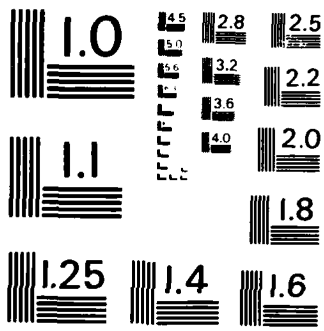
AD-A161 412 ON PREDICTION OF HARMONIZABLE STABLE PROCESSES(U) NORTH CAROLINA UNIV AT CHAPEL HILL CENTER FOR STOCHASTIC PROCESSES S CAMBANIS ET AL. JUL 85 TR-110 1/1

UNCLASSIFIED AFOSR-TR-85-0950 F49620-82-C-0009 F/G 12/1 NL

END

FILMED

DTIC


1.0
4.5
5.0
5.6
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

(2)

AD-A161 412

CENTER FOR STOCHASTIC PROCESSES

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

ON PREDICTION OF HARMONIZABLE STABLE PROCESSES

by

S. Cambanis and A.G. Miamee

Technical Report No. 110

July 1985

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT UNLIMITED
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Rept. No. 110	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR-
6a. NAME OF PERFORMING ORGANIZATION Center for Stochastic Processes; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR
6c. ADDRESS (City, State and ZIP Code) Statistics Dept., 321 PH 039A, UNC Chapel Hill, NC 27514	7b. ADDRESS (City, State and ZIP Code) Bldg 410 Bolling AFB DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-82-C-0009
8c. ADDRESS (City, State and ZIP Code) Bolling AFB Washington, DC 20332	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 61102F; PROJECT NO.: 2304; TASK NO.: A5; WORK UNIT NO.:
11. TITLE (Include Security Classification) ON PREDICTION OF HARMONIZABLE STABLE PROCESSES	

12. PERSONAL AUTHOR(S)
S. Cambanis and A.G. Miamee

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
technical	FROM 9/84 TO 8/85	July 1985	43

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Harmonizable stable processes, moving average representation, linear predictors, angle, isotropic complex random variables

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

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.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED X SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Brian W Woodruff Peggy Ravitch	22b. TELEPHONE NUMBER (Include Area Code) 919 962-2307 (202) 767-5027; 22c. OFFICE SYMBOL NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE

unclassified

SECURITY CLASSIFICATION OF THIS PAGE

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.

AMS Subject Classification (1980). Primary 60G10, 60G25. Secondary 42A20, 42A50.

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 unwielding 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), Miamee 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 SαS process $X = \{X_n, n=0,\pm 1,\dots\}$ with spectral measure a finite measure μ on $(-\pi,\pi]$ is defined through its finite dimensional characteristic functions

$$E \exp\{i \operatorname{Re} \sum_{n=k}^{j} z_n X_n\} = \exp\{-\int_{-\pi}^{\pi} |\sum_{n=k}^{j} z_n e^{-in\theta}|^\alpha d\mu(\theta)\}.$$

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 SαS measure on the Borel subsets of $(-\pi,\pi]$ with

$$E \exp\{i \operatorname{Re} \int_{-\pi}^{\pi} f dZ\} = \exp\{-\int_{-\pi}^{\pi} |f|^\alpha d\mu\}$$

for all $f \in L^\alpha(\mu)$ (see Hosoya (1982), Cambanis (1983)). The correspondance $f \to \int_{-\pi}^{\pi} f dZ$ is an isomorphism between $L^\alpha(\mu)$ 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 $L^\alpha(\mu)$, 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 $(r-is)f$ we have

$$E \exp\{i(r \operatorname{Re}\int f dZ + s \operatorname{Im}\int f dZ)\} = \exp\{-(r^2+s^2)^{\alpha/2}\int |f|^\alpha d\mu\}.$$

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 SαS 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 SαS if X_1 and X_2 are jointly SαS with radially symmetric distribution, i.e. $E \exp\{i(r_1X_1+r_2X_2)\} = = \exp\{-c(r_1^2 + r_2^2)^{\alpha/2}\}$, or in complex notation with $r = r_1 + ir_2$,

$$E \exp\{i \operatorname{Re} \bar{r}X\} = \exp\{-c|r|^{\alpha}\}.$$

A complex process $X = \{X_t,\ t \in T\}$ is called isotropic SαS if every finite complex linear combination $\sum z_n X_{t_n}$ is a complex isotropic SαS r.v. Then there exists a measure space (I, Σ, μ) and complex functions $f_t \in L^{\alpha}(\mu)$, $t \in T$, such that

$$E \exp\{i \operatorname{Re} \bar{r} \sum_{n=1}^{N} z_n X_{t_n}\} = \exp\{-|r|^{\alpha} \|\sum_{n=1}^{N} z_n f_{t_n}\|_{\alpha}^{\alpha}\}$$

where $\|f\|_{\alpha}$ denotes $\|f\|_{L^{\alpha}(\mu)}$ (Hardin (1982)). Equivalently, if Z is a complex, independently scattered, isotropic SαS measure on (I, Σ, μ), i.e. for all disjoint sets $I_1, \ldots, I_n \in \Sigma$ of finite μ-measure, $Z(I_1), \ldots, Z(I_n)$ are independent with $E \exp\{i \operatorname{Re} \bar{r}Z(I_k)\} = \exp\{-|r|^{\alpha}\mu(I_k)\}$, so that for all $f \in L^{\alpha}(\mu)$,

$$E \exp\{i \operatorname{Re} \bar{r} \int f dZ\} = \exp\{-|r|^{\alpha} \|f\|_{\alpha}^{\alpha}\},$$

then the stochastic process $\{\int_I f_t(s)dZ(s),\ t \in T\}$ 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 correspondance $X_t \leftrightarrow f_t$ extends to an isomorphism between M^X and the subspace $\overline{sp}\{f_t,\ t \in T\}$ of $L^{\alpha}(\mu)$. M^X is then a complex isotropic SαS space and every Y in M^X is represented by some f in $L^{\alpha}(\mu)$. (For general, not necessarily isotropic, complex SαS 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 $0<\alpha\leq 2$ and let $Y\in M^X$ be represented by $f\in L^\alpha(\mu)$. Then the pair (Re Y, Im Y) has the same distribution as the pair $\sqrt{2}\,\|f\|_\alpha R^{1/2}(N_1,N_2)$, where R is a positive $\alpha/2$ stable r.v. with $E\exp(-uR)=\exp(-u^{\alpha/2})$, $u>0$, and is independent of the iid standard normal r.v.'s N_1 and N_2. Moreover, for all $0<p<\alpha$,

$$(E|Y|^p)^{1/p}=\left\{\frac{p2^p\,\Gamma(p/2)\,\Gamma(-p/\alpha)}{\alpha\Gamma(-p/\alpha)}\right\}^{1/p}\|f\|_\alpha \triangleq C_{p,\alpha}\|f\|_\alpha .$$

(ii) Assume $1<\alpha\leq 2$ and let $Y_1, Y_2\in M^X$ be represented by $f_1,f_2\in L^\alpha(\mu)$. Then for every $1<p<\alpha$,

$$\frac{EY_1Y_2^{<p-1>}}{E|Y_2|^p}=\frac{\int f_1 f_2^{<\alpha-1>}d\mu}{\int|f_2|^\alpha d\mu},$$

where for $\beta>0$ and complex $z\neq 0$, $z^{<\beta>}=|z|^{\beta-1}\bar{z}$.

