

# A REPRESENTATION THEORY FOR A CLASS OF VECTOR AUTOREGRESSIVE MODELS FOR FRACTIONAL PROCESSES

SØREN JOHANSEN \*

*University of Copenhagen*

Department of Economics  
University of Copenhagen  
Studiestræde 6  
DK-1455 Copenhagen K  
Denmark  
email: sjo@math.ku.dk

---

\*I would like to thank the Danish Social Sciences Research Council for continuing support, and the members of the ESF-EMM network for discussions. As usual the insight and help of Søren Tolver Jensen is greatly appreciated in connection with Lemma 10. I am grateful for comments from an anonymous referee which helped clarify some explanations, and to Bent Nielsen who helped with the proof of Lemma 11. Finally Massimo Franchi had some good comments to the almost final version.

Running headline:

## FRACTIONAL AUTOREGRESSIVE PROCESSES

Abstract:

Based on an idea of Granger (1986), we analyze a new vector autoregressive model defined from the fractional lag operator  $1 - (1 - L)^d$ . We first derive conditions in terms of the coefficients for the model to generate processes which are fractional of order zero. We then show that if there is a unit root, the model generates a fractional process  $X_t$  of order  $d$ ,  $d > 0$ , for which there are vectors  $\beta$  so that  $\beta' X_t$  is fractional of order  $d - b$ ,  $0 < b \leq d$ . We find a representation of the solution which demonstrates the fractional properties. Finally we suggest a model that allows for a polynomial fractional vector, that is, the process  $X_t$  is fractional of order  $d$ ,  $\beta' X_t$  is fractional of order  $d - b$  and a linear combination of  $\beta' X_t$  and  $\Delta^b X_t$  is fractional of order  $d - 2b$ . The representations and conditions are analogous to the well known conditions for  $I(0)$ ,  $I(1)$  and  $I(2)$  variables.

Mailing address:

Søren Johansen

Department of Economics

University of Copenhagen

Stuejstræde 6

DK-1455 Copenhagen K

Denmark

email: sjo@math.ku.dk

# 1 INTRODUCTION AND MOTIVATION

Since the definition of fractional processes by Granger and Joyeux (1980) and Hosking (1981), attention has been given to the estimation of the fractional order of a time series, see for instance Beran (1994), with minimal assumptions on the processes. The problem considered in this paper is different.

We propose a vector autoregressive model that allows for a fractional process as solution, but also allows for modelling cofractional relations and adjustments to these, so that economic hypotheses can be formulated within a model, that can be tested against the data. In this way we have a platform for making model based inference on coefficients, relations and fractional order.

The contribution of the present paper is to study the solution of this autoregressive model in order to find conditions under which the process is fractional of order zero and conditions for the solution to be fractional of order  $d > 0$ , but allowing for linear combinations  $\beta'X_t$ , that are fractional of order  $d - b$ ,  $0 < b \leq d$ . We extend the results to a model that further allows polynomial fractionality, that is, a linear combination of  $\Delta^b X_t$  and  $\beta'X_t$ , which is fractional of order  $d - 2b$ . Such models are well known for  $I(0)$ ,  $I(1)$ , and  $I(2)$  variables, see for example Johansen (1996), and we formulate the models so that results for the usual cointegrating models can be carried over to the new framework.

## 1.1 Granger's model for fractional processes

Granger (1986) suggested an autoregressive model

$$A^*(L)\Delta^d X_t = (1 - \Delta^b)\Delta^{d-b}\alpha\beta'X_{t-1} + d(L)\varepsilon_t, \quad (1)$$

where  $\varepsilon_t$  is independent identically distributed with mean zero and positive definite variance  $\Omega > 0$ , denoted i.i.d.  $(0, \Omega)$ . We have used the matrix  $\alpha\beta'$  instead of  $-\gamma\alpha'$  from Granger (1986).

He noted that the 'lag operator'

$$L_b = 1 - (1 - L)^b, \quad (2)$$

plays the role of the usual lag operator ( $b = 1$ ,  $L_1 = L$ ) in models for fractional processes. This model has been analyzed by Lyhagen (1998) from the point of view of finding the properties of the solution, but unfortunately the results and their proof are not correct.<sup>†</sup>

Dittmann (2004) attempts to derive model (1) from a moving average form, assuming the processes are fractional, using the results of Engle and Granger (1987). However, the results are not correctly proved.<sup>‡</sup>

---

<sup>†</sup>The proof is attempted by the usual transformation from  $X_t$  to  $\beta'X_t$  and  $\beta'_\perp \Delta X_t$ , but for fractional processes a different method is needed. This becomes apparent when it is stated that the matrix  $D(z) = \text{diag}((1 - z)^{d_i} - (1 - z))$  is of full rank for  $z = 1$ .

<sup>‡</sup>A counter example to Lemma 1 in the paper Engle and Granger (1987) is given by taking  $G(\lambda) = \text{diag}(1 + \lambda, \lambda^2, \lambda^2)$ . What is missing in Lemma 1 is a condition corresponding to the  $I(1)$  condition of cointegration (see Johansen 1996, Theorem 4.5.)

This type of model can for instance be derived as follows. Let us assume that  $X_t$  is fractional of order  $d$ , and that there are  $r$  linear combinations,  $\beta$ , that are fractional of order  $d - b$ . Let  $\xi$  be  $p \times (p - r)$  so that  $(\xi, \beta)$  has rank  $p$ , then the assumptions are

$$\begin{aligned}\xi' \Delta^d X_t &= u_{1t}, \\ \beta' \Delta^{d-b} X_t &= u_{2t},\end{aligned}$$

where for simplicity  $u_t = (u'_{1t}, u'_{2t})'$  is i.i.d.  $(0, \Sigma)$ . This formulation was used by Breitung and Hassler (2002) and allows for modelling and estimating both the cofraction vectors,  $\beta$ , and the ‘common trends’ vectors,  $\xi$ . For a  $p \times m$  matrix  $a$  we define  $a_\perp$  to be a  $p \times (p - m)$  matrix of rank  $p - m$ , for which  $a'a_\perp = 0$ . When  $(\xi, \beta)$  has full rank  $p$ , the identity

$$\xi_\perp (\beta' \xi_\perp)^{-1} \beta' + \beta_\perp (\xi' \beta_\perp)^{-1} \xi' = I_p,$$

shows that

$$\begin{aligned}\Delta^d X_t &= \beta_\perp (\xi' \beta_\perp)^{-1} u_{1t} + \xi_\perp (\beta' \xi_\perp)^{-1} \Delta^b u_{2t} \\ &= \beta_\perp (\xi' \beta_\perp)^{-1} u_{1t} + \xi_\perp (\beta' \xi_\perp)^{-1} u_{2t} - \xi_\perp (\beta' \xi_\perp)^{-1} (1 - \Delta^b) u_{2t} \\ &= (1 - \Delta^b) \alpha \beta' \Delta^{d-b} X_t + \varepsilon_t,\end{aligned}\tag{3}$$

where  $\varepsilon_t = \beta_\perp (\xi' \beta_\perp)^{-1} u_{1t} + \xi_\perp (\beta' \xi_\perp)^{-1} u_{2t}$ , is i.i.d. The adjustment matrix  $\alpha = -\xi_\perp (\beta' \xi_\perp)^{-1}$  satisfies  $\beta' \alpha = -I_r$ , and (3) is a special case of model (1) with  $A^*(z) = 1$ , apart from the lagged  $X_t$ . It is now a natural idea to make the model more flexible by adding a lag structure, and Granger suggested to add lags of  $\Delta^d X_t$ .

