, Y_2 = \int (1/2,1) dZ, N = \text{sp} Y_1, Y_2$ and $Y = \int (0,2/3) dZ$. An easy calculation shows that

$$a_r(Y_1,Y_2) = E(Y_1,Y_2) = Y_1 + \frac{1}{3} Y_2,$$

$$a_e(Y_1,Y_2) = Y_1 + (\frac{1}{3})^{\frac{1}{1-1}} Y_2,$$

$$m(Y_1,Y_2) = Y_1 + \frac{1}{1+2^{\frac{1}{1-1}}} Y_2,$$

and that they have distinct directions (i.e. coefficients of $Y_2$).

An interesting infinite dimensional case where all projections have the
same direction arises when \( X = \|X_t \|_t, t \geq 0 \) is a so-called \( \alpha \)-sub-Gaussian process, i.e. \( X_t = A^{1/2} G_t, t \geq 0 \), where \( G = \|G_t \|_t, t \geq 0 \) is a zero-mean Gaussian process independent of the positive \( \alpha/2 \)-stable r.v. \( A \), with \( \mathbb{E} \exp \left( -uA \right) = \exp \left( -u^{\alpha /2} \right), u > 0 \).

**Proposition 8.** Let \( X \) be \( \alpha \)-sub-Gaussian with \( 1 < \alpha < 2 \), \( N \) a subspace of \( M^X \) and \( Y \subset M^X N \). Then

\[
m(Y|N) = a_r(Y|N) = \mathbb{E}(Y|N) = c a_r(Y|N)
\]

for some constant \( c \) (depending on \( Y \) and \( N \)).

**Proof.** We have \( M^X = A^{1/2} M^G \) and thus \( N = A^{1/2} L \) for some subspace \( L \) of \( M^G \), and \( Y = A^{1/2} W \) for some \( W \subset L \). Also, from Corollary 2.3 in Cambanis and Miller (1981),

\[
[Y_1, Y_2]_t = \frac{\mathbb{E}(W_1 W_2)}{2^{\alpha/2} [\mathbb{E}(W^2)_1]^{1-\alpha/2}}, \quad ||Y||_Y = \left( \frac{1}{2} \mathbb{E}(W^2)_1 \right)^{1/2}.
\]

The expression of the norm shows that \( m(Y|N) = A^{1/2} m(W|L) = A^{1/2} \mathbb{E}(W|L) \). The expression for the covariation then shows that for all \( n = A^{1/2} e \subset N, e \subset L \),

\[
[m(Y|N), n] = \frac{\mathbb{E}(\mathbb{E}(W|L) e)}{2^{\alpha/2} [\mathbb{E}(\mathbb{E}(W^2|L))^{1-\alpha/2}} = \frac{\mathbb{E}(W)}{2^{\alpha/2} [\mathbb{E}(W^2)_1]^{1-\alpha/2}} = [Y, n],
\]

so that \( a_r(Y|N) = m(Y|N) \). As for sub-Gaussian processes conditional expectations are linear (cf. Hardin (1982a)), we have by Proposition 7(iii), \( a_r(Y|N) = \mathbb{E}(Y|N) \).

We also see that

\[
[n, m(Y|N)] = \frac{\mathbb{E}(\mathbb{E}(W|L))}{2^{\alpha/2} [\mathbb{E}(\mathbb{E}(W^2|L))^{1-\alpha/2}} = \frac{\mathbb{E}(W)}{2^{\alpha/2} [\mathbb{E}(W^2)_1]^{1-\alpha/2}} = c^{-1} [n, Y|N],
\]

from which it follows that \( a_r(Y|N) = c^{-1} m(Y|N) \).
3. Regularity and orthogonal moving average representation.

In this section we obtain criteria for regularity of harmonizable $S$-$S$ processes $X_n$, and a Wold decomposition. We first present and discuss the results, and then prove them.

A process is called regular when its remote past is empty and singular when its remote past contains all the (linear) information. Specifically, let us denote by $M_n^X$, resp. $M^X$, the closure in probability, or in $L^p$ norm or in $L^p(\cdot)$ norm (cf. Proposition 1), of the linear span of $\{X_k, k\leq n\}$, resp. $\{X_k, k\geq n\}$. The remote past of $X$ is the subspace $M_{\infty}^X = \bigcap_n M_n^X$. $X$ is called regular if $M_{\infty}^X = \emptyset$ and singular if $M_{\infty}^X = M^X$. $H^\ell$ denotes the space of Hardy functions in the unit disk. Spectral and time domain criteria for regularity are given in the following Theorem 1. For a harmonizable $S$-$S$ process $X$ with $1 < 2$ and spectral measure $\mu$, the following are equivalent.

(1) $X$ is regular.

(2) $\mu(\cdot) = f(\cdot)d\nu$ and $\int_{H^\ell} \log f(\cdot) d\nu > -\infty$.

(3) $\mu(\cdot) = f(\cdot) d\nu$ and $f(\cdot) = \langle \cdot, \cdot \rangle$ where $\cdot \in H^\ell$.

(4) $X$ has a moving average representation $X_n = \sum_{k=0}^{\infty} a_k V_{n-k}$, where the process $V = \{V_n, n \geq 0\}$ is jointly stationary with $X$, satisfies $M_n^X = M_n^V$, and has mutually orthogonal r.v.'s.

(5) The one step ahead linear predictor $X_{n+1} \mid n$ of $X_{n+1}$ based on $\{X_k, k \leq n\}$ is given by $\hat{X}_{n+1} \mid n = \sum_{k=1}^{\infty} a_k V_{n+1-k}$, where the process $V = \{V_n\}$ is jointly stationary with $X$, satisfies $M_n^X = M_n^V$, and has mutually orthogonal r.v.'s.

These criteria extend to the case $1 = 2$ the well known criteria for regularity in the Gaussian case $\mu = \delta$. While the spectral domain criteria (2) and (3) are nearly identical to those in the Gaussian case, the time domain
criteria (4) and (5) exhibit significant differences with their Gaussian counterparts. The series in (4) and (5) converge in norm, or equivalently in $p$-th order mean $O(p)$.