Proof. (i) Since $Y\in M^X$ we have $Y=\int f dZ$ for some $f\in L^\alpha(\mu)$. Putting $r=r_1+ir_2$ we obtain

$$E\exp\{i(r_1Y_1+r_2Y_2)\}=E\exp\{i\,\mathrm{Re}\,\bar{r}\int f dZ\}=\exp\{-(r_1^2+r_2^2)^{\alpha/2}\|f\|_\alpha^\alpha\}.$$

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

$$E\exp\{i\sqrt{2}\,\|f\|_\alpha R^{1/2}(r_1N_1+r_2N_2)\}=E\exp\{-\|f\|_\alpha^2R(r_1^2+r_2^2)\}=\exp\{-(r_1^2+r_2^2)^{\alpha/2}\|f\|_\alpha^\alpha\}.$$

(ii) The calculation is similar to that on page 357 of Köthe (1969). For complex numbers $z\neq 0$ and w, and real λ we have $\frac{d|z+\lambda w|^p}{d\lambda}\Big/_{\lambda=0}=p|z|^{p-2}\mathrm{Re}(zw)$ and thus

$$\frac{d}{d\lambda}\{|z+\lambda w|^p-i|z+\lambda iw|^p\}_{\lambda=0}=p|z|^{p-2}\{\mathrm{Re}(zw)-i\,\mathrm{Re}(ziw)\}$$

$$=p|z|^{p-2}zw=pz^{<p-1>}w.$$

Using part (i) we obtain

$$pEY_1Y_2^{<p-1>} = \frac{d}{d\varepsilon}\{E|Y_2+\varepsilon Y_1|^p - i\, E|Y_2+\varepsilon iY_1|^p\}_{\varepsilon=0}$$

$$= C^p_{p,\alpha}\frac{d}{d\varepsilon}\{\|f_2+\varepsilon f_1\|^p_\alpha - i\|f_2+\varepsilon if_1\|^p_\alpha\}_{\varepsilon=0}$$

$$= C^p_{p,\alpha}\, p\|f_2\|^{p-\alpha}_\alpha \int |f_2|^{\alpha-2}\{Re(\bar{f}_2f_1) - i\, Re(\bar{f}_2 if_1)\}d\mu$$

$$= p\, C^p_{p,\alpha}\|f_2\|^{p-\alpha}_\alpha \int f_1f_2^{<\alpha-1>}d\mu .$$

Coupled with $E|Y_2|^p = C^p_{p,\alpha}\|f_2\|^p_\alpha$, this establishes (ii).

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

Now putting $\|Y\|_\alpha = \|f\|_\alpha$ we have that $\|Y\|_\alpha^{1\wedge\alpha}$ defines a norm when $1\le\alpha\le2$ and a quasi-norm when $0<\alpha<1$ on M^X, which metrizes convergence in probability (Cambanis (1983)) and which, by (i) of Proposition 1, is equivalent to convergence in $L^p(\Omega)$, $0<p<\alpha$.

When $1<\alpha\le2$ and $Y_1, Y_2 \in M^X$ are represented by $f_1, f_2 \in L^\alpha(\mu)$, the covariation of Y_1 with Y_2 is defined by

$$[Y_1, Y_2]_\alpha = \int f_1f_2^{<\alpha-1>}d\mu$$

and by (ii) of Proposition 1 we have for all $1\le p<\alpha$ (provided $Y_2\neq0$)

$$\frac{[Y_1, Y_2]_\alpha}{\|Y_2\|^\alpha_\alpha} = \frac{EY_1Y_2^{<p-1>}}{E|Y_2|^p} .$$

By Hölder's inequality we have $|[Y_1,Y_2]_\alpha| \le \|f_1\|_\alpha \|f_2\|_\alpha^{\alpha-1} = \|Y_1\|_\alpha \|Y_2\|_\alpha^{\alpha-1}$ with equality if and only if $Y_1 = zY_2$ for some complex z. The covariation of a harmonizable process is

$$[X_n,X_m]_\alpha = \int_{-\pi}^{\pi} e^{-i(n-m)\lambda} d\mu(\lambda),$$

the familiar form of the covariance of a stationary process. In the Gaussian case $\alpha = 2$ 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 $[Y_1,Y_2]_\alpha = 0$ and $[Y_2,Y_1]_\alpha = 0$. When $[Y_1,Y_2]_\alpha = 0$ we say that Y_2 is orthogonal to Y_1, $Y_2 \perp Y_1$, 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 $L^p(\ )$, $1 \le p < \alpha$ (see Cambanis et al. (1985)) for a discussion in the real case). While independence and orthogonality are equivalent in the Gaussian case $\alpha=2$, when $1<\alpha<2$ independence implies mutual orthogonality but the converse is not generally true. This is because when $0<\alpha<2$, 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. [μ] (Cambanis (1983)), while mutual orthogonality merely means that $\int f_1 f_2^{<\alpha-1>} d\mu = 0 = \int f_2 f_1^{<\alpha-1>} d\mu$.

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

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

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

Proof. For any two jointly SαS complex r.v.'s Y_1,Y_2, it is shown in Cambanis (1983) that $E(Y_2|Y_1) = cY_1$ iff $[Y_2 - cY_1, \operatorname{Re}(\bar{z}Y_1)]_\alpha = 0$ for all complex z in which case $c = [Y_2,Y_1]_\alpha/[Y_1,Y_1]_\alpha$, i.e. iff

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

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

$$\int|f_1|^\alpha d\mu \cdot\int f_2(\operatorname{Re}\bar{z} f_1)^{<\alpha-1>}d\mu = \int f_2f_1^{<\alpha-1>}d\mu\cdot\int f_1(\operatorname{Re}\bar{z} f_1)^{<\alpha-1>}d\mu.$$

Now from

$$E \exp\{i \operatorname{Re} (\bar{z}_1Y_1 + \bar{z}_2Y_2)\} = \exp\{-\int|\bar{z}_1f_1 + \bar{z}_2f_2|^\alpha d\mu\}$$

$$= \exp\{-\int\left|\bar{z}_1|f_1| + \bar{z}_2f_2e^{-i\arg f_1}\right|^\alpha d\mu\}$$

it follows that (Y_1,Y_2) is also represented by $(|f_1|, f_2e^{-i\arg f_1})$. Thus without loss of generality we may take f_1 to be real, whence $\operatorname{Re}(\bar{z} f_1) = (\operatorname{Re}\bar{z})f_1$ 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 $1<\alpha\leq 2$ is as follows. We define a complex valued <u>cosine of the angle of</u> $Y_1\neq 0$

<u>with</u> $Y_2 \neq 0$ <u>by</u>

$$\cos_\alpha(Y_1,Y_2) \triangleq \left[\frac{Y_1}{\|Y_1\|_\alpha}, \frac{Y_2}{\|Y_2\|_\alpha} \right]_\alpha = \frac{[Y_1,Y_2]_\alpha}{\|Y_1\|_\alpha \|Y_2\|_\alpha^{\alpha-1}} .$$

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

$$\cos_\alpha(Y_1,Y_2) = \frac{EY_1Y_2^{<p-1>}}{(E|Y_1|^p)^{1/p}(E|Y_2|^p)^{\frac{p-1}{p}}} = E\left(\frac{Y_1}{\|Y_1\|_p} \cdot \frac{Y_2^{<p-1>}}{\|Y_2^{<p-1>}\|_p}\right) \triangleq \cos_p(Y_1,Y_2)$$

for all $1<p<\alpha$. Thus the cosine defined through covariation agrees with that defined through $L^p(\cdot)$ for all $1<p<\alpha$. 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 $Y_1 = zY_2$ 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*

$$\rho(N_1,N_2) = \sup\{|\cos(Y_1,Y_2)|: \quad Y_1 \in N_1, \quad Y_2 \in N_2\}$$

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

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

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

or by $d_p(N_1,N_2) = \inf\{(E|Y_1-Y_2|^p)^{\frac{1}{p}\wedge 1}: Y_1 \in N_1, Y_2 \in N_2\}$, $0<p<\alpha$. In view of Proposition 1.(i) we have $d_p(N_1,N_2) = \mathrm{Const}(p,\alpha)d_\alpha(N_1,N_2)$. We say that N_1 and N_2 are at positive distance and write $d(N_1,N_2) > 0$ if $d_\alpha(N_1,N_2) > 0$, or equivalently if $d_p(N_1,N_2) > 0$ for some $0<p<\alpha$. We now show that when $1<\alpha\leq 2$ 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 $1<\alpha\le 2$ and N_1,N_2 be two subspaces of M^X. The following are equivalent.