**EXAMPLE 1.** As a simple example of model (1) consider the univariate model with one lag and  $b = d$ ,

$$\Delta^d X_t = (1 - \Delta^d) \gamma_1 X_{t-1} + \gamma_2 \Delta^d X_{t-1} + \varepsilon_t.\tag{4}$$

The characteristic function, which is not a polynomial unless  $d$  is an integer, is

$$\pi(z) = (1 - z)^d - \gamma_1 (1 - (1 - z)^d) z - \gamma_2 (1 - z)^d z.\tag{5}$$

The process is stationary if the roots of  $\pi(z) = 0$  are greater than one in absolute value. This criterion involves solving an unpleasant transcendental equation. Finding the roots of this equation is not a standard problem. ■

Thus model (1) is an autoregressive model for fractional processes, but its lag structure is inconvenient to analyze in the sense that the stochastic properties of the solution generated by the equations are not easily reflected in properties of the coefficients.

Therefore we propose a slightly different model with different choice of lag structure, for which we get a feasible algebraic analysis, in the sense that we get a verifiable criterion for the solution of the model equations to be fractional of various orders.

## 1.2 An alternative autoregressive model for fractional processes

We propose an autoregressive model,  $VAR_{d,b}(k)$ ,  $k = 0, 1, \dots$ ,  $0 < b \leq d$ , of the form

$$A(L_b)\Delta^d X_t = (1 - \Delta^b)\Delta^{d-b}\alpha\beta'X_t + \varepsilon_t,$$

or, if  $A(z) = I_p - \sum_{i=1}^k \Gamma_i z^i$ , the model is

$$\Delta^d X_t = \Delta^{d-b}\alpha\beta' L_b X_t + \sum_{i=1}^k \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t. \quad (6)$$

We assume throughout that  $\varepsilon_t$  is i.i.d.  $(0, \Omega)$  in  $p$  dimensions. This model preserves the main structure of (1) in that it allows for modeling of cofractionality and adjustment. We have dropped the lag on  $X_t$ , which seems superfluous because  $L_b$  is already a lag operator. The essential difference is, that we have replaced the usual lag operator in the polynomial  $A^*(L)$  by the new lag operator, see (2).

Note that the model (6) is not a fractional ARIMA model, that is, it is not of the form

$$D(L)\Delta^d X_t = B(L)\varepsilon_t,$$

where  $D(L)$  and  $B(L)$  are finite order lag polynomials. It is, however, a fractional ARIMA model in the new lag operator, because it can be expressed as  $\Pi^*(L_b)\Delta^{d-b}X_t = \varepsilon_t$ , for  $\Pi^*(u) = (1 - u)I_p - \alpha\beta'u - \sum_{i=1}^k \Gamma_i(1 - u)u^i$ .

EXAMPLE 1. (continued) The simple univariate example from (4) is changed into

$$\Delta^d X_t = \gamma_1 L_d X_t + \gamma_2 \Delta^d L_d X_t + \varepsilon_t.$$

The characteristic function is

$$\begin{aligned} \pi(z) &= (1 - z)^d - \gamma_1(1 - (1 - z)^d) - \gamma_2(1 - z)^d(1 - (1 - z)^d) \\ &= 1 - u - \gamma_1 u - \gamma_2(1 - u)u = \pi^*(u), \quad u = 1 - (1 - z)^d. \end{aligned}$$

Note that  $\pi(\cdot)$  is not a polynomial but  $\pi^*(\cdot)$  is a second degree polynomial in the variable  $u = 1 - (1 - z)^d$ . Hence  $\pi(z) = 0$  is equivalent to  $\pi^*(u) = 0$ , which is the usual polynomial equation, which is well understood. We just have to understand the mapping  $z \mapsto 1 - (1 - z)^d$  and its range in the complex plane  $\mathbb{C}$ ; see Appendix A.3. ■

Thus the idea, to get a tractable theory, is to replace  $A^*(L)$  in (1) by a polynomial in the lag operator  $L_b$ , or equivalently, replace in the usual cointegrated VAR model the difference operator  $\Delta$  by  $\Delta^b$ . This trick allows us to carry over the result from the  $I(0)/I(1)/I(2)$  theory and develop the analogous theory for fractional processes.

This paper does not deal with inference, but the models (1) and (6) have the advantage that for known fractional orders, it is easy maximize the Gaussian likelihood applying the usual reduced rank techniques, which are well known from the  $I(1)$  model. This leaves a function of only one or two variables  $(b, d)$  to be optimized.

For model (1) this was noted by Lyhagen (1998) and Lasak (2005). Lyhagen derived the asymptotic distribution of the test for no cointegration, when the fractional order is known, and Lasak derived the limit distribution of this test in case the fractional order  $b$  is unknown and  $d = 1$ . Model (1) has been applied by Davidson (2002) and Breitung and Hassler (2002).

### 1.3 An overview of the paper

In the next section we give the basic properties of the fractional operator and define the class of fractional processes  $\mathcal{F}(d)$  and the notion of a cofractional vector. In section 3 we analyze the solution of  $VAR_{d,b}(k)$  and start by finding a criterion for the solution of the univariate  $AR_d(1)$  model to be fractional of order zero in terms of the roots of a characteristic polynomial. Next we give the theory for the cofractional model  $VAR_{d,b}(k)$  and show that the model allows for fractional processes of order  $d$ , for which  $\beta'X_t$  is fractional of order  $d - b$ . Finally in section 4 we define and analyze a model for polynomially cofractional processes, with the property that  $X_t$  is fractional of order  $d$ ,  $\beta'X_t$  fractional of order  $d - b$  and finally that a linear combination of  $\beta'X_t$  and  $\Delta^b X_t$  is fractional of order  $d - 2b$ .