The spectral domain criterion (2) was established in Cambanis and Soltani (1984) and has the feature of being independent of the index of stability $\gamma$. The spectral density factorization criterion (3) does depend on $\gamma$, does not require $\gamma$ to be outer, even though this may be added to it without loss of generality, and leads to the following

**Corollary 1.** If $0 < f \cdot L^1$, then $f$ is factorable as $f = \bigvee$ with $\bigvee H^1$, if and only if $\log f \cdot L^1$.

The time domain criterion (4) provides a "unique" orthogonal moving average representation in terms of a $S_sS$ process $V$. As shown in the proof of Theorem 1, the $S_sS$ process $V$ is in fact harmonizable with Lebesgue spectral measure, and up to a fixed multiple the weights $a_k$ in the moving average are the Fourier coefficients of the outer factor $\gamma$ of the spectral density $f$. The necessity of the moving average representation (4) is a refinement in the discrete time case of a continuous-time result in Cambanis and Soltani (1984) (Theorem 3.1). In sharp contrast with the Gaussian case where the r.v.'s of $V$ are independent, in the non-Gaussian stable case the process $V$ never has independent r.v.'s; this is the discrete-time analog of a continuous-time result in Theorem 3.1 of Cambanis and Soltani (1984). Thus the moving average obtained here is the best extension to stationary harmonizable stable processes of the result for stationary Gaussian processes. More specifically, we can prove the following

**Proposition 9.** (i) A harmonizable $S_sS$ process $V$ with $1 > \gamma$ is regular if and only if it has a moving average representation $X_n = \sum_{k=0}^{\infty} a_k V_{n-k}$ where $a_0 > 0$. 

and \( V = :V_n \) is jointly stationary with \( X \), satisfies \( M_n^X = M_n^V \), and has mutually orthogonal r.v.'s with norm one.

(ii) The representation in (i) is unique.

(iii) No harmonizable non-Gaussian S.S process \( X \) with \( 0 < \delta \leq 2 \) is the moving average of an independent S.S process \( V \) with \( M_n^X = M_n^V \).

The time domain criterion (5) expresses the one step ahead linear predictor as the one term truncation of the moving average. Its necessity is implicit in Hosoya (1982) and Cambanis and Soltani (1984). In sharp contrast with the Gaussian case, however, in the non-Gaussian stable case the \( m \)-term truncation of the moving average does not generally produce the \( m \)-step ahead linear predictor for \( m > 2 \). The linear predictor \( \hat{X}_{n+m,n} \) of \( X_{n+m} \) based on \( \langle X_k, k \leq n \rangle \) is the best metric approximation to \( X_{n+m} \) in \( M_n^X \) :

\[
\| X_{n+m} - \hat{X}_{n+m,n} \|_2 = \inf \| X_{n+m} - Y \|_2 \quad Y \in M_n^X
\]

or equivalently, by Proposition 1,

\[
E(X_{n+m} - \hat{X}_{n+m,n}) = \inf E(X_{n+m} - Y) \quad Y \in M_n^X,
\]

and is uniquely determined by

\[
[X_k, X_{n+m} - \hat{X}_{n+m,n}] = 0 \quad \text{for all } k \leq n.
\]

cf. Singer (1970a). In particular, the \( m \)-step ahead linear predictor is given by the \( m \)-term truncated moving average:

\[
X_{n+m,n} = \sum_{k=0}^{m-1} a_k V_{n+m-k}
\]

if and only if

\[
[V_{j+k}, a_k V_{n+m-k}] = 0 \quad \text{for all } j \leq n,
\]

since \( M_n^X = M_n^V \); or equivalently if and only if

\[
\int \left( \sum_{k=0}^{m-1} a_k e^{ik\zeta} \right)^2 \cdot |d\zeta| = 0 \quad \text{for all } \zeta \in \mathbb{C}.
\]
where the $a_k$'s are the Fourier coefficients of the outer factor $f$ of the spectral density $f$.

Putting together the (clearly unique) decomposition into independent regular and singular components obtained in Cambanis and Soltani (1984), Theorem 4.2, along with Theorem 1 and Proposition 9 we have the following

**Theorem 2. Wold decomposition.** Let $\{X_n\}$ be a (non-singular) harmonizable $\mathcal{S}_t\mathcal{S}$ process with $1<\alpha<2$. Then there is a unique 4-variate harmonizable $\mathcal{S}_t\mathcal{S}$ process $\{X_n, Y_n, Z_n, V_n\}$ such that

$$X_n = Y_n + Z_n = \sum_{k=0}^{\infty} a_k V_{n-k} + Z_n,$$

$\{Y_n\}$ is regular, $\{Z_n\}$ is singular and independent of $\{Y_n\}$ and of $\{V_n\}$, $a_0>0$, and $\{V_n\}$ are orthogonal and satisfy $M^n_X = M^n_Y + M^n_V$

Of course we also have that $M^n_X - M^n_Y$ is independent of $M^n_V = M^n_Y$, and $Z_n$ is the metric projection of $X_n$ onto $M^n_X$. In the Gaussian case $\alpha=2$ the innovations $\{V_n\}$ are independent, and the $m$-step ahead linear or regression predictors are obtained by $m$-term truncation of the right hand side of the Wold decomposition. In the non-Gaussian stable case the Wold decomposition described in Theorem 2 has substantially weaker consequences and in particular provides only the one step ahead linear predictors. For general $\mathcal{S}_t\mathcal{S}$ processes Wold decompositions with stronger properties, called "right", "left" and "independent" Wold decompositions, are defined and studied in Cambanis et al. (1985), to which the reader is referred for definitions and details. However harmonizable $\mathcal{S}_t\mathcal{S}$ processes can not have any of these stronger Wold decompositions.

**Proposition 10.** A harmonizable $\mathcal{S}_t\mathcal{S}$ process does not have a right, left or
independent Wold decomposition.

**Proof of Theorem 1.** The equivalence of (1) and (2) is shown in Cambanis and Soltani (1984). We first show that (2) \( \Rightarrow \) (3) \( \Rightarrow \) (1).