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

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

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

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)$\Leftrightarrow$(ii). Now to show (ii) $\Rightarrow$(i), it suffices to show that $d(N_1,N_2) = 0$ implies $\rho(N_1,N_2) = 1$. Assume $d(N_1,N_2) = 0$. Then there exist $Y_n\in N_1$, $Z_n\in N_2$, such that $||Y_n||_\alpha = 1 = ||Z_n||_\alpha$ and $||Y_n-Z_n||_\alpha \to 0$. By Hölder's inequality $|[Y_n-Z_n,Z_n]_\alpha| \le ||Y_n-Z_n||_\alpha \, ||Z_n||_\alpha^{\alpha-1}$. It follows that $[Y_n-Z_n, Z_n]_\alpha \to 0$ and thus $[Y_n, Z_n]_\alpha \to 1$, i.e. $\cos(Y_n,Z_n) \to 1$ and $\rho(N_1,N_2) = 1$.

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 $L^\alpha(\mu)$ which represent N_1,N_2. Then (i) implies that for all $f_1\in M_1$, $f_2\in M_2$ with $||f_1||_\alpha = 1 = ||f_2||_\alpha$, we have $||f_1-f_2||_\alpha \ge \varepsilon > 0$. By the uniform convexity of $L^\alpha(\mu)$ (cf. Köthe (1969), § 26.7]) there is $\delta = \delta(\varepsilon) > 0$ such that $||f_1+f_2||_\alpha \le 2(1-\delta)$. Thus for every $0<\lambda<1$ we have

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

$$\le \frac{1}{\lambda}((1-\lambda)||f_2||+\lambda||f_1+f_2||_\alpha-||f_2||_\alpha)$$

$$= ||f_1+f_2||_\alpha-||f_2||_\alpha = ||f_1+f_2||_\alpha-1 \le 1-2\delta.$$

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

$$\frac{d}{d\lambda}\,||f_2+\lambda f_1||_{\alpha/\lambda=0} = \mathrm{Re}\int f_1 f_2^{<\alpha-1>}d\mu.$$

It follows that

$$\mathrm{Re} \int f_1 f_2^{<\alpha-1>} d\mu \leq 1-2\delta.$$

Since this is true for every $f_1 \in M_1$ with $||f_1||_\alpha = 1$, replacing f_1 by $f_1 \exp\{-i \arg (\int f_1 f_2^{<\alpha-1>} d\mu)\}$ we obtain

$$\left|\int f_1 f_2^{<\alpha-1>} d\mu\right| \leq 1-2\delta.$$

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

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

$$\begin{aligned}
&[n, Y-P(Y|N)]_\alpha = 0, && \text{for all } n \in N,\\
&[n,Y]_\alpha = [n,P(Y|N)]_\alpha, && \text{for all } n \in N,\\
&[Y,n]_\alpha = [P(Y|N),n]_\alpha, && \text{for all } n \in N,\\
&[Y-P(Y|N),n]_\alpha = 0, && \text{for all } n \in N.
\end{aligned}$$

When $1<\alpha<2$ 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 $1<\alpha<2$ there are three possible ways of defining projection, via the first three conditions above. The first condition leads to the metric projection $m(Y|N)$, which is the unique element in N minimizing the distance to Y from N, in any of the applicable metrics discussed above, for instance

$$||Y-m(Y|N)||_\alpha = \inf_{n \in N} ||Y-n||_\alpha,$$

and is uniquely determined by

$$[n, Y-m(Y|N)]_\alpha = 0 \quad \text{for all } n \in N,$$

(cf. Singer (1970a)). Using the second and third conditions we define the left, resp. right, angle projection of Y onto N as an element $a_\ell(Y|N)$, resp. $a_r(Y|N)$, of N which satisfies

$$[n,Y]_\alpha = [n, a_\ell(Y|N)]_\alpha \quad \text{for all } n \in N,$$

resp.

$$[Y,n]_\alpha = [a_r(Y|N),n]_\alpha \quad \text{for all } n \in N.$$

The following properties further justify the terminology used.

Proposition 4. If $a_\ell(Y|N)$, resp. $a_r(Y|N)$, exists then $\|a_\ell(Y|N)\|_\alpha \le \|Y\|_\alpha$, resp. $\|a_r(Y|N)\|_\alpha \le \|Y\|_\alpha$, and if moreover $a_\ell(Y|N) \neq 0$, resp. $a_r(Y|N) \neq 0$, then the left, resp. right, angle projection direction minimizes the left, resp. right, angle, i.e.

$$\sup_{\substack{n \in N \\ \|n\|_\alpha = 1}} |\cos(n,Y)| = \left|\cos\left(\frac{a_\ell(Y|N)}{\|a_\ell(Y|N)\|_\alpha}, Y\right)\right| = \frac{\|a_\ell(Y|N)\|_\alpha^{\alpha-1}}{\|Y\|_\alpha^{\alpha-1}},$$

resp.

$$\sup_{\substack{n \in N \\ \|n\|_\alpha = 1}} |\cos(Y,n)| = \left|\cos\left(Y, \frac{a_r(Y|N)}{\|a_r(Y|N)\|_\alpha}\right)\right| = \frac{\|a_r(Y|N)\|_\alpha}{\|Y\|_\alpha}.$$

Proof. If $n \in N$, $\|n\|_\alpha = 1$, and $a_\ell(Y|N) \neq 0$ we have

$$\cos(n,Y) = \frac{[n,Y]_\alpha}{\|Y\|_\alpha^{\alpha-1}} = \frac{[n,a_\ell(Y|N)]_\alpha}{\|Y\|_\alpha^{\alpha-1}} = \frac{\|a_\ell(Y|N)\|_\alpha^{\alpha-1}}{\|Y\|_\alpha^{\alpha-1}} \cos\left(n, \frac{a_\ell(Y|N)}{\|a_\ell(Y|N)\|_\alpha}\right)$$

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.

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

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

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

$$\sup_{\substack{n \in N \\ ||n||_\alpha = 1}} |\cos(n,Y)| = |\cos(\delta_\ell,Y)|,$$

(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 δ_ℓ that

$$a_\ell(Y|N) = [\delta_\ell,Y]^{\frac{1}{\alpha-1}}\,\delta_\ell$$

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

Let $f \in L^\alpha(\mu)$ and the subspace M of $L^\alpha(\mu)$ represent Y and N, and let g represent $n \in N$ with $||n||_\alpha = 1$. We have $\cos(n,Y) = \int g f^{<\alpha-1>} d\mu \cdot ||f||_\alpha^{1-\alpha}$, and thus to show the existence and uniqueness of δ_ℓ it is equivalent to show that there is a unique $g_\ell \in M$ with $||g_\ell||_\alpha = 1$ such that

$$S = \sup \{ |\int g f^{<\alpha-1>} d\mu| : g \in M, \|g\|_\alpha = 1 \} = |\int g_\ell f^{<\alpha-1>} d\mu|.$$

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

$$\frac{1}{2} \int (h_1+h_2) f^{<\alpha-1>} d\mu = S.$$

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

$$\int h f^{<\alpha-1>} d\mu = \frac{1}{\|h_1+h_2\|_\alpha} \int (h_1+h_2) f^{<\alpha-1>} d\mu \geq \frac{S}{1-\delta} > S$$

contradicting the definition of the supremum S. Hence the uniqueness of g_ℓ is established.