In the Appendix we have collected some mathematical tools: The notion of a regular (or holomorphic) function, and a few results on some special complex functions. We suggest a convenient formalism for expressing the solution of difference equations, which involve fractional differences, as a function of (infinitely many) initial values and the errors of the equations. We end the Appendix with a result from Grenander and Rosenblatt (1956) which is the key tool in the paper and a trigonometric inequality.

## 2 FRACTIONAL PROCESSES

We let  $C_n$  be a sequence of  $p \times p$  matrices for which  $\sum_{n=0}^{\infty} \|C_n\|^2 < \infty$ , where  $\|C_n\|^2 = \text{tr}(C_n' C_n)$ . We define  $C(z) = \sum_{n=0}^{\infty} z^n C_n$ ,  $|z| < 1$  and let  $\varepsilon_t$  be a sequence of  $p$ -dimensional i.i.d. variables with  $E(\varepsilon_t) = 0$  and  $\text{Var}(\varepsilon_t) = \Omega > 0$ . This allows us to define the stationary linear process  $X_t = \sum_{n=0}^{\infty} C_n \varepsilon_{t-n}$  with mean zero and finite variance  $\sum_{k=0}^{\infty} C_k \Omega C_k'$ . In order to have an expression for the spectral density, we assume that  $C(z)$  can be extended to the boundary  $|z| = 1$  by continuity so that the spectral density is

$$f_X(\lambda) = \frac{1}{2\pi} C(e^{-i\lambda}) \Omega C(e^{i\lambda})'.$$

If  $\sum_{n=0}^{\infty} \|C_n\| < \infty$  we can define  $I(0)$ , or  $\mathcal{F}(0)$ , by the condition that  $C(1) = \sum_{n=0}^{\infty} C_n \neq 0$ , see Johansen (1996), but for the processes studied here, the sum in the sense of  $\lim_{N \rightarrow \infty} \sum_{n=0}^N C_n$  need not be defined. Instead, when  $C(z)$  is continuous for  $|z| \leq 1$ , we have  $C(1) = \lim_{r \rightarrow 1} \sum_{n=0}^{\infty} r^n C_n$ , which we use in the definition of  $\mathcal{F}(0)$ .

We first define an  $\mathcal{F}(0)$  process, that is, a process which is fractional of order zero.

DEFINITION 1 *If  $\sum_{n=0}^{\infty} \|C_n\|^2 < \infty$  and  $C(z) = \sum_{n=0}^{\infty} C_n z^n, |z| < 1$  can be extended to a continuous function on the boundary  $|z| = 1$ , we call the stationary linear process  $X_t = \sum_{k=0}^{\infty} C_n \varepsilon_{t-n}$  fractional of order zero,  $F(0)$ , if the spectrum at zero  $f_X(0) = \frac{1}{2\pi} C(1)\Omega C(1)' \neq 0$ .*

*For such processes, we denote  $F(0)_+$  the class of processes of the form*

$$X_t^+ = C(L)_+ \varepsilon_t = C(L) \varepsilon_t 1\{t \geq 1\} = \begin{cases} \sum_{n=0}^{t-1} C_n \varepsilon_{t-n} & t = 1, 2, \dots \\ 0 & t = 0, -1, \dots \end{cases} \quad (7)$$

Note that we do not assume in Definition 1, that  $f_X(0) > 0$ , so that any linear combination is also fractional of order zero, and we do not assume that  $\text{diag}(f_X(0)) > 0$ , which would imply that all components of  $X_t$  were fractional of order zero. The class  $\mathcal{F}(0)$  will be used to define fractional processes of other orders and examples will be discussed below. Note that  $X_t^+ = \sum_{n=0}^{t-1} C_n \varepsilon_{t-n}$ ,  $t = 1, 2, \dots$  is not stationary, because nothing is assumed about the values before time 1, but  $X_t = \sum_{n=0}^{\infty} C_n \varepsilon_{t-n}$  is stationary. The difference satisfies  $\text{Var}(X_t - X_t^+) = \sum_{n=t}^{\infty} C_n \Omega C_n' \rightarrow 0$ , as  $t \rightarrow \infty$ , and we say that  $X_t^+$  is asymptotically stationary.

We next turn to fractional processes of order  $d$ , see for example Brockwell and Davis (1991, Chapter 13.2) or Beran (1994). The binomial expansion

$$(1 - z)^{-d} = \sum_{n=0}^{\infty} (-1)^n \binom{-d}{n} z^n, \quad |z| < 1, \quad d \in R,$$

defines the coefficients

$$(-1)^n \binom{-d}{n} = \frac{d(d+1) \dots (d+n-1)}{n!} = \frac{\Gamma(d+n)}{\Gamma(d)\Gamma(n+1)},$$

which are  $O(n^{d-1})$ . This shows that  $\sum_{n=0}^{\infty} \binom{-d}{n}^2 < \infty$  for  $d < 1/2$ , so that for a sequence of i.i.d. variables  $\varepsilon_t$  with mean zero and finite variance we can define

$$\Delta^{-d} \varepsilon_t = (1 - L)^{-d} \varepsilon_t = \sum_{n=0}^{\infty} (-1)^n \binom{-d}{n} \varepsilon_{t-n}, \quad d < \frac{1}{2}, \quad (8)$$

as a stationary process with finite variance. For  $d \geq 1/2$  the infinite sum does not exist, but we can define a non-stationary process by the operator  $\Delta_+^{-d}$ , see Appendix A.4,

$$\Delta_+^{-d} \varepsilon_t = \sum_{n=0}^{t-1} (-1)^n \binom{-d}{n} \varepsilon_{t-n} = \varepsilon_t + d\varepsilon_{t-1} + \dots + (-1)^{t-1} \binom{-d}{t-1} \varepsilon_1, \quad t = 1, \dots, T,$$

see for instance Robinson and Marinucci (2001) who use the notation  $\Delta^d \varepsilon_t 1\{t \geq 1\}$ .

DEFINITION 2 *We say that  $X_t$  is fractional of order  $d$ , and write  $X_t \in F(d)$  if conditionally on the past  $\{X_s, s \leq 0\}$ ,  $\Delta_+^d X_t - \mu_t \in F(0)_+$  for some function  $\mu_t$  of the past.*

Note that we do not assume that all components of  $X_t$  have the same order of fractionality. This is in line with the definition that a process is  $I(1)$  if the difference is  $I(0)$ . This allows the components to be either  $I(1)$  or  $I(0)$  in the cointegrated VAR.

The term  $\mu_t$  is needed because when solving difference equations, we get a contribution from initial values, see Appendix A.5. If we condition on these, we get a deterministic term.