Assume (2). Then \( \phi \) can be defined as in Cambanis and Soltani (1984) (Eq. (5.4) or Remark 5.1): Since \( \log f \in L^1 \), the function

\[
\phi(z) = \exp \left( \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{i0 + e \log f(\rho)}}{e^{i0 + e \rho}} d\rho \right), \quad |z| < 1,
\]

is outer, and for a.e. \( \sigma \), \( \lim_{r \uparrow 1} \phi(e^{i0}) = \phi(\cdot) \) and \( \phi(\cdot)^{-1} = f(\cdot) \) (cf. Rudin (1966), Theorem 17.16).

Now assume (3). Consider the linear isometry \( U_1 : L^\alpha(f) \rightarrow M^X \) defined by \( U_1(f) = \int_\mathbb{R} f d\mathcal{L} \), which is onto (cf. Section 2). Also note that in view of (3), \( U_2 : L^\alpha(f) \rightarrow L^\alpha(\mathcal{L}) \) defined by \( U_2(g) = g(\cdot) \) is a linear isometry (which is not necessarily onto). Then \( U = U_2 U_1^{-1} : M^X \rightarrow L^\alpha \) is a linear isometry (which is not necessarily onto). Since \( U(X_n) = U_2 U_1^{-1}(X_n) = U_2(e^{-in\sigma}) = e^{-in\sigma} \phi(\cdot) \), we have for all \( n \),

\[
U(M^X_n) = L^\alpha - \text{sp} \{ e^{-ik\sigma}, \quad k \leq n \} \\
= L^\alpha - \text{sp} \{ e^{ij\sigma}, \quad j \geq -n \} \\
\leq L^\alpha - \text{sp} \{ e^{i\sigma}, \quad j \geq -n \} \quad (\text{as } H^\alpha) \\
= L^\alpha - \text{sp} \{ e^{-i\sigma}, \quad k \geq n \}.
\]

Thus \( n M^X_n : U^{-1}[\cdot, L^\alpha - \text{sp} e^{-ik\sigma}, \quad k \leq n] \), and in order to show (1): \( n M^X_n = \{0\} \), it suffices to show \( L^\alpha - \text{sp} e^{-ik\sigma}, \quad k \geq n \) for all \( n \). Then

\[
\int_{-\infty}^{\infty} h(n) e^{-ij\sigma} d\sigma = 0 \quad \text{for all } j \geq -n.
\]
Since \( h \in L^\alpha \subset L^1 \) and all its Fourier coefficients are zero, \( h = 0 \).

We now show \((1) \iff (4)\). First assume \((1)\). Then choose an outer factor \( \phi \) in \((3)\), so that \( \phi \neq 0 \) a.e. It follows that the isometry \( U_2 \) considered in the previous paragraph is onto, and hence so is the isometry \( U : M^X \to L^\alpha \). Thus \( V_n = U^{-1}(e^{-in\theta}) \) are well defined and satisfy \( M_n^X = M_n^V \). Since \( \phi \in H^\alpha \), it has a Fourier series

\[
\phi(\theta) = \sum_{k=0}^{\infty} a_k e^{ik\theta}
\]

which converges in \( L^\alpha \), and thus

\[
X_n = U^{-1}(e^{-in\theta}(\psi)) = \sum_{k=0}^{\infty} a_k U^{-1}(e^{-i(n-k)\theta}) = \sum_{k=0}^{\infty} a_k V_{n-k}.
\]

Also, in view of the isomorphism \( U \), we have

\[
E \exp \left\{ i \operatorname{Re} \sum_{n=1}^{N} z_n V_n \right\} = \exp \left\{ -\int_{-\pi}^{\pi} \left| \sum_{n=1}^{N} z_n e^{-in\theta} \right| d\theta \right\}.
\]

Thus \( V = (V_n) \) is harmonizable \( S_\alpha S \) with Lebesgue spectral measure and thus mutually orthogonal r.v.'s:

\[
[V_k, V_n]_\alpha = \int_{-\pi}^{\pi} e^{-ik\theta} \left| \sum_{n=1}^{N} z_n e^{-in\theta} \right| d\theta = 0 \quad \text{for all } k \neq n.
\]

This shows \((4)\). The joint stationarity of \( X, V \) is evident from \( X_n = U^{-1}(e^{-in\theta}(\psi)) \) and \( V_n = U^{-1}(e^{-in\theta}) \).

Conversely assume \((4)\). Since \( M_n^V = M_n^X \), each \( V_n \) belongs to \( M^X \) and is thus of the form

\[
V_n = \int_{-\pi}^{\pi} g_n(\theta) d\theta, \quad g_n \in L^1(\mathbb{R}).
\]

Now for all \( n, m \) we have
Thus for each \( m \) the Fourier transforms of the finite measures \( g_m <\alpha-1> \) and 
\( e^{-i(n-m)\theta} g_0 <\alpha-1> \) coincide, hence these measures are identical, i.e.
\[
 g_m(n) = e^{-in\theta} g_0(n) \quad \text{a.e. [\( \theta \)]}
\]
(since \( z = w \text{ iff } z <\beta> = w <\beta> \)). From the orthogonality of the \( V_n \)'s we have
\[
 0 = [V_n, V_0]_\alpha = \left\{ \int_{-\pi}^\pi g_m(n) g_0(n) <\alpha-1> d\mu = \int_{-\pi}^\pi e^{-in\theta} g_0(n) |<\alpha-1> d\mu, \quad n \neq 0. \right.
\]
It follows from the Riesz theorem that the measure \( |g_0(\theta)|<\alpha> d\mu(n) \) is absolutely continuous with respect to Lebesgue measure: \( |g_0(n)|<\alpha> d\mu(\theta) = c(\theta) d\theta \), and then the above equality reduces to \( \int_{-\pi}^\pi e^{-in\theta} c(\theta) d\theta = 0 \) for all \( n \neq 0 \), which in turn implies that \( c(\theta) = \text{positive constant} = c^{t} \), say, a.e. [Leb]. Thus
\[
 |g_0(\theta)|<\alpha> d\mu(n) = c^{t} d\theta.
\]
Note that \( H_n^{X} \) equals \( M_n^{V} \), which is isomorphic under the stochastic integral to
\[
 L^{t}(\mu) - \text{sp}(e^{-ik\theta}, k <n) = (L^{t} - \text{sp}(e^{-ik\theta}, k <n)) g_0,
\]
which is in turn isomorphic to \( L^{t} - \text{sp}(e^{-ik\theta}, k <n) \) under the correspondence \( h \cdot g_0 \leftrightarrow \chi \). Thus in order to show that \( X \) is regular it is equivalent to show that \( n L^{t} - \text{sp}(e^{-ik\theta}, k <n) = 0 \), which has been done in the third paragraph of this proof. Thus (1) is shown.
We finally show that (4) \( \prec \) (5). Assume (4). Put \( Y = \sum_{k=1}^{n} a_{k} V_{n+1-k} \).