The unique maximizing element g_ℓ must satisfy $(d/d\varepsilon)F(g_\ell + \varepsilon g)_{/\varepsilon=0}$ for all $g \in M$, where

$$F(g) = \int g f^{<\alpha-1>} d\mu + \lambda\left(\int |g|^\alpha d\mu - 1\right),$$

(see Luenberger (1969), pp. 188-189) i.e. $\int g f^{<\alpha-1>} d\mu + \lambda\alpha \int g g_\ell^{<\alpha-1>} d\mu = 0$, $g \in M$. Putting $g=g_\ell$, we find $\lambda\alpha = -\int g_\ell f^{<\alpha-1>} d\mu$, and thus the condition becomes

$$\int g f^{<\alpha-1>} d\mu - \int g g_\ell^{<\alpha-1>} d\mu \cdot \int g_\ell f^{<\alpha-1>} d\mu = 0 \quad \text{for all } g \in M.$$

Expressing it in terms of the space M^X, this condition becomes $[n,Y]_\alpha - [n,\lambda_\ell]_\alpha [\lambda_\ell, Y]_\alpha = 0$ for all $n \in N$. 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 $Y \in M^X \setminus N$. Then $E(Y|N) \in N$ if and only if there exists $\bar{Y} \in N$ such that $[Y-\bar{Y},Z]_\alpha = 0$ for all $Z \in N$ and then $\bar{Y} = E(Y|N)$.

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

$$\phi(r_1,r_2) = E\exp\{i\,\mathrm{Re}(\bar{r}Y + Z)\} = E\exp\{i(r_1Y_1+r_2Y_2 + Z_1)\}$$

$$= \exp\{-\int |\bar{r}f + g|^\alpha d\mu\}.$$

Then

$$E\{e^{i\mathrm{Re}(Z)}Y\} = -i\left(\frac{\partial\phi}{\partial r_1} + i\frac{\partial\phi}{\partial r_2}\right)_{r_1=0=r_2}$$

$$= i\alpha\exp(-\int|g|^\alpha d\mu)\{\int|g|^{\alpha-2}\mathrm{Re}(\bar{g}f)d\mu + i\int|g|^{\alpha-2}\mathrm{Re}(-i\bar{g}f)d\mu\}$$

$$= i\alpha\exp(-\int|g|^\alpha d\mu)\cdot\int f g^{<\alpha-1>} d\mu$$

$$= i\alpha\exp(-\|Z\|_\alpha^\alpha)\cdot[Y,Z]_\alpha$$

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

$$E\{e^{i\,\mathrm{Re}(Z)}[E(Y|N)-\bar{Y}]\} = E\{e^{i\,\mathrm{Re}(Z)}(Y-\bar{Y})\}$$

$$= i\alpha\exp(-\|Z\|_\alpha^\alpha)\cdot[Y-\bar{Y},Z]_\alpha.$$

Now $E(Y|N)\in N$ iff for some $\bar{Y}\in N$ we have $E(Y|N) = \bar{Y}$ or equivalently LHS = 0 for all $Z\in N$, i.e. RHS = 0 for all $Z\in N$. $\square$

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 $\alpha=2$, 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 $Y\in M^X\setminus N$, and $1<\alpha<2$.

(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 $I = [0,1]$, μ = Lebesgue, $f_1 = 1_{[0,2/3]}$, $f_2 = 1_{[1/3,1]}$, $f = 1_{[0,1]}$, $Y = \int f dZ$, $Y_i = \int f_i dZ$, $i=1,2$, $N = sp\{Y_1,Y_2\}$. 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 $[Y,n]_\alpha = [a_r,n]$ for all $n\in N$, i.e. for all $n = xY_1 + yY_2 = \int(xf_1 + yf_2)dZ$. Thus a,b must satisfy

$$\int_0^1 (af_1 + bf_2)(xf_1 + yf_2)^{<\alpha-1>} = \int_0^1 (xf_1 + yf_2)^{<\alpha-1>} \quad \text{for all } x,y.$$

Putting x=0 and y=0 gives a=b=2/3, and then putting x=y gives the contradiction $2^\alpha=4$. A genuinely complex isotropic example can be provided by taking $I = (-\pi,\pi]$, μ = Lebesgue, $Y_k = \int_{-\pi}^{\pi} e^{-ik\theta}dZ(\theta)$, k=0,1,2, $N = sp\{Y_0,Y_1\}$, $Y = Y_0+Y_1+Y_2$, and reaching a contradiction likewise (using a property shown in Example 4.5 in Cambanis et al. (1985)).

(ii) Suppose there are $a_1, a_2 \in N$ such that for all $n \in N$: $[Y,n]_\alpha = [a_1,n]_\alpha$ and $[Y,n]_\alpha = [a_2,n]_\alpha$. Then $[a_1,n]_\alpha = [a_2,n]_\alpha$ and $[a_1-a_2,n]_\alpha=0$ for all $n \in N$. Taking $n=a_1-a_2$ gives $\|a_1-a_2\|^\alpha = 0$ and thus $a_1=a_2$.

(iii) If $E(Y|N) \in N$, by Proposition 6 we have for all $n \in N$, $[Y-E(Y|N),n]_\alpha=0$, i.e. $[Y,n]_\alpha =[E(Y|N),n]_\alpha$, and thus by (ii), $a_r(Y|N) = E(Y|N)$.

The following examples show that in the non-Gaussian stable case $1<\alpha<2$, 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 $I = [0,1]$, μ = Lebesgue, $1<\alpha<2$, $Y = \int 1_{(0,2/3)}dZ$, $W = \int 1_{(0,1)}dZ$, $N = sp\{W\}$. It is easily seen that

$$a_r(Y|W) = E(Y|W) = \frac{2}{3}W, \quad a_\ell(Y|W) = \left(\frac{2}{3}\right)^{\frac{1}{\alpha-1}} W, \quad m(Y|W) = \frac{1}{1 + 2^{\frac{1}{\alpha-1}}} W,$$

and hence they are all distinct.

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

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

$$a_\ell(Y|Y_1,Y_2) = Y_1 + \left(\frac{1}{3}\right)^{\frac{1}{\alpha-1}} Y_2,$$

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

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 $X = \{X_t,\ t \in T\}$ is a so-called α-sub-Gaussian process, i.e. $X_t = A^{1/2}G_t,\ t \in T$, where $G = \{G_t,\ t \in T\}$ is a zero-mean Gaussian process independent of the positive $\alpha/2$-stable r.v. A, with $E \exp(-uA) = \exp(-u^{\alpha/2})$, $u>0$.

Proposition 8. Let X be α-sub-Gaussian with $1<\alpha<2$, N a subspace of M^X and $Y \in M^X\ N$. Then

$$m(Y|N) = a_r(Y|N) = E(Y|N) = c\, a_\ell(Y|N)$$

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

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

$$[Y_1,Y_2]_\alpha = \frac{E(W_1 W_2)}{2^{\alpha/2}[EW_2^2]^{1-\alpha/2}}, \quad \|Y\|_\alpha = (\tfrac{1}{2}EW^2)^{1/2}.$$

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

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

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), $a_r(Y|N) = E(Y|N)$. We also see that

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

$$= \left(\frac{EW^2}{E|E(W|L)|^2}\right)^{1-\alpha/2} \frac{E(\ell W)}{2^{\alpha/2}[EW^2]^{1-\alpha/2}} =: c^{-1}[n,Y]_\alpha,$$

from which it follows that $a_\ell(Y|N) = c^{-1}m(Y|N)$.