EXAMPLE 2. Let  $\varepsilon_t$  be i.i.d. in  $p$  dimensions with mean zero and finite variance. The equations

$$\Delta^d X_t = \varepsilon_t, \quad t = 1, \dots, T \quad (9)$$

have the interpretation

$$\begin{aligned} E_{t-1} X_t &= (1 - \Delta^d) X_t = L_d X_t, \\ \text{Var}_{t-1} X_t &= \Omega, \end{aligned}$$

where the subscript indicates that we take conditional mean and variance given the past  $\{X_s, s \leq t-1\}$ . Note that for the equations to make sense we must assume that the initial values of  $X_s, s \leq 0$  are such that the expression  $\Delta^d X_t = \sum_{n=0}^{\infty} (-1)^n \binom{d}{n} X_{t-n}$  converges. Obviously this would be the case if all initial values were zero, but for applications that is probably not so useful.

The equations can be solved, see Appendix A.5, by applying  $\Delta_+^{-d}$  to both sides of the equations. We then find, using  $\Delta_+^{-d} \Delta_+^d = 1_+$ ,

$$X_t = \mu_t + \Delta_+^{-d} \varepsilon_t, \quad t = 1, \dots, T,$$

where  $\mu_t = -\Delta_+^{-d} \Delta_-^d X_t = E_0 X_t$  is a function of initial values  $\{X_s, s \leq 0\}$ .

When  $d < 1/2$ , the stationary process  $X_t^* = \Delta^{-d} \varepsilon_t$ , see (8), with spectrum

$$f_X(\lambda) = \frac{1}{2\pi} |1 - e^{-i\lambda}|^{-2d} \Omega,$$

is a solution of (9). We note that  $X_t^*$  and  $X_t$  are fractional of order  $d$ , and that the spectral density of  $X_t^*$  has a pole of order  $2d$  at  $\lambda = 0$  for  $d > 0$ , in the sense that  $\lambda^{2d} f_X(\lambda) \rightarrow \frac{1}{2\pi} \Omega > 0$ , as  $\lambda \rightarrow 0$ . ■

DEFINITION 3 *If  $X_t \in F(d)$  and there exists a vector  $\beta$  so that  $\beta' X_t \in F(d-b)$  for some  $b, 0 < b \leq d$ , we call  $X_t$  cofractional with cofraction vector  $\beta$ .*

Note that the zero vector is not a cofraction vector by this definition, because the process  $X_t = 0$  is not fractional of any order. The linear combination  $\beta$  could be the unit vector  $(1, 0, 0, \dots, 0)'$  in which case the first component of  $X_t$  is  $\mathcal{F}(d-b)$ .

The definition of fractional and cofractional process used here are chosen with the following point of view in mind. When we fit the autoregressive fractional model we first check that the model type as defined by  $d, b, k$ , and the error structure is adequate. Based on the model we then make inference about the rank of  $\alpha\beta'$  and



the roots of the characteristic function. If we find  $r < p$ , we infer that there are unit roots and that there are linear combinations of the data that are fractional of order  $d - b$ . Only then do we start the identification of the cofractional vectors  $\beta$ , which of course may contain a unit vector.

A process  $Y_t \in \mathcal{F}(d)$  is also called fractionally integrated, or just integrated, of order  $d$ , and fractionally cointegrated, or just cointegrated,  $\mathcal{CI}(d, b)$ , but integration gives connotations of nonstationarity and it is therefore perhaps better to use fractional and cofractional. The linear combination  $\beta' X_t$  does not eliminate nonstationarity but reduces the order of fractionality, that is, the pole at zero frequency of the spectrum.

**EXAMPLE 3.** As a simple example of fractional and cofractional processes we consider the process  $X_t = (X_{1t}, X_{2t}, X_{3t})'$  defined by its moving average representation for  $t = 1, \dots, T$

$$\begin{aligned} X_{1t} &= \Delta_+^{-0.4} \varepsilon_{1t} - \Delta_+^{-0.2} \varepsilon_{2t} + \varepsilon_{3t}, \\ X_{2t} &= \Delta_+^{-0.4} \varepsilon_{1t} + \Delta_+^{-0.2} \varepsilon_{2t} + \varepsilon_{3t}, \\ X_{3t} &= \Delta_+^{-0.2} \varepsilon_{1t} + \varepsilon_{2t} + \varepsilon_{3t}. \end{aligned}$$

where  $\varepsilon_t$  is i.i.d.  $(0, I_3)$ . The transfer function for  $\Delta^{0.4} X_t$  at  $\lambda = 0$  is

$$C(1) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

and the spectrum is  $C(1)'C(1) \neq 0$ , so that  $\Delta^{0.4} X_t$  is fractional of order zero according to Definition 1, and  $X_t$  is fractional of order 0.4, according to Definition 2. The vectors

$$\beta' = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

define the bivariate process  $Y_t = \beta' X_t$  which is fractional of order 0.2, so that  $\beta$  is a cofraction vector, see Definition 3. Note that  $X_{3t}$  is fractional of a lower order than  $X_t$ , and that is expressed by  $\beta$  containing the unit vector  $(0, 0, 1)$ .

Finally note that there is another type of fractionality in this example because if

$$\beta(L)' = \begin{pmatrix} \Delta^{0.2} & 0 & -1 \\ 0 & \Delta^{0.2} & -1 \end{pmatrix},$$

then  $\beta(L)' X_t$  is fractional of order 0, so we have polynomial cofractionality between levels  $X_t$  and differences  $\Delta^{0.2} X_t$ . ■

We conclude this section with a definition of polynomial cofractionality, to which we shall return in section 4.

**DEFINITION 4** *If  $X_t \in F(d)$  and  $\beta' X_t \in F(d - b)$  for some  $b$ ,  $0 < b \leq d$ , and we can find  $\gamma_1 \neq 0$  and  $\gamma_2$  so that  $\gamma_1' \beta' X_t + \gamma_2' \Delta^b X_t \in F(c)$  for some  $0 \leq c < d - b$ , then we say that  $X_t$  has a polynomial cofraction vector  $\gamma_1' \beta' + \gamma_2' \Delta^b$ .*

Note that if  $X_t \in \mathcal{F}(d)$ , then of course  $\Delta^b X_t \in \mathcal{F}(d - b)$ , but this is not called polynomial cofractionality, because the levels have to enter a cofractional relation.

The models we propose only allow for fractional processes of order  $d$  and  $d - b$ , and polynomial cofractional vectors of order  $d - 2b$ . This allows for a simpler representation theory and a simpler statistical theory.