Then \( X_{n+1} - Y = a_{0} V_{n+1} \) and the orthogonality of the \( V_{k} \)'s imply

\[ [V_{k}, X_{n+1} - Y] = 0 \]

for all \( k \prec n \), and by linearity and continuity \( [W, X_{n+1} - Y] = 0 \) for all \( W \prec M_{n}^{V} = M_{n}^{X} \), including in particular all \( X_{k} \)'s with \( k \prec n \). It follows that \( Y = X_{n+1,n} \).

Conversely assume (5). Put \( W_{n+1} = X_{n+1} - X_{n+1,n} \). Then \( [W, W_{n+1}] = 0 \) for all \( W \prec M_{n}^{V} = M_{n}^{X} \). Also from \( X_{n+1} = W_{n+1} + \sum_{k=1}^{n} a_{k} V_{n+1-k} \) we obtain

\[ M_{n+1}^{V} = M_{n+1}^{V} + \text{sp}(W_{n+1}) \]

and from \( M_{n+1}^{X} = M_{n+1}^{V} \) by a standard argument

\[ M_{n+1}^{X} = M_{n+1}^{V} + \text{sp}(V_{n+1}) \].

Hence \( W_{n+1} = M_{n+1}^{X} \) can be written in the form \( W_{n+1} = Y + c V_{n+1} \) where \( Y \prec M_{n}^{V} \), and thus \( W_{n+1} - c V_{n+1} = Y \prec M_{n}^{V} \). It then follows that

\[ [W_{n+1} - c V_{n+1}, W_{n+1}]_{\alpha} = 0 \quad \text{and} \quad [W_{n+1} - c V_{n+1}, V_{n+1}]_{\alpha} = 0 \]

i.e.

\[ \|W_{n+1}\|_{\alpha}^{2} = c [V_{n+1}, W_{n+1}]_{\alpha} \quad \text{and} \quad [W_{n+1}, V_{n+1}]_{\alpha} = c \|V_{n+1}\|_{\alpha}^{2} \].

Then \( c \neq 0 \). For if \( c = 0 \Rightarrow W_{n+1} = 0 \Rightarrow X_{n+1} = X_{n+1,n} \prec M_{n}, \Rightarrow M_{n+1}^{X} = M_{n}^{X} \).

\[ M_{n}^{V} > [V_{n+1}, V_{n+1}]_{\alpha} = 0 \Rightarrow V_{n} = 0 \Rightarrow X_{n} = 0 \]

i.e. \( X \) is the zero process. Then multiplying the above equalities together we obtain

\[ [V_{n+1}, W_{n+1}]_{\alpha} = \|V_{n+1}\|_{\alpha}^{2} \quad \text{and} \quad [W_{n+1}, V_{n+1}]_{\alpha} = c \|V_{n+1}\|_{\alpha}^{2} \]

Writing \( V_{n+1} = f_{n+1} dZ \) and \( W_{n+1} = g_{n+1} dZ \), we thus have

\[ \int f_{n+1}^{< \alpha-1} d_{n} \cdot \int g_{n+1}^{< \alpha-1} d_{n} = \int f_{n+1}^{< \alpha-1} d_{n} \cdot \int g_{n+1}^{< \alpha-1} d_{n} \]

Dropping for simplicity the subscripts, this means that equality holds in Hölder's inequalities \( \int |f|^{\alpha} d_{n} \leq \|f\|_{\alpha}^{\alpha} \|g\|_{\alpha}^{\alpha-1} \)

\[ \int |g|^{\alpha} d_{n} \leq \|g\|_{\alpha}^{\alpha} \|f\|_{\alpha}^{\alpha-1} \],

so that \( |g(n)| = r|f(0)| \) a.e. \( [i_{n}] \), for some \( r > 0 \),
and thus \( g(\omega) = re^{i\phi} f(\omega) \) a.e. \([\omega]\). Substituting in the above equation we obtain

\[
\left| \int e^{i\phi} f(\omega) d\mu \right|^2 = \left( \int f(\omega) d\mu \right)^2
\]

and an elementary argument shows that \( e^{i\phi} \) is a complex constant for a.e. w.r.t. \( f \). Thus \( g_{n+1}(\omega) = z_{n+1} f_{n+1}(\omega) \) a.e. \([\omega]\) for some complex constant \( z_{n+1} \), and \( W_{n+1} = z_{n+1} V_{n+1} \). The joint stationarity of \( X \) and \( V \) implies

\[ [X_{n+1} - X_{n+1}, n, V_{n+1}] = [W_{n+1}, V_{n+1}]_\alpha = z_{n+1} \]

is independent of \( n \). Hence putting \( z_{n+1} = a_0 \) we obtain \( X_{n+1} = \sum_{k=0}^n a_k V_{n+1-k} \). Thus (4) is established and the proof of the theorem is complete.

Proof of Proposition 9. (i) and (ii). Assume (1) of Theorem 1 and consider two moving average representations as in (4): \( X_n = \sum_{k=0}^\infty a_k V_{n-k} = \sum_{k=0}^\infty b_k U_{n-k} \).