3. Regularity and orthogonal moving average representation.

In this section we obtain criteria for regularity of harmonizable SαS processes $X = \{X_n\}$, 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 $\|\cdot\|_\alpha$ norm or in $L_p(\Omega)$ norm (cf. Proposition 1), of the linear span of $\{X_k, k\le n\}$, resp. $\{X_k, k<+\infty\}$. The remote past of X is the subspace $M_{-\infty}^X = \cap_n M_n^X$. X is called regular if $M_{-\infty}^X = \{0\}$ and singular if $M_{-\infty}^X = M^X$. H^α 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 SαS process X with $1<\alpha\le 2$ and spectral measure μ, the following are equivalent.

(1) X is regular.

(2) $d\mu(\theta) = f(\theta)d\theta$ and $\int_{-\pi}^{\pi} \log f(\theta)\, d\theta > -\infty$.

(3) $d\mu(\theta) = f(\theta)\, d\theta$ and $f(\cdot) = |\phi(\cdot)|^\alpha$ where $\phi \in H^\alpha$.

(4) X has a moving average representation $X_n = \sum_{k=0}^{\infty} a_k V_{n-k}$, 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.

(5) The one step ahead linear predictor $\hat{X}_{n+1,n}$ of X_{n+1} based on $\{X_k, k\le n\}$ is given by $\hat{X}_{n+1,n} = \sum_{k=1}^{\infty} a_k V_{n+1-k}$, 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 $1<\alpha<2$ the well known criteria for regularity in the Gaussian case $\alpha=2$. 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 $\|\cdot\|_\alpha$ norm, or equivalently in p-th order mean $0<p<\alpha$.

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

Corollary 1. If $0\le f\in L^1$, then f is factorable as $f=|\phi|^\alpha$ with $\phi\in H^\alpha$, if and only if $\log f\in L^1$.

The time domain criterion (4) provides a "unique" orthogonal moving average representation in terms of a SαS process V. As shown in the proof of Theorem 1, the SαS 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 ϕ 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 specificall, we can prove the following

Proposition 9. (i) A harmonizable SαS process X with $1<\alpha<2$ is regular if and only if it has a moving average representation $X_n=\sum_{k=0}^{\infty}a_kV_{n-k}$ 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 SαS process X with $0<\alpha<2$ is the moving average of an independent SαS process V with $M^X = M^V$.

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 $m\geq 2$. The linear predictor $\hat{X}_{n+m,n}$ of X_{n+m} based on $\{X_k, k\leq n\}$ is the best metric approximation to X_{n+m} in M_n^X:

$$\|X_{n+m} - \hat{X}_{n+m,n}\|_\alpha = \inf\{\|X_{n+m} - Y\|_\alpha : Y \in M_n^X\}$$

or equivalently, by Proposition 1,

$$E|X_{n+m} - \hat{X}_{n+m,n}|^p = \inf\{E|X_{n+m} - Y|^p : Y \in M_n^X\}, \quad 0<p<\alpha,$$

and is uniquely determined by

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

cf. Singer (1970a). In particular, the m-step ahead linear predictor is given by the m-term truncated moving average: $\hat{X}_{n+m,n} = \sum_{k=m}^{\infty} a_k V_{n+m-k}$ if and only if

$$[V_j, \sum_{k=0}^{m-1} a_k V_{n+m-k}]_\alpha = 0 \qquad \text{for all } j\leq n,$$

since $M_n^X = M_n^V$; or equivalently if and only if

$$\int_{-\pi}^{\pi} e^{i\ell\lambda}\Big(\sum_{k=0}^{m-1} a_k e^{ik\lambda}\Big)^{<\alpha-1>} d\lambda = 0 \qquad \text{for all } \ell\geq m,$$

where the a_k's are the Fourier coefficients of the outer factor φ 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 SαS process with $1<\alpha\leq 2$. Then there is a unique 4-variate harmonizable SαS 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 $M_n^X = M_{-\infty}^X + M_n^V$

Of course we also have that $M_{-\infty}^X = M^Z$ is independent of $M^V = M^Y$, and Z_n is the metric projection of X_n onto $M_{-\infty}^X$.

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 SαS 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 SαS processes can not have any of these stronger Wold decompositions.

Proposition 10. A harmonizable SαS 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 ϕ can be defined as in Cambanis and Soltani (1984) (Eq. (5.4) or Remark 5.1): Since $\log f \in L^1$, the function

$$\phi(z) = \exp \{\frac{1}{2\pi\alpha} \int_{-\pi}^{\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z} \log f(\theta)d\theta\}, \quad |z| < 1,$$

is outer, and for a.e. θ, $\lim_{r\uparrow 1} \phi(re^{i\theta}) = \phi(\theta)$ and $|\phi(\theta)|^{\alpha} = f(\theta)$ (cf. Rudin (1966), Theorem 17.16).

Now assume (3). Consider the linear isometry $U_1: L^{\alpha}(f) \to M^X$ defined by $U_1(f) = \int_{-\pi}^{\pi} f dZ$, which is onto (cf. Section 2). Also note that in view of (3), $U_2: L^{\alpha}(f) \to L^{\alpha}(Leb) \stackrel{\Delta}{=} L^{\alpha}$ defined by $U_2(g) = g\phi$ is a linear isometry (which is not necessarily onto). Then $U = U_2U_1^{-1}: M^X \to L^{\alpha}$ is a linear isometry (which is not necessarily onto). Since $U(X_n) = U_2[U_1^{-1}(X_n)] = U_2(e^{-in\theta}) = e^{-in\theta}\phi(\theta)$, we have for all n,

$$\begin{aligned} U(M_n^X) &= L^{\alpha} - \overline{sp}\{e^{-ik\theta}\phi(\theta), \quad k \le n\} \\ &= L^{\alpha} - \overline{sp}\{e^{ij\theta}\phi(\theta), \quad j \ge -n\} \\ &\subseteq L^{\alpha} - \overline{sp}\{e^{ij\theta}, \ j \ge -n\} \quad (\text{as } \phi \in H^{\alpha}) \\ &= L^{\alpha} - \overline{sp}\{e^{-ik\theta}, \ k \le n\}. \end{aligned}$$

Thus $\bigcap_n M_n^X \subseteq U^{-1}[\bigcap_n L_{\alpha}\text{-}\overline{sp}\{e^{-ik\theta}, k \le n\}]$, and in order to show (1): $\bigcap_n M_n^X = \{0\}$, it suffices to show $\bigcap_n L^{\alpha}\text{-}\overline{sp}\,\{e^{-ik\theta}, k\le n\} = \{0\}$. Let $h \in L^{\alpha}\text{-}\overline{sp}\{e^{-ik\theta}, k\le n\}$ for all n. Then

$$\int_{-\pi}^{\pi} h(\theta)e^{-ij\theta}d\theta = 0 \qquad \text{for all } j \ge -n.$$

Since $h \in L^{\alpha} \subset L^{1}$ and all its Fourier coefficients are zero, $h = 0$.