### 3 AUTOREGRESSIVE MODEL FOR COFRACTIONAL PROCESSES

In this section we first analyze the univariate case with one lag for  $b = d$ , which we call  $AR_d(1)$ , and find a criterion for the solution to be  $\mathcal{F}(0)$ . Next we analyze the model  $VAR_{d,b}(k)$  and find criteria for the solution to be a cofractional process.

#### 3.1 The $AR_d(1)$ model and a condition for $\mathcal{F}(0)$

As the simplest case of (6) we take  $b = d$ ,  $k = 0$ , and  $p = 1$ , and the model is

$$\Delta^d X_t = \gamma(1 - \Delta^d)X_t + \varepsilon_t, \quad t = 1, \dots, T.$$

or

$$X_t = \rho L_d X_t + \varepsilon_t = \rho(1 - (1 - L)^d)X_t + \varepsilon_t, \quad t = 1, \dots, T, \quad (10)$$

with  $\rho = 1 + \gamma$ . We assume that  $\varepsilon_t$  is univariate i.i.d.  $(0, \sigma^2)$ . Note that for  $\rho = 1$  we get a model with unit root of fractional multiplicity:  $\Delta^d X_t = \varepsilon_t$ . First, however, we want to find the set of  $\rho$  for which (10) has a solution which is fractional of order zero. We define the characteristic function, which is a polynomial only if  $d$  is a nonnegative integer,

$$\pi(z) = 1 - \rho(1 - (1 - z)^d).$$

The equations (10) determine the value of  $X_t$  as a function of  $\varepsilon_1, \dots, \varepsilon_t$ , and the infinitely many initial values by successive substitution and the solution is given in Theorem 5. We allow any complex value of  $\rho$  so  $X_t$  may be complex. In order to discuss the order of fractionality of (10), we need the set  $\mathbb{C}_d$  defined as the image of the unit disk under the mapping  $z \mapsto 1 - (1 - z)^d$ , see Appendix A.3 and Figure 1.

<Figure 1 about here>

The main result in Theorem 5 generalizes the well known case of  $d = 1$ , where the solution of  $X_t = \rho X_{t-1} + \varepsilon_t$  is given by  $X_t = \rho^t X_0 + \sum_{n=0}^{t-1} \rho^n \varepsilon_{t-n}$  for any  $\rho$ , and if  $|\rho| < 1$  the process  $X_t^* = \sum_{n=0}^{\infty} \rho^n \varepsilon_{t-n}$  is a stationary solution, and if  $\rho = 1$  we get the random walk  $\sum_{i=1}^t \varepsilon_i + X_0$ . For the fractional process in Theorem 5, the stationarity condition  $|\rho| < 1$  is replaced by  $\rho^{-1} \notin \mathbb{C}_d$ , which reduces to  $(1 - 2^d)^{-1} < \rho < 1$  if  $\rho$  is real. Note that the set of parameter values that give an  $\mathcal{F}(0)$  process depends on  $d$ .

**THEOREM 5** 1. For any  $\rho \in \mathbb{C}$ , there exists a  $\delta > 0$ , so that  $\pi(z)^{-1} = \sum_{n=0}^{\infty} \psi_n z^n$ , for  $|z| < \delta$ . The solution of (10) has the representation

$$X_t = \pi_+(L)^{-1} \varepsilon_t + \mu_t = \sum_{n=0}^{t-1} \psi_n \varepsilon_{t-n} + \mu_t, \quad t = 1, 2, \dots \quad (11)$$

where  $\mu_t = -\pi_+^{-1}(L)\pi_-(L)X_t$  is a function of infinitely many initial values  $X_0, X_{-1}, \dots$ .

2. If  $\rho^{-1} \notin \mathbb{C}_d$ , see Appendix 6.3,  $\pi(z)^{-1}$  is bounded and continuous on the closed unit disk, and has the expansion  $\pi(z)^{-1} = \sum_{n=0}^{\infty} \psi_n z^n$  for  $|z| < 1$ , where  $\sum_{n=0}^{\infty} |\psi_n|^2 < \infty$ . It follows that

$$X_t^* = \pi(L)^{-1} \varepsilon_t = \sum_{n=0}^{\infty} \psi_n \varepsilon_{t-n} \quad (12)$$

is stationary with mean zero and finite variance. The process  $X_t^*$  is a solution of (10) for all  $t$  with spectrum

$$f_{X^*}(\lambda) = \frac{1}{2\pi} \frac{\sigma^2}{|1 - \rho(1 - (1 - e^{-i\lambda})^d)|^2}, \quad (13)$$

so that  $f_{X^*}(0) > 0$ , and  $X_t^* \in \mathcal{F}(0)$ .

3. If  $\rho = 1$ , then  $\pi(1) = 0$  and

$$X_t = \Delta_+^{-d} \varepsilon_t + \mu_t, \quad t = 1, 2, \dots \quad (14)$$

where  $\mu_t = -\Delta_+^{-d} \Delta_-^d X_t$  is a function of initial values, and  $X_t \in \mathcal{F}(d)$ .

**Proof.** Because  $\pi(0) \neq 0$ , we can expand around  $z = 0$  in a neighborhood,  $|z| < \delta$ , so that  $\psi(z) = \sum_{n=0}^{\infty} \psi_n z^n$ ,  $|z| < \delta$ . We introduce the operators

$$1_+ x_t = x_t 1_{\{T \geq 1\}} = \begin{cases} x_t, & t = 1, 2, \dots \\ 0, & t = \dots, -1, 0 \end{cases}$$

and  $1_- = 1 - 1_+$ ,  $\pi_+(L) = \pi(L)1_+$ , and  $\pi_-(L) = \pi(L)1_-$ , see Appendix A.4. The equations are

$$\pi(L)X_t = \pi_+(L)X_t + \pi_-(L)X_t = \varepsilon_t, \quad t = 1, \dots, T.$$

Now apply  $\pi_+(L)^{-1}$  on both sides, using  $\pi_+(L)^{-1}\pi_+(L) = 1_+$ , and we find, with  $X_t^+ = X_t 1_+$ ,

$$X_t^+ + \pi_+(L)^{-1}\pi_-(L)X_t = \pi_+(L)^{-1}\varepsilon_t = \sum_{n=0}^{t-1} \psi_n \varepsilon_{t-n}.$$

Finally define  $\mu_t = -\pi_+(L)^{-1}\pi_-(L)X_t$ , which is a function of initial values. This completes the proof of (11).