Since the metric projection \( \hat{X}_{n,n-1} = \sum_{k=1}^\infty a_k V_{n-k} = \sum_{k=1}^\infty b_k U_{n-k} \) of \( X_n \) onto \( M_n^X = M_n^V = M_n^U \) is unique we obtain \( a_k V_n = b_k U_n \). By absorbing in \( V_n \), resp. \( U_n \), the phase of \( a_0 \), resp. \( b_0 \), we may assume without loss of generality that \( a_0 = b_0 = 0 \). Since \( |V_n| = 1 = |U_n| \), it follows that \( |a_0| = |b_0| \) and hence \( a_0 = b_0 \). Thus we have \( V_n = U_n \) for all \( n \) and hence \( \sum_{k=1}^\infty (a_n-b_n) V_k = 0 \) which implies \( a_n = b_n \) by the orthogonality of the \( V_n \)'s. This shows both (i) and (ii).

(iii). Suppose on the contrary that \( X_n = \sum_{k=0}^\infty a_k V_{n-k} \) where the r.v.'s \( V_n \) are independent. Since \( V_n \sim M_n^X \), they are of the form \( V_n = \sum_{n} f_n(\omega) d\mu(\omega) \), \( f_n \in L'('\omega) \), and the mutual independence of the \( V_n \)'s implies the \( f_n \)'s have mutually disjoint supports, say \( E_n \), (see Cambanis (1983) for the complex case considered here). It then follows from \( X_n = \sum_{n} e^{-i\omega} e^{-i\omega} d\mu(\omega) \) that for all \( n \),

\[
e^{-i\omega} = \sum_{k=0}^\infty a_k f_{n-k}(\omega) \text{ in } L'(\omega).
\]
Thus for all \( k \geq 0 \) and all \( n \), on \( \mathbb{E}_{n-k} \): 
\[
e^{-in} = a_k f_{n-k}(\cdot) \text{ a.e. } [\mu],
\]
or, equivalently,
\[
on \mathbb{E}_m: e^{-i(m+j)} = a_j f_m(\cdot) \text{ a.e. } [\mu], j \geq 0.
\]

If all \( f_m = 0 \) then all \( V_n = 0 \) and \( X = 0 \). Thus for some \( m \), \( \sum_{m} |f_m|^2 \mu > 0 \). It follows from the displayed equality that then \( a_j \neq 0 \), \( j > 0 \), which in turn implies \( \sum_{E_m} f_m \cdot d \mu > 0 \) for all \( m \). Now fix an arbitrary \( m \), and some \( \mathbb{E}_m \) with an \( \mathbb{E}_m \) neighbourhood of positive \( \mu \) measure. Then \( f_m(\cdot) = e^{-i(m+j)} / a_j \) for all \( j > 0 \) implies \( a_j = e^{-ij} a_0 \) for all \( j > 0 \). But since this should hold for each such \( m \) on each of the disjoint sets \( \mathbb{E}_m \), it leads to an obvious contradiction. Thus (iii) is proven.

Proof of Proposition 10. An independent Wold decomposition (WD) is precluded by Theorem 2. Assume now \( X \) has a left WD: 
\[
X_n = Y_n + Z_n = \sum_{k=0}^\infty a_k V_{n-k} + Z_n,
\]
along with the WD described in Theorem 2. Then \( Z_n \) is the metric projection of \( X_n \) onto \( \mathbb{M}_n \) (Cambanis et al. 1985) hence \( Z_n = Z_n \) and thus also \( Y_n = Y_n \). It follows that \( \mathbb{V}_n = \mathbb{M}_n = \mathbb{M}_n = \mathbb{M}_n \) and thus \( V \) has a left WD, since \( V \) does.

Similarly assuming \( X \) has a right WD it follows that so does \( V \). But it has been proven in Example 4 of Cambanis et al. (1985), that a harmonizable S.S process with Lebesgue spectral measure, such as \( V \) of Theorems 1 and 2, has no left nor right WD. Thus the proof of the Proposition is complete.
4. Positive angle and distance between past and future.

In this section we give spectral and analytic criteria for a harmonizable
S¹S process \( X = \{X_n\} \) to have positive angle or distance between past and future,
and we discuss its ramifications.

In view of stationarity, the location of the "present" is not important,
and thus the past \( P^X \) and future \( F^X \) of \( X \) are defined as the closure in probability
of the linear spans of \( \{X_n, n \leq 0\} \) and of \( \{X_n, n \geq 1\} \) respectively. We say
that past and future of \( X \) are at positive angle, or that \( X \) has positive angle,
if \( \rho(P^X, F^X) < 1 \) or \( \rho(F^X, P^X) < 1 \). We also say that past and future are at
positive distance, or that \( X \) has positive distance, if \( d(P^X, F^X) > 0 \). Finally
\( X \) is called minimal if \( X_n \) cannot be perfectly interpolated from \( \{X_k, k \neq n\} \), i.e.
if \( X_n \) does not belong to the closure in probability of the linear span of
\( \{X_k, k \neq n\} \).

Theorem 3. For a harmonizable S¹S process \( X = \{X_n\} \) with \( 1 < \omega < 2 \) and spectral
measure \( \mu \) the following are equivalent and imply that \( X \) is regular and
minimal.

1. \( X \) has positive angle: \( \rho(P^X, F^X) < 1 \) or \( \rho(F^X, P^X) < 1 \).

2. \( X \) has positive distance: \( d(P^X, F^X) > 0 \).

3. \( \{X_n\} \) is a Schauder basis for \( M^X \).

4. \( d_{\mu}(\varnothing) = f(\varnothing) d\alpha, L^0(f) \subset L^1 \), and the Fourier series of every \( g \in L^0(f) \)
   converges to \( g \) in \( L^0(f) \).

5. \( d_{\mu}(\varnothing) = f(\varnothing) d\alpha \) and the spectral density \( f \) satisfies

   \[
   \left( \frac{1}{|I|} \int_I f(u) du \right) \left( \frac{1}{|I|} \int_I f(u)^{\alpha-1} du \right)^{\alpha-1} \leq k
   \]

   for some constant \( k \) and all intervals \( I \) with length \( |I| \) (which are
   allowed to wrap around \( \pm \pi \)).
(6) The conjugation operator, considered on real trigonometric polynomials, is bounded in $L'(\cdot)$.