We now show (1)<=>(4). First assume (1). Then choose an outer factor ϕ 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^{α}, and thus

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

Also, in view of the isomorphism U, we have

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

Thus $V = \{V_n\}$ is harmonizable SαS with Lebesgue spectral measure and thus mutually orthogonal r.v.'s:

$$[V_k, V_n]_{\alpha} = \int_{-\pi}^{\pi} e^{-ik\theta} e^{in\theta} d\theta = 0 \quad \text{for all } k \neq n.$$

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

$$V_n = \int_{-\pi}^{\pi} g_n(\theta) dZ(\theta), \quad g_n \in L^{\alpha}(\mu).$$

Now for all n,m we have

$$[X_n, V_m]_\alpha = \begin{Bmatrix} 0, & m>n \\ a_{n-m}, & m\le n \end{Bmatrix} = [X_{n-m}, V_0]_\alpha ,$$

i.e.

$$\int_{-\pi}^{\pi} e^{-in\theta} g_m^{<\alpha-1>} d\mu = \int_{-\pi}^{\pi} e^{-i(n-m)\theta} g_0^{<\alpha-1>} d\mu .$$

Thus for each m the Fourier transforms of the finite measures $g_m^{<\alpha-1>}d\mu$ and $e^{im\theta}g_0^{<\alpha-1>}d\mu$ coincide, hence these measures are identical, i.e. $g_m^{<\alpha-1>}(\theta) = e^{im\theta}g_0^{<\alpha-1>}(\theta) = (e^{-im\theta}g(\theta))^{<\alpha-1>}$ a.e. $[\mu]$, and thus

$$g_m(\theta) = e^{-im\theta} g_0(\theta) \quad \text{a.e. } [\mu]$$

(since $z = w$ iff $z^{<\beta>} = w^{<\beta>}$). From the orthogonality of the V_n's we have

$$0 = [V_n, V_0]_\alpha = \int_{-\pi}^{\pi} g_n g_0^{<\alpha-1>} d\mu = \int_{-\pi}^{\pi} e^{-in\theta} |g_0|^\alpha d\mu, \quad n \neq 0.$$

It follows from the Riesz theorem that the measure $|g_0(\theta)|^\alpha d\mu(\theta)$ is absolutely continuous with respect to Lebesgue measure: $|g_0(\theta)|^\alpha d\mu(\theta) = c(\theta)d\theta$, 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 $c(\theta)$ = positive constant = c^α, say, a.e. [Leb]. Thus

$$|g_0(\theta)|^\alpha d\mu(\theta) = c^\alpha d\theta .$$

Note that M_n^X equals M_n^V, which is isomorphic under the stochastic integral to

$$L^\alpha(\mu) - \overline{sp}\{g_k,\ k\le n\} = (L^\alpha - \overline{sp}\{e^{-ik\theta},\ k\le n\})\cdot g_0,$$

which is in turn isomorphic to $L^\alpha - \overline{sp}\ \{e^{-ik\theta},\ k\le n\}$ under the correspondence $h\cdot g_0 \leftrightarrow ch$. Thus in order to show that X is regular it is equivalent to show that $\bigcap_n L^\alpha - \overline{sp}\ \{e^{-ik\theta},\ k\le n\} = \{0\}$, which has been done in the third paragraph of this proof. Thus (1) is shown.

We finally show that (4) $\Leftrightarrow$ (5). Assume (4). Put $Y = \sum_{k=1}^{\infty} a_k V_{n+1-k} \in M_n^V = M_n^X$. Then $X_{n+1} - Y = a_0 V_{n+1}$ and the orthogonality of the V_k's imply $[V_k, X_{n+1}-Y]_\alpha = 0$ for all $k \le n$, and by linearity and continuity $[W, X_{n+1} - Y]_\alpha = 0$ for all $W \in M_n^V = M_n^X$, including in particular all X_k's with $k \le n$. It follows that $Y = X_{n+1,n}$.

Conversely assume (5). Put $W_{n+1} = X_{n+1} - \hat{X}_{n+1,n}$. Then $[W, W_{n+1}]_\alpha = 0$ for all $W \in M_n^X = M_n^V$. Also from $X_{n+1} = W_{n+1} + \sum_{k=1}^{\infty} a_k V_{n+1-k}$ we obtain $M_{n+1}^X = M_n^V + sp(W_{n+1})$, and from $M_{n+1}^X = M_{n+1}^V$ by a standard argument $M_{n+1}^X = M_n^V + sp(V_{n+1})$. Hence $W_{n+1} \in M_{n+1}^X$ can be written in the form $W_{n+1} = Y + cV_{n+1}$ where $Y \in M_n^V$, and thus $W_{n+1} - cV_{n+1} = Y \in M_n^V$. It then follows that

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

i.e.

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

Then $c \neq 0$. For if $c = 0 \Rightarrow W_{n+1} = 0 \Rightarrow X_{n+1} = \hat{X}_{n+1,n} \in M_n^X \Rightarrow M_{n+1}^X = M_n^X \Rightarrow$

$\Rightarrow M_{n+1}^V = M_n^V \Rightarrow [V_{n+1}, V_{n+1}]_\alpha = 0 \Rightarrow V_n = 0 \Rightarrow X_n = 0$ i.e. X is the zero

process. Then multiplying the above equalities together we obtain $[V_{n+1}, W_{n+1}]_\alpha [W_{n+1}, V_{n+1}]_\alpha = ||V_{n+1}||_\alpha^\alpha \, ||W_{n+1}||_\alpha^\alpha$. Writing $V_{n+1} = \int f_{n+1} dZ$ and $W_{n+1} = \int g_{n+1} dZ$, we thus have

$$\int f_{n+1} g_{n+1}^{<\alpha-1>} d\mu \cdot \int g_{n+1} f_{n+1}^{<\alpha-1>} d\mu = \int |f_{n+1}|^\alpha d\mu \cdot \int |g_{n+1}|^\alpha d\mu .$$

Dropping for simplicity the subscripts, this means that equality holds in Hölder's inequalities $\int |f|\,|g|^{\alpha-1} d\mu \le ||f||_\alpha \, ||g||_\alpha^{\alpha-1}$,

$\int |g|\,|f|^{\alpha-1} d\mu \le ||g||_\alpha ||f||_\alpha^{\alpha-1}$, so that $|g(\theta)| = r|f(\theta)|$ a.e. $[\mu]$, for some $r>0$, and thus $g(\theta) = re^{i\phi(\theta)}f(\theta)$ a.e. $[\mu]$. Substituting in the above equation we obtain

$$\left|\int e^{i\phi}|f|^{\alpha}d\mu\right|^2 = \left(\int |f|^{\alpha}d\mu\right)^2$$

and an elementary argument shows that $e^{i\phi(\theta)}$ is a complex constant for a.e. θ w.r.t. $|f|^{\alpha}d\mu$. Thus $g_{n+1}(\theta) = z_{n+1}f_{n+1}(\theta)$ a.e. $[\mu]$ 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 $[X_{n+1} - \hat{X}_{n+1,n}, V_{n+1}]_{\alpha} = [W_{n+1}, V_{n+1}]_{\alpha} = z_{n+1}$ is independent of n. Hence putting $z_{n+1} = a_0$ we obtain $X_{n+1} = \sum_{k=0}^{\infty} a_k V_{n+1-k}$. 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): $X_n = \sum_{k=0}^{\infty} a_k V_{n-h} = \sum_{k=0}^{\infty} b_k U_{n-k}$.