If  $\rho^{-1} \notin \mathbb{C}_d$ , the equation  $\pi(z) = 0$  (or  $\rho^{-1} = 1 - (1 - z)^d$ ) has no solution for  $|z| \leq 1$ , and  $\psi(z)$  is regular for  $|z| < 1$  and continuous for  $|z| \leq 1$ , when  $d > 0$ . Thus the expansion of  $\psi(z)$  is valid for  $|z| < 1$ , and by Lemma 10, the continuity on  $|z| \leq 1$  implies that  $\sum_{n=0}^{\infty} |\psi_n|^2 < \infty$ . We use the coefficients  $\psi_n$  to construct the process

$$X_t^* = \pi(L)^{-1} \varepsilon_t = \sum_{n=0}^{\infty} \psi_n \varepsilon_{t-n},$$

which is a stationary process with mean zero, finite variance, and spectrum given by (13). The value for  $\lambda = 0$ , is  $\sigma^2 |1 - \rho|^{-2} / 2\pi > 0$ , so that  $X_t^*$  is  $\mathcal{F}(0)$ . Because the coefficients  $\psi_n$  are the coefficients in  $\pi(z)^{-1}$ , it follows that  $X_t^*$  solves (10) for all  $t$ . Finally, (14) is a special case of (11) using  $(1 - z)^{-1} = \sum_{n=0}^{\infty} z^n$ ,  $|z| < 1$ . Equation (14) shows that  $X_t \in \mathcal{F}(d)$ . ■

EXAMPLE 4. Consider the special case with  $d = 0.4$ , that is, the model

$$X_t = \rho(1 - (1 - L)^{0.4})X_t + \varepsilon_t, \quad t = 1, \dots, T.$$

The boundary of the set  $\mathbb{C}_{0.4}$  intersects the negative axis at the point  $1 - 2^{0.4} = -0.32$ . If we take  $\rho = -2$ , then  $\rho^{-1} = -0.5$  is outside  $\mathbb{C}_{0.4}$ , see Figure 1, and the process  $X_t$  is  $\mathcal{F}(0)$ . Note that there are values of  $\rho$  arbitrarily close to  $z = 1$ , and inside the unit circle, for which the process (10) is  $\mathcal{F}(0)$  and hence stationary. Finally for  $\rho = 1$ , we find  $\Delta^{0.4} X_t = \varepsilon_t$  which admits a solution  $X_t^* = \sum_{k=0}^{\infty} (-1)^k \binom{-0.4}{k} \varepsilon_t$  with spectrum  $\frac{\sigma^2}{2\pi} |1 - e^{-i\lambda}|^{-0.8} \propto \lambda^{-0.8}$ ,  $\lambda \rightarrow 0$ , so that although  $X_t$  is stationary it is fractional of order 0.4. ■

Theorem 5 can easily be extended to the model for the  $p$ -dimensional process given by  $A(L_d)X_t = \varepsilon_t$ . We formulate the result in

COROLLARY 6 *Let  $X_t$  be given by  $A(L_d)X_t = \sum_{i=0}^k A_i L_d^i X_t = \varepsilon_t$ ,  $t = 1, \dots, T$ . If the roots of  $|\sum_{i=0}^k A_i u^i| = 0$  are outside  $\mathbb{C}_d$ , then  $X_t$  is  $\mathcal{F}(0)$ .*

**Proof.** The stacked process  $\tilde{X}_t = (X_t', L_d X_t', \dots, L_d^{k-1} X_t')'$  is a  $VAR_d(1)$  process with coefficient matrix  $\tilde{A}$ , say. By a linear transformation we can reduce the system to its Jordan form, where it is seen, using Theorem 5, that the condition on the (possibly complex) roots means that all components are  $\mathcal{F}(0)$ . ■

The proof of this result can also be given as the proof of Theorem 8 where we study the cofractional model  $VAR_{d,b}(k)$ .

We conclude this section by giving a criterion for checking the condition that  $\rho^{-1} \notin \mathbb{C}_d$ . The boundary of  $\mathbb{C}_d$  is described by the curves

$$\begin{aligned} x(\psi) &= 1 - (2 \cos \psi)^d \cos(d\psi), \\ y(\psi) &= \pm (2 \cos \psi)^d \sin(d\psi), \end{aligned}$$

for  $0 \leq \psi \leq \min(\frac{\pi}{d}, \frac{\pi}{2})$ . Let  $\rho^{-1} = x + iy$ . Because of the symmetry of  $\mathbb{C}_d$ , there is no loss of generality in assuming that  $y > 0$ .

LEMMA 7 *The function*

$$g(\psi) = \frac{x(\psi)}{y(\psi)} = \frac{1 - (2 \cos \psi)^d \cos(d\psi)}{(2 \cos \psi)^d \sin(d\psi)}, \quad 0 < \psi < \min\left(\frac{\pi}{d}, \frac{\pi}{2}\right).$$

is strictly increasing with  $\mathbb{R}$ . Thus for any  $(x, y)$  with  $y > 0$ , we can determine  $\psi_u$  so that

$$g(\psi_u) = x/y.$$

Then

$$\rho^{-1} \notin \mathbb{C}_d \text{ if and only if } y(\psi_u) < y.$$

**Proof.** Note that for  $0 < \psi < \min(\frac{\pi}{d}, \frac{\pi}{2})$ ,  $y(\psi) > 0$ , see Figure 1, so that  $g(\psi)$  is well defined. It is seen that  $\psi \rightarrow 0$  implies  $g(\psi) \rightarrow -\infty$ , and  $\psi \rightarrow \min(\frac{\pi}{2}, \frac{\pi}{d})$  implies  $g(\psi) \rightarrow \infty$ , so the range is  $\mathbb{R}$ . Next we find the derivatives

$$\begin{aligned} \dot{x}(\psi) &= \frac{d}{d\psi}x(\psi) = d2^d(\cos \psi)^{d-1} \sin((1+d)\psi), \\ \dot{y}(\psi) &= \frac{d}{d\psi}y(\psi) = d2^d(\cos \psi)^{d-1} \cos((1+d)\psi), \end{aligned}$$

and

$$\begin{aligned} \dot{g}(\psi) &= (y(\psi)\dot{x}(\psi) - \dot{y}(\psi)x(\psi))/y^2 \\ &= [(2 \cos \psi)^d \sin(d\psi)d2^d(\cos \psi)^{d-1} \sin((1+d)\psi) \\ &\quad - (1 - (2 \cos \psi)^d \cos(d\psi))d2^d(\cos \psi)^{d-1} \cos((1+d)\psi)]/y^2 \\ &= d2^d(\cos \psi)^{d-1}[2^d(\cos \psi)^{d+1} - \cos((1+d)\psi)]/y^2. \end{aligned}$$