The first three equivalent conditions are time domain conditions, while the last three are frequency domain conditions. The equivalence of the spectral conditions (4), (5) and of (6) with $\mu$ absolutely continuous, is a well known result in Hunt et al. (1973). Here we provide a simple proof of the equivalence of the weaker condition (6) (where $\mu$ is not assumed absolutely continuous) with (4) via the time domain criterion given in Corollary 2, while of course the proof in Hunt et al. (1973) is analytic. Let us recall that the conjugate of a Fourier series $\sum a_n e^{in\theta}$ is defined by $\sum_{n \neq 0} -i \text{sgn}(n) a_n e^{in\theta}$.

In the proof of Theorem 3 use will be made of the following property which is valid in general normed linear spaces and says that two subspaces are at a positive distance if and only if the algebraic projection from their algebraic sum onto either subspace is a bounded operator.

**Proposition 11.** If $M$ and $N$ are subspaces of a normed linear space, the following are equivalent.

(i) $d(M,N) = \inf \{ \|X-Y\| : \|X\| = 1 = \|Y\|, X \in M, Y \in N \} > 0$.

(ii) There is a constant $k$, such that $\|X\| \leq k \|X+Y\|$ for all $X \in M, Y \in N$.

**Proof.** (ii) clearly implies (i), with the $\inf \geq k^{-1}$. We now show that "not (ii)" implies "not (i)". Assume (ii) is not satisfied. Then there are $x_n \in M, y_n \in N$ such that $0 < n \|x_n - y_n\| < \|x_n\|$. It follows that $\|y_n\| > 0$ for $n \geq 2$, and
and hence (i) is also not satisfied.

We now obtain the following useful result.

Corollary 2. With $X$ as in Theorem 3 the following are equivalent.

(i) $X$ has positive distance: $d(P_X, F_X) > 0$.

(ii) There is a constant $k$ such that

$$\left| \sum_{n=-k}^{m} c_n x_n \right|_\alpha \leq k \left| \sum_{n=-k}^{m} c_n x_n \right|_\alpha$$

for all $0 < k < m$, $0 < m' < m$, and complex numbers $c_n$ (and we may take $k = 0$ or $m = 0$).

(iii) There is a constant $k$ such that

$$\left| \sum_{n=k}^{m'} c_n x_n \right|_\alpha \leq k \left| \sum_{n=k}^{m'} c_n x_n \right|_\alpha$$

for all $k < k' < m' < m$.

Proof. The equivalence of (i) and (ii) is an immediate consequence of Proposition 11, and the equivalence of (ii) and (iii) follows from the stationarity of $X$.

Proof of Theorem 3. The equivalence of (1) and (2) is shown in Proposition 3. The equivalence of (2) and (3) follows from Corollary 2 and the fact that (iii) in Corollary 2 is a characterization of a two-sided Schauder basis, cf. Singer (1970).
We next show that (3) implies that $X$ is regular and minimal. An argument similar to the one below has been used in Miane and Niemi (1985).

Assume (3). To show that $X$ is minimal: $X_n \neq \text{sp}\{X_k, k \neq n\}$, it suffices to show that $\text{sp}\{X_k, k \neq n\} \neq M^X$. Assume on the contrary that $\text{sp}\{X_k, k \neq n\} = M^X$ for some $n$, and hence by stationarity for all $n$, so that $\cap_n \text{sp}\{X_k, k \neq n\} = M^X$.

In fact we will show that $\cap_n \text{sp}\{X_k, k \neq n\} = \{0\}$, namely that $X$ is $J_0$-regular. Indeed if $Y \in \cap_n \text{sp}\{X_k, k \neq n\}$ then by (3) it can be written uniquely as $Y = \sum_n c_n X_n$, and since for each $n$, $Y \in \text{sp}\{X_k, k \neq n\}$ we have $c_n = 0$ and thus $Y = 0$.

Hence $X$ is $J_0$-regular, and thus minimal as well as regular, since $\cap_n \text{sp}\{X_k, k \neq n\} \subset \cap_n \text{sp}\{X_k, k \neq n\} = \{0\}$.

Now we show (3) $\iff$ (4). First assume (3). Then $X$ is regular and by Theorem 1, we have $d_\mu(\theta) = f_\theta d\mu$. Since $X$ is also minimal, it follows from Theorem 3.3 in Pourahmadi (1984) that $f_{\theta}^{-\frac{1}{\alpha-1}} \in L^1$, and thus if $g \in L^\alpha(f)$, by Hölder's inequality,

$$\int |g| = \int |g| f_\theta^{\frac{1}{\alpha-1}} \leq \left( \int |g| f_\theta \right)^{\frac{1}{\alpha}} \left( \int f_\theta^{-\frac{1}{\alpha-1}} \right)^{\frac{\alpha-1}{\alpha}} < \infty$$

and hence $g \in L^1$. Thus $L^\alpha(f) \subset L^1$. Now by (3), every $Y \in M^X$ has a unique representation $Y = \sum_n c_n X_n$ in $M^X$. Using the linear isomorphism $Y = g dZ \leftrightarrow g$ between $M^X$ and $L^\alpha(f)$, it follows that every $g \in L^\alpha(f)$ has a unique representation $a(\cdot) = \sum_n c_n e^{-in\cdot}$ in $L^\alpha(f)$. But from the above displayed inequality the convergence is also in $L^1$, from which it follows that $c_n$ is in fact the $n$-th Fourier coefficient of $g$. Thus the Fourier series of every $g \in L^\alpha(f)$ converges to $g$ in $L^\alpha(f)$, and (4) is shown. Conversely assume (4) is satisfied. Fix $Y \in M^X$. Then $Y = g dZ$ for some $g \in L^\alpha(f)$, and by (4), $a(\cdot) = \sum_n c_n e^{-in\cdot}$ in $L^\alpha(f)$. It follows that $Y = \sum_n c_n X_n$ in $M^X$. We now show this representation of $Y$ is unique. Assume we also have $Y = \sum_n c'_n X_n$. Then $g(\cdot) = \sum_n c'_n e^{-in\cdot}$ in $L^\alpha(f)$.
Using Lemma 3.1 in Miamee (1985), it follows that this convergence is also in $L^1$ and hence $a_n = \hat{g}_n$, the $n$-th Fourier coefficient of $g$. Thus every $Y \in M^X$ has a unique representation $Y = \sum_n \hat{g}_n X_n$, showing (3).