Since the metric projection $\hat{X}_{n,n-1} = \sum_{k=1}^{\infty} a_n V_{n-k} = \sum_{k=1}^{\infty} b_k U_{n-k}$ of X_n onto $M_n^X = M_n^V = M_n^U$ is unique we obtain $a_0 V_n = b_0 U_n$. 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 $V_n \in M^X$, they are of the form $V_n = \int_{-\pi}^{\pi} f_n(\theta)dZ(\theta)$, $f_n \in L^{\alpha}(\mu)$, 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 $X_n = \int_{-\pi}^{\pi} e^{-in\theta}dZ(\theta)$ that for all n,

$$e^{-in\theta} = \sum_{k=0}^{\infty} a_k f_{n-k}(\theta) \quad \text{in } L^{\alpha}(\mu).$$

Thus for all $k \geq 0$ and all n, on E_{n-k}: $e^{-in\cdot} = a_k f_{n-k}(\cdot)$ a.e. [μ], or, equivalently,

$$\text{on each } E_m: \; e^{-i(m+j)\cdot} = a_j f_m(\cdot) \quad \text{a.e. } [\mu],\; j \geq 0.$$

If all $f_m = 0$ then all $V_m = 0$ and $X = 0$. Thus for some m, $\int_{E_m} |f_m|^{\alpha} d\mu > 0$. It follows from the displayed equality that then $a_j \neq 0$, $j \geq 0$, which in turn implies $\int_{E_m} |f_m|^{\alpha} d\mu > 0$ for all m. Now fix an arbitrary m, and some $\theta \in E_m$ with an E_m neighbourhood of positive μ measure. Then $f_m(\theta) = e^{-i(m+j)\theta}/a_j$ for all $j \geq 0$, implies $a_j = e^{-ij\theta} a_0$ for all $j \geq 0$. But since this should hold for each such θ on each of the disjoint sets 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 $M_{-\infty}^X$ (Cambanis et al. (1985)) hence $Z_n' = Z_n$ and thus also $Y_n' = Y_n$. It follows that $M_n^V = M_n^Y = M_n^{Y'} = M_n^{V'}$ and thus V has a left WD, since V' does. Similarly assuming X has a righ WD it follows that so does V. But it has been proven in Example 4 of Cambanis et al.(1985), that a harmonizable SαS 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 $S\alpha S$ 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 $\{X_n, n \le 0\}$ and of $\{X_n, n \ge 1\}$ 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 $S\alpha S$ process $X = \{X_n\}$ with $1<\alpha\le 2$ and spectral measure μ 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) $d\mu(\theta) = f(\theta)d\theta$, $L^\alpha(f) \subset L^1$, and the Fourier series of every $g \in L^\alpha(f)$ converges to g in $L^\alpha(f)$.

(5) $d\mu(\theta) = f(\theta)d\theta$ and the spectral density f satisfies

$$(A_\alpha) \qquad \left(\frac{1}{|I|}\int_I f(\theta)d\theta\right)\left(\frac{1}{|I|}\int_I f(\theta)^{\frac{-1}{\alpha-1}}\,d\theta\right)^{\alpha-1} \le k$$

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 $L^1(\mu)$.

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 μ 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 μ 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 $\sum_n a_n e^{in\theta}$ is defined by $\sum_{n\neq 0} -i \operatorname{sgn}(n) a_n e^{in\theta}$.

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

$$\left\| \frac{X_n}{\|X_n\|} - \frac{Y_n}{\|Y_n\|} \right\| = \left\| \frac{X_n}{\|X_n\|} - \frac{Y_n}{\|X_n\|} + \frac{Y_n}{\|X_n\|} - \frac{Y_n}{\|Y_n\|} \right\|$$

$$\leq \frac{\|X_n - Y_n\|}{\|X_n\|} + \frac{\big| \|Y_n\| - \|X_n\| \big|}{\|X_n\|} \leq \frac{2}{\|X_n\|} \|X_n - Y_n\| \leq \frac{2}{n} \to 0$$

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

$$\Big\| \sum_{n=-k'}^{m'} c_n X_n \Big\|_\alpha \leq k \Big\| \sum_{n=-k}^{m} c_n X_n \Big\|_\alpha$$

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

(iii) There is a constant k such that

$$\Big\| \sum_{n=k'}^{m'} c_n X_n \Big\|_\alpha \leq k \Big\| \sum_{n=k}^{m} c_n X_n \Big\|_\alpha$$

for all $k \leq k' \leq m' \leq m$.

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 Miamee and Niemi (1985). Assume (3). To show that X is minimal: $X_n \notin \overline{sp}\{X_k, k \neq n\}$, it suffices to show that $\overline{sp}\{X_k, k \neq n\} \subsetneq M^X$. Assume on the contrary that $\overline{sp}\{X_k, k\neq n\} = M^X$ for some n, and hence by stationarity for all n, so that $\cap_n \overline{sp}\{X_k, k\neq n\} = M^X$. In fact we will show that $\cap_n \overline{sp}\{X_k, k\neq n\} = \{0\}$, namely that X is J_0-regular. Indeed if $Y \in \cap_n \overline{sp}\{X_k, k\neq n\}$ then by (3) it can be written uniquely as $Y = \sum_n c_n Y_n$, and since for each n, $Y \in \overline{sp}\{X_k, k\neq n\}$ we have $c_n = 0$ and thus Y=0. Hence X is J_0-regular, and thus minimal as well as regular, since $\cap_n \overline{sp}\{X_k, k\leq n\} \subset \cap_n \overline{sp}\{X_k, k\neq n\} = \{0\}$.

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

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

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 $Y = \int g dZ \longleftrightarrow g$ between M^X and $L^\alpha(f)$, it follows that every $g \in L^\alpha(\mu)$ has a unique representation $g(\cdot) = \sum_n c_n e^{-in\cdot}$ 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 $Y = \int g dZ$ for some $g \in L^\alpha(f)$, and by (4), $g(\cdot) = \sum_n \hat{g}_n e^{-i\cdot}$ in $L^\alpha(f)$. It follows that $Y = \sum_n \hat{g}_n X_n$ 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 Szëgo (1960), pp. 129-130, shows that (6) is equivalent to the boundedness of the truncation operator T from $L^\alpha(\mu)$ into itself defined by

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

which, considering the isomorphism $e^{-in\theta} \longleftrightarrow X_n$, 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

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

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

$$(P_m) \qquad \hat{X}_{n+m,n} = \sum_{k=0}^{\infty} c_{m,k} X_{n-k},$$

i.e. it can be realized by the filter $\phi_m = \{\phi_{m,k}\}_{k\geq 0}$ acting on the observed part of X. These series converge with respect to the norm $||\cdot||_\alpha$, or equivalently in $L^p(\Omega)$. While condition (A_α) guarantees that all functions in $L^\alpha(f)$ have convergent Fourier series, and all r.v.'s 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

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

(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 $\hat{X}^r_{n+m,n}$ exists then it is given by

$$\hat{X}^r_{n+m,n} = \sum_{k=m}^{\infty} a_k V_{n+m-k}.$$

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

$$\hat{X}^r_{n+m,n} = E(X_{n+m}|M^X_n) = \sum_{k=m}^{\infty} a_k V_{m+n-k}.$$

Proof. (i) If the m-step ahead right angle predictor $\hat{X}^r_{n+m,n}$ 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. (This is because the density of V is simply the Lebesgue measure, which clearly satisfies the (A_α) condition of Theorem 3). Thus one can write $\hat{X}^r_{n+m,n}$ as a convergent series

$$\hat{X}^r_{n+m,n} = \sum_{k=m}^{\infty} c_k V_{n+m-k}.$$

Now by the definition of right angle projection we have

$$[\hat{X}^r_{n+m,n}, Y] = [X_{n+m}, Y]$$

for every Y in $M^X_n = M^V_n$ and in particular for $Y = V_k$, with $k \le n$; i.e. we have

$$[\hat{X}^r_{n+m,n}, V_k] = [X_{n+m}, V_k], \quad k \le n.$$

This shows that

$$a_i = c_i, \quad \text{for all } i \ge m.$$

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

(P_1) $$\hat{X}_{n+1,1} = \sum_{k=0}^{\infty} d_k X_{n-k},$$

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

(AP_m) $$\hat{X}^r_{n+m,n} = \sum_{k=0}^{\infty} e_{k,m} X_{n-k}.$$

The equivalence of (I) and (P_1) can be established in the time domain. Indeed assume that (I) holds. The orthogonality of V_n's implies that for every Y in $M_n^X = M_n^V$ 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 (P_1) is satisfied.