This is positive by the trigonometric inequality in Lemma 11. ■

### 3.2 The cofractional $VAR_{d,b}(k)$ model

We next turn to the model  $VAR_{d,b}(k)$  and want to find the conditions under which the solution is fractional of order  $d$  and  $\beta'X_t$  fractional of order  $b$ . The model is

$$\Delta^d X_t = \alpha\beta' \Delta^{d-b} L_b X_t + \sum_{i=1}^k \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t, \quad t = 1, \dots, T,$$

where  $0 < b \leq d$ . We define the characteristic polynomial

$$\Pi(z) = (1-z)^d I_p - \alpha\beta'(1 - (1-z)^b)(1-z)^{d-b} - \sum_{i=1}^k \Gamma_i (1 - (1-z)^b)^i (1-z)^d$$

and the polynomial

$$\Pi^*(u) = (1-u)I_p - \alpha\beta'u - \sum_{i=1}^k \Gamma_i (1-u)u^i. \quad (15)$$

These are related via the transformation  $u = 1 - (1 - z)^b$  so that

$$\Pi(z) = (1 - z)^{d-b} \Pi^*(1 - (1 - z)^b), \quad |z| \leq 1.$$

Note that  $\Pi$  and  $\Pi^*$  have a unit root if  $r < p$  because  $\Pi(1) = \Pi^*(1) = -\alpha\beta'$  has determinant zero, and that  $\Pi^*(u)$  is the characteristic polynomial for the usual cointegration model ( $d = b = 1$ ,  $L_d = L$ ). We define  $\Gamma = I - \sum_{i=1}^k \Gamma_i$  and formulate the main result, where the conditions are given on the roots of the polynomial  $|\Pi^*(u)| = 0$ , and the coefficients of the model.

**THEOREM 8** *Assume that  $|\Pi^*(u)| = 0$  implies that either  $u = 1$  or  $u \notin \mathbb{C}_b$ , that  $\alpha$  and  $\beta$  have rank  $r < p$ , and that  $|\alpha'_\perp \Gamma \beta_\perp| \neq 0$ , then*

$$X_t = C \Delta_+^{-d} \varepsilon_t + \Delta_+^{-(d-b)} Y_t + \mu_t, \quad t = 1, \dots, T,$$

where

$$C = \beta_\perp (\alpha'_\perp \Gamma \beta_\perp)^{-1} \alpha'_\perp. \quad (16)$$

The process  $Y_t$  is fractional of order zero, and has continuous spectrum which at zero frequency is given by  $C^* \Omega C^{*'} / 2\pi \neq 0$ , where

$$C^* = -(\bar{\beta} \bar{\alpha}' + C \Gamma \bar{\beta} \bar{\alpha}' + \bar{\beta} \bar{\alpha}' \Gamma C + C (\Gamma \bar{\beta} \bar{\alpha}' \Gamma + \sum_{i=1}^k i \Gamma_i) C), \quad (17)$$

and  $\mu_t = -\Pi_+^{-1}(L) \Pi_-(L) X_t$  depends on initial values. Thus  $X_t$  is fractional of order  $d$ , while  $\Delta^b X_t$  and  $\beta' X_t$  are fractional of order  $d - b$ .

There is no  $\gamma' X_t$  which is fractional of lower order than  $d - b$ , and even though  $\beta' X_t$  and  $\Delta^b X_t$  are both  $\mathcal{F}(d - b)$  it holds that if  $\gamma'_1 \beta' X_t + \gamma'_2 \Delta^b X_t$  is fractional of lower order than  $d - b$ , then  $\gamma_1 = 0$  and  $\gamma_2 = \beta \xi$  for some  $\xi$ , so that  $X_t$  does not allow for cofractionality.

**Proof.** From Theorem 3 in Johansen (in press) it follows that because

$$|\alpha'_\perp \frac{d}{du} \Pi^*(u)|_{u=1} \beta_\perp| = |\alpha'_\perp (-I_p - \alpha \beta' + \sum_{i=1}^k \Gamma_i) \beta_\perp| = (-1)^{p-r} |\alpha'_\perp \Gamma \beta_\perp| \neq 0,$$

it holds that

$$\Pi^*(u)^{-1} = C \frac{1}{1-u} + C^* + (1-u)H(u), \quad 0 < |u-1| < \delta, \quad (18)$$

for some  $\delta > 0$ , where  $C$  and  $C^*$  are given by (16) and (17). Furthermore, the function  $H^*(u)$  defined by

$$H^*(u) = C^* + (1-u)H(u)$$

is regular in  $\mathbb{C}$  with poles at the roots where  $|\Pi^*(u)| = 0$ , but no singularity at  $u = 1$ . The function  $f(z) = 1 - (1 - z)^b$  is regular for  $|z| < 1$ , and continuous for  $|z| \leq 1$ ,

when  $b > 0$ . If the remaining roots are outside  $\mathbb{C}_b$ , then there is no  $z$  in the closed unit disk for which  $1 - (1 - z)^b = u_i$ . Hence the compound function

$$F(z) = H^*(f(z)) = H^*(1 - (1 - z)^b), \quad |z| \leq 1 \quad (19)$$

is continuous for  $|z| \leq 1$  and regular without singularities on the open unit disk,  $|z| < 1$ . It therefore has an expansion  $F(z) = \sum_{n=0}^{\infty} F_n z^n$ ,  $|z| < 1$ , where the coefficients satisfy  $\sum_{n=0}^{\infty} \|F_n\|^2 < \infty$ , see Lemma 10. We apply these coefficients to define  $Y_t = F(L)\varepsilon_t = \sum_{n=0}^{\infty} F_n \varepsilon_{t-n}$  as a stationary process with mean zero, finite variance, and continuous spectral density

$$f_Y(\lambda) = \frac{1}{2\pi} F(e^{-i\lambda}) \Omega F(e^{i\lambda})' = \frac{1}{2\pi} H^*(1 - (1 - e^{-i\lambda})^b) \Omega H^*(1 - (1 - e^{i\lambda})^b)'$$

For  $\lambda = 0$  we get

$$\frac{1}{2\pi} F(1) \Omega F(1)' = \frac{1}{2\pi} H^*(1) \Omega H^*(1)' = \frac{1}{2\pi} C^* \Omega C^*'$$

The inequality

$$\Omega - \alpha(\alpha' \Omega^{-1} \alpha)^{-1} \alpha' = \Omega \alpha_{\perp} (\alpha'_{\perp} \Omega \alpha_{\perp})^{-1} \alpha'_{\perp} \Omega \geq 0 \quad (20)$$

shows that

$$\beta' C^* \Omega C^* \beta > \beta' C^* \alpha (\alpha' \Omega^{-1} \alpha)^{-1} \alpha' C^* \beta = (\alpha' \Omega^{-1} \alpha)^{-1} > 0, \quad (21)$$

because  $\beta' C^* \alpha = -I_r$ , see (17). This shows that  $f_Y(0) \neq 0$ , so that  $Y_t \in \mathcal{F}(0)$ .