The equivalence of (4), (5) and (6), with $d_\mu(\theta) = f(\theta)d\theta$ added on is established in Hunt et al. (1973), Theorem 1. Here we shall show that the weaker statement in (6) is equivalent to (2). An argument as in Helson and Szégo (1960), pp. 129-130, shows that (6) is equivalent to the boundedness of the truncation operator $T$ from $L^\infty(\mu)$ into itself defined by

$$T(\sum_{n=-\infty}^{\infty} c_n e^{in\theta}) = \sum_{n>0} c_n e^{in\theta},$$

which, considering the isomorphism $e^{-in\theta} \leftrightarrow X_n$, is equivalent to part (ii) of Proposition 11 and hence $d(P_X,F_X) > 0$, i.e. (2) which completes the proof.

Condition (3) is the crucial one. It means that every r.v. $Y$ in the linear space $M^X$ of the sequence $X = \{X_n\}$ can be written uniquely as a converging series in terms of the r.v.'s $X_n$: $Y = \sum_n b_n X_n$. Thus every linear estimator based on an observed part of $X$ can be realized by a unique linear filter acting on $X$. In particular, under any of the equivalent conditions of Theorem 3, which are stronger than those in Theorem 1, the moving average representation of Theorem 1 can be inverted to express the sequence of innovations $\{V_n\}$ as a convergent series

$$(1) \quad V_n = \sum_{k=0}^{\infty} b_k X_{n-k},$$

and the $m$-step ahead linear predictor $\hat{X}_{n+m,m}$ of $X_{n+m}$ based on $X_k$, $k \leq n$, can be written in the form

$$(P_m) \quad \hat{X}_{n+m,n} = \sum_{k=0}^{\infty} c_{m,k} X_{n-k}.$$
i.e. it can be realized by the filter \( \ell_m = \ell_m \forall k \geq 0 \) acting on the observed part of \( X \). These series converge with respect to the norm \( \| \cdot \|_{\alpha} \), or equivalently in \( L^p(\cdot) \). While condition \( (A_\alpha) \) guarantees that all functions in \( L^\alpha(f) \) have convergent Fourier series, and all random variables in \( M^X \) can be written as convergent series in terms of the r.v.'s \( X_n \), substantially weaker conditions can be found, in between those in Theorems 1 and 3, which are sufficient for the innovations to have a convergent series representation \( (I) \) in terms of the observed values of the process \( X \) itself. However we postpone a discussion on this to the end of this section, in order to first explore the relationship between the existence of \( (I) \) and the existence of auto-regressive representations of the predictors.

The metric predictor of a harmonizable symmetric processes has been considered in Hosoya (1982) and Cambanis and Soltani (1984) and the one step ahead metric predictor \( \hat{X}_{n+1,n} \) has been obtained. In terms of our results here, the one step ahead metric predictor can be written as

\[
\hat{X}_{n+1,n} = \sum_{k=1}^{\infty} a_k V_{n+1-k}^{-1}
\]

(cf. Theorem 1.5)). The problem of obtaining the m-step ahead metric predictor \( \hat{X}_{n+m,n} \) in the general case is still open, cf. Cambanis and Soltani (1984) for more details.

Now we consider the right angle m-step predictor \( \hat{X}^r_{n+m,n} \) which is the right angle projection of \( X_{n+m} \) on \( M^X_n \):

\[
\hat{X}^r_{n+m,n} = a_r(X_{n+m}; M^X_n).
\]

While this right angle predictor may not exist (Proposition 7(i)), the following proposition shows that, when it exists, it is in fact the truncation of the moving average given in Theorem 1, extending to this predictor a nice
property from the Gaussian case.

**Proposition 12.** (i) If the \( m \)-step right angle predictor \( \hat{X}_{n+m,n}^r \) exists then it is given by

\[
\hat{X}_{n+m,n}^r = \sum_{k=m}^{\infty} a_k V_{n+m-k}.
\]

(ii) If the regression \( E(X_{n+m} | M^X) \) is linear then the right angle predictor \( \hat{X}_{m+n,n}^r \) exists and we have

\[
\hat{X}_{n+m,n}^r = E(X_{n+m} | M^X) = \sum_{k=m}^{\infty} a_k V_{n+m-k}.
\]

**Proof.** (i) If the \( m \)-step ahead right angle predictor \( \hat{X}_{n+m,n}^r \) exists then it is in \( M^X_n \) and hence in \( M^V_n \). Now Theorem 3, applied to the innovation process \( V_n \), shows that \( V_n \) is a Schauder basis for \( M^X_n \). (This is because the density of \( V \) is simply the Lebesgue measure, which clearly satisfies the \((A_\alpha)\) condition of Theorem 3). Thus one can write \( \hat{X}_{n+m,n}^r \) as a convergent series

\[
\hat{X}_{n+m,n}^r = \sum_{k=m}^{\infty} c_k V_{n+m-k}.
\]

Now by the definition of right angle projection we have

\[
[\hat{X}_{n+m,n}^r, Y] = [X_{n+m}, Y]
\]

for every \( Y \) in \( M^X_n \) and in particular for \( Y = V_k \), with \( k > n \); i.e. we have

\[
[\hat{X}_{n+m,n}^r, V_k] = [X_{n+m}, V_k], \quad k > n.
\]

This shows that

\[
a_i = c_i, \quad \text{for all} \quad i > m.
\]

which completes the proof of (i). The proof of (ii) is now immediate from part (ii) of Proposition 7.
The last proposition shows that when the right angle predictor exists it can be obtained through a filter exactly similar to the standard one in the Gaussian case. This can be used to show that the problem of inverting the moving average representation

\[ X_n = \sum_{k=0}^{\infty} a_k V_{n-k} \]

to obtain a moving average representation for the innovations

(I) \[ V_n = \sum_{k=0}^{\infty} b_k X_{n-k} \]
is equivalent to the existence of a series representation of the one step ahead metric predictors \( \hat{X}_{n+1,1} \) in terms of the observed values of \( X \) itself