Conversely, assume (P_1) is satisfied. From (3) and (4) of Theorem 1, we have

$$X_{n+1} - \hat{X}_{n+1,n} = a_0 V_{n+1},$$

and thus by (P_1),

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

with $b_0 = a_0^{-1}$ and $b_k = -d_{k-1}/d_0$, $k \geq 1$. 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 (1984), together with the representation of $\hat{X}^r_{n+m,n}$ 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 ϕ^{-1} has a series expansion

$$\text{(F)} \qquad \phi^{-1}(\theta) = \sum_{k=0}^{\infty} a_k e^{-ik\theta},$$

converging in $L^{\alpha}(f)$.

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 $\alpha=2$). Following are some sufficient conditions which are easier to check. The simplest is the one suggested by Masani (1960):

$$\text{(M)} \qquad f \in L^{\infty} \quad \text{and} \quad f^{-1} \in L^1.$$

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

(B) $f = hg$ where h satisfies (A_α) and $g > 0$ satisfies (M).

The following are yet weaker conditions.

Proposition 13. Let X be regular harmonizable SαS with $1 < \alpha < 2$, and let ϕ 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 $h_i, g_i > 0$, h_i satisfies condition (A_α), and g_i satisfies condition (M), i = 1,2.

(b) $g_1h_1 \le f \le g_2h_2$ where $h_i, g_i > 0$, $g_1^{-1} \in L^1$, $g_2 \in L^\infty$, $L^\alpha(h_1) = L^\alpha(h_2)$ and h_1 satisfies (A_α)

Proof. (a) Clearly $f \ge h_ig_i$, for i = 1,2, so $f^{-1} \le (h_ig_i)^{-1}$, and hence

$$|\phi^{-1}|^\alpha h_i = \frac{h_i}{f} \le g_i^{-1}.$$

Thus $\phi^{-1} \in L^\alpha(h_i)$. Now since h_i satisfies (A_α) by Theorem 3 we see that $(\phi^{-1})^N$, the N-th Fourier partial sum of ϕ^{-1}, converges to ϕ^{-1} in $L^\alpha(h_i)$ and hence in $L^\alpha(h_ig_i)$ (because $g_i \in L^\infty$), i.e.

$$\int |(\phi^{-1})^N - \phi^{-1}|^\alpha f_ig_i \to 0, \quad i = 1,2.$$

Adding these two together we get

$$\int |(\phi^{-1})^N - \phi^{-1}|^\alpha f \to 0,$$

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_m). We know that $|1 + e^{i\theta}|^p$ satisfies (A_2) for $-1 < p < 1$, by Helson and Szegö (1960), and (M) for $0 \leq p < 1$. 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

Akutowicz, E.J. (1957). On an explicit formula in linear least squares prediction. Math. Scand. 5, 261-266.

Bloomfield, P. (1984). On series representations for linear predictors. Ann. Probab. 3, 226-233.

Cambanis, S. (1983). Complex symmetric stable variables and processes. Contributions to statistics: Essays in Honor of Norman L. Johnson, P.K. Sen (Ed.), North Holland, New York, pp. 63-79.

Cambanis, S., Hardin, C.D. Jr., and Weron, A. (1984) Ergodic properties of stationary stable processes. Center for Stochastic Processes, Tech. Rept. No. 59, Statistics Dept., Univ. of North Carolina, Chapel Hill.

Cambanis, S., Hardin, D.C. Jr. and Weron, A. (1985). Innovations and Wold decompositions of stable sequences. Center for Stochastic Processes, Tech. Rept. No. 106, Univ. of North Carolina, Chapel Hill.

Cambanis, S. and Miller, G. (1981). Linear problems in pth order and stable processes. SIAM J. Appl. Math. 41, 43-69.

Cambanis, S. and Soltani, A.R. (1984). Prediction of stable processes: Spectral and moving average representations, Z. Wahrsch. verw. Geb. 66, 593-612.

Cline, H. and Brockwell, J. (1985). Linear prediction of ARMA processes with infinite variance. Stochastic Proc. Appl. 19, 281-296.

Hannan, E.J. and Kanter, M. (1977). Autoregressive processes with infinite variance. J. Appl. Probability, 14, 411-415.

Hardin, D.C. (1982). On the spectral representation of symmetric stable processes. J. Multivariate Anal. 12, 385-401.

Hardin, C.D. (1982a). On the linearity of regression. Z. Wahrsch. verw. Geb. 61, 293-302.

Helson, M. and Szegö, G. (1960). A problem in the prediction theory. Ann. Math. Pura Appl. 51, 107-138.

Hosoya, Y. (1978). Discrete time stable processes and their certain properties, Ann. Probab. 6, 94-105.

Hosoya, Y. (1982). Harmonizable stable processes, Z. Wahrsch. verw. Geb. 60, 517-533.

Hunt, R., Muckenhoupt, B. and Wheeden, R.L. (1973). Weighted norm inequalities for the conjugate function and Hilbert transform. Trans. Amer. Math. Soc. 176, 227-251.

Köthe, G. (1969). Topological Vector Spaces I. Springer-Verlag, New York.

LePage, R. (1980). Multidimensional infinitely divisible variables and processes. Part I: Stable case. Tech. Rept. No. 292. Statistics Dept., Stanford University.

Luenberger, D.G. (1969). Optimization by Vector Space Methods. Wiley, New York.

Masani, P.R. (1960). The prediction theory of multivariate stochastic processes III. Acta. Math. 104, 141-162.

Masani, P.R. (1966). Recent trands in multivariate prediction theory. Multivariate Analysis I, P.R. Krishnaiah (Ed.), Academic Press, New York pp. 351-382.

Masry, E. and Cambanis, S. (1984). Spectral density estimation for stationary stable processes. Stochastic Proc. Appl. 18, 1-31.

Miamee, A.G. and Niemi, H. (1985). On the angle for stationary random fields. Center for Stochastic Processes. Tech. Rept. No. 92, Statistics Dept., Univ. of North Carolina, Chapel Hill.

Miamee, A.G. and Salehi, H. (1983). On an explicit representation of the linear predictor of a weakly stationary stochastic sequences. Bol. Soc. Math. Mexicana 28, 81-93.

Pourahmadi, M. (1984). On minimality and interpolation of harmonizable stable processes. SIAM J. Appl. Math. 44, 1023-1030.

Pourahmadi, M. (1984a). Infinite order autoregressive representations of multivariate stationary stochastic processes. Center for stochastic processes. Tech. Rept. No. 75, Statistics Dept. Univ. of North Carolina, Chapel Hill.

Pourahmadi, M. (1985). A matricial extension of the Helson-Szegö theorem and its application in multivariate prediction. J. Multivariate Anal. (to appear)

Rozanov, Yu. A. (1967). Stationary Random Processes. Holden-Dav. San Francisco.

Rudin, W. (1966). Real and Complex Analysis. McGraw-Hill, New York.

Singer, I. (1970). Bases in Banach Spaces I. Springer-Verlag, New York.

Singer, I. (1970a). Best Approximation in Normed Linear Space by Elements of Linear Subspaces. Springer-Verlag, New York.

Urbanik, K. (1970). Harmonizable sequences of random variables. Colloques Internationaux du C.N.R.S. 186, 345-361.

Weron, A. (1985). Harmonizable stable processes on groups: Spectral, ergodic and interpolation properties. Z. Wahrsch verw. Geb. 68, 473-491.

END

FILMED

1-86

DTIC