(R1) \[ \hat{X}_{n+1,1} = \sum_{k=0}^{\infty} d_k n-k \]

and this is equivalent to the series representation of all the existing right angle predictors \( \hat{X}_{n+m,n} \) in terms of the observed values of \( X \) itself

(APm) \[ \hat{X}_{n+m,n} = \sum_{k=0}^{\infty} e_{k,n} X_{n-k} \]

The equivalence of (I) and (R1) can be established in the time domain. Indeed assume that (I) holds. The orthogonality of \( V_n \)'s implies that for every \( Y \) in \( M^X_n = M_n \) we have \( [Y, V_{n+1}]_\alpha = 0 \). Using (4) of Theorem 1, we obtain

\[ [V_{n+1}, V_{n+1}]_\alpha = [ \sum_{k=0}^{\infty} b_k X_{n+1-k}, V_{n+1}]_\alpha = b_0 [X_{n+1}, V_{n+1}]_\alpha \]

\[ = b_0 [ \sum_{k=0}^{\infty} a_k X_{n+1-k}, V_{n+1}]_\alpha = b_0 a_0 [V_{n+1}, V_{n+1}]_\alpha \]

and thus \( b_0 = a_0^{-1} > 0 \). Now it follows from \( b_0 X_{n+1} = V_{n+1} - \sum_{k=1}^{\infty} b_k X_{n+1-k} \)

that \( X_{n+1,n} = b_0^{-1} \sum_{k=1}^{\infty} b_k X_{n+1-k} = \sum_{k=0}^{\infty} (-b_{k+1}/b_0) X_{n-k} \) and thus (P1) is satisfied.
Conversely, assume \((P_1)\) is satisfied. From (3) and (4) of Theorem 1, we have

\[ X_{n+1} - X_{n+1,n} = a_0 V_{n+1}, \]

and thus by \((P_1)\),

\[ V_{n+1} = a_0^{-1} (X_{n+1} - \sum_{k=0}^{\infty} d_{k-1} X_{n-k}) = \sum_{k=0}^{\infty} b_k X_{n+1-k}, \]

with \(b_0 = a_0^{-1}\) and \(b_k = -d_{k-1}/d_0, \ k > 1\). So (I) holds. The equivalence of (I) and \((AP_m)\) follows from an appropriate adjustment in the proof the corresponding fact for the second order case, as given in Bloomfield (1904), together with the representation of \(\hat{X}_{n+m,n}\) given in part (i) of Proposition 12.

Considering the isomorphism between the time domain and spectral domain we see that a necessary and sufficient condition for (I) to hold is that \(\phi^{-1}\) has a series expansion

\[ \phi^{-1}(r) = \sum_{k=0}^{\infty} a_k e^{-ik\theta}, \]

converging in \(L_1(f)\).

While condition (F) is necessary and sufficient for the convergent series representations (I) and \((P_1)\) of interest to us here, it is not easily checked (and no easily checked necessary and sufficient condition is available even when \(m=2\)). Following are some sufficient conditions which are easier to check. The simplest is the one suggested by Masani (1960):

\[ f \in L^\infty \text{ and } f^{-1} \in L^1. \]

A different condition is given in Theorem 3: \((A_\alpha)\). The fact that \((A_\alpha)\) implies the convergent series representation \((P_1)\) has also been shown in Pourahmadi (1985). A weaker condition, generalizing both conditions \((A_\alpha)\) and \((M)\), can be proved similarly to Theorem 4 in Bloomfield (1984), where the
case $\alpha = 2$ is considered:

$$(B) \quad f = hq \text{ where } h \text{ satisfies } (A_\alpha) \text{ and } g > 0 \text{ satisfies } (M).$$

The following are yet weaker conditions.

**Proposition 13.** Let $X$ be regular harmonizable $S\alpha S$ with $1 < \alpha < 2$, and let $\phi$ be the outer factor of $f$ (cf. Theorem 1.3). Then any of the following conditions implies (F).

(a) $f = h_1g_1 + h_2g_2$ where $h_1,g_1 > 0$, $h_1$ satisfies condition $(A_\alpha)$, and $g_1$ satisfies condition $(M)$, $i = 1,2$.

(b) $g_1h_1 - f \leq g_2h_2$ where $h_1,g_1 > 0$, $g_1^{-1}$, $L^1$, $g_2 : L^\alpha$, $L^\alpha(h_1) = L^\alpha(h_2)$ and $h_1$ satisfies $(A_\alpha)$.

**Proof.** (a) Clearly $f = h_1g_1$, for $i = 1,2$, so $f^{-1} \leq (h_1g_1)^{-1}$, and hence

$$\int h_1 = f^{-1}g_1^{-1}.$$

Thus $f^{-1} \in L^\alpha(h_1)$. Now since $h_1$ satisfies $(A_\alpha)$ by Theorem 3 we see that $(\cdot^{-1})^N$, the $N$-th Fourier partial sum of $\cdot^{-1}$, converges to $\cdot^{-1}$ in $L^\alpha(h_1)$ and hence in $L^\alpha(h_1,g_1)$ (because $g_1 \in L^\alpha$), i.e.

$$|\cdot^{-1}f_0|^N - \cdot^{-1}|f_0|g_1 > 0, \quad i = 1,2.$$

Adding these two together we get

$$|\cdot^{-1}f_0|^N - \cdot^{-1}|f_0|g_1 > 0,$$

which completes the proof of (a). (b) can be proved by adjusting the proof in Bloomfield (1985).
As an application of Proposition 13 one can verify that a second order stationary stochastic process with spectral density

\[ f(\theta) = |1 + e^{i\theta}|^{1.5} + |1 + e^{i\theta}|^{0.5} \]

has the representations (I), (P) and (Q). We know that \(|1 + e^{i\theta}|^p\) satisfies (A) for \(-1 < p < 1\), by Helson and Szegö (1960), and (M) for \(0 < p < 1\). Thus we can, for example, take \(g_1 = |1 + e^{i\theta}|^{0.6}\), \(g_2 = |1 + e^{i\theta}|^{0.5}\), \(h_1 = |1 + e^{i\theta}|^{0.9}\) and \(h_2 = 1\).
REFERENCES