We find from (18) that

$$\Pi^{-1}(z) = C(1 - z)^{-d} + (1 - z)^{-d+b} F(z),$$

and we next apply  $\Pi_+^{-1}(L)$  to the equation  $\Pi(L)X_t = \varepsilon_t$  and find the solution

$$X_t^+ = C \Delta_+^{-d} \varepsilon_t + \Delta_+^{-(d-b)} Y_t - \Pi_+^{-1}(L) \Pi_-(L) X_t.$$

This shows that  $X_t$  is  $\mathcal{F}(d)$  because  $C \neq 0$ , and  $\Delta_+^{d-b} \beta' X_t = \beta' Y_t^+$  is in  $\mathcal{F}(0)_+$ , because  $Y_t \in \mathcal{F}(0)$ .

Next consider a linear combination  $\gamma' X_t$ . For this to be fractional of order less than  $d$ , it must satisfy  $\gamma' C = 0$ , that is,  $\gamma = \beta \xi$  for some  $\xi$ . Then we find

$$\gamma' X_t = \xi' \beta' X_t = \Delta_+^{-(d-b)} \xi' \beta' Y_t^+ + \xi' \beta' \mu_t.$$

The stationary process  $\xi' \beta' Y_t$  has spectrum at frequency zero which is bounded below by  $\xi' (\alpha' \Omega^{-1} \alpha)^{-1} \xi$ , see (21). Thus the spectrum is only zero if  $\xi = 0$ , which means that no linear combination has lower fractional order than  $d - b$ .

Finally consider

$$\begin{aligned} Z_t &= \gamma_1' \beta' X_t + \gamma_2' \Delta^b X_t \\ &= \Delta_+^{-(d-b)} [\gamma_1' \beta' Y_t^+ + \gamma_2' C \varepsilon_t^+ + \Delta_+^b \gamma_2' Y_t^+] + (\gamma_1' \beta' + \gamma_2' \Delta^b) \mu_t + \gamma_2' \Delta_-^b X_t. \end{aligned}$$

The spectral density of  $U_t = \gamma'_1 \beta' Y_t + \gamma'_2 C \varepsilon_t + \Delta^b \gamma'_2 Y_t$  at zero frequency is

$$f_U(0) = (\gamma'_1 \beta' C^* + \gamma'_2 C) \Omega (C' \gamma_2 + C^{*'} \beta \gamma_1).$$

Because of the inequality (20), this is bounded below by

$$(\gamma'_1 \beta' C^* + \gamma'_2 C) \alpha (\alpha' \Omega^{-1} \alpha)^{-1} \alpha' (C' \gamma_2 + C^{*'} \beta \gamma_1) = \gamma'_1 (\alpha' \Omega^{-1} \alpha)^{-1} \gamma_1$$

because  $C\alpha = 0$ , and  $\beta' C^* \alpha = -I_r$ , see (17). If  $Z_t$  were fractional of lower order than  $d - b$ , then  $f_U(0) = 0$ , which implies that  $\gamma_1 = 0$  and hence  $\gamma'_2 C = 0$ , so that  $\gamma_2 = \beta \xi$  for some  $\xi$ . ■

Note that the last result implies that model (6) does not allow for polynomial fractional cointegration, see Engsted and Johansen (1999) for the  $I(1)$  case. A model that allows this possibility is analyzed in the next section.

We could easily include the case  $r = p$  in the formulation and we would get  $C = 0$ , and  $C^* = -\bar{\beta} \bar{\alpha}' = -(\alpha \beta')^{-1} = \Pi(1)^{-1}$ . In this sense, the result of Theorem 8 contains Corollary 6, as a special case for  $r = p$ .

## 4 A MODEL FOR POLYNOMIAL COFRACTIONAL PROCESSES

We consider the  $VAR_{d,b}(k)$  model (6) reparametrized as

$$\Delta^d X_t = \Delta^{d-2b} (\alpha \beta' L_b X_t - \Gamma \Delta^b L_b X_t) + \sum_{i=1}^k \Psi_i \Delta^d L_b^i X_t + \varepsilon_t, \quad (22)$$

where we assume that  $\alpha$  and  $\beta$  have rank  $r < p$ , and that  $\alpha'_\perp \Gamma \beta_\perp = \xi \eta'$  of rank  $s < p - r$ , and that  $0 < 2b \leq d$ .

This is an analogue of the model for  $I(2)$  variables, which we get for  $d = 2$ ,  $b = 1$ ; see Johansen (1997). We define the directions

$$(\beta, \beta_1, \beta_2) = (\beta, \bar{\beta}_\perp \eta, \beta_\perp \eta_\perp), \quad (\alpha, \alpha_1, \alpha_2) = (\alpha, \bar{\alpha}_\perp \xi, \alpha_\perp \xi_\perp). \quad (23)$$

Note that  $(\beta, \beta_1, \beta_2)$  and  $(\alpha, \alpha_1, \alpha_2)$  are orthogonal decompositions of  $R^p$ . From (22) we find the characteristic function

$$\Pi(z) = (1-z)^d I_p - (1-(1-z)^b)(1-z)^{d-2b} (\alpha \beta' - \Gamma(1-z)^b) - \sum_{i=1}^k \Psi_i (1-z)^d (1-(1-z)^b)^i$$

defined for  $|z| \leq 1$ , and the polynomial  $\Pi^*(u)$ , see (15), which is reparametrized as

$$\Pi^*(u) = (1-u)^2 I_p - u \alpha \beta' + \Gamma u (1-u) - \sum_{i=1}^k \Psi_i (1-u)^2 u^i,$$

and related to  $\Pi(z)$  by

$$\Pi(z) = (1-z)^{d-2b} \Pi^*(u), \quad u = 1 - (1-z)^b.$$



**THEOREM 9** Assume that  $|\Pi^*(u)| = 0$  implies that either  $u = 1$  or  $u \notin \mathbb{C}_b$ , and that  $\alpha$  and  $\beta$  have rank  $r < p$ . If  $\alpha'_\perp \Gamma \beta_\perp = \xi \eta'$  has rank  $s < p - r$  and if the condition  $|\alpha'_2 \theta \beta_2| \neq 0$  holds, where

$$\theta = \Gamma \bar{\beta} \bar{\alpha}' \alpha \beta' + I_p - \sum_{i=1}^k \Psi_i - \Gamma,$$

then

$$X_t = C_2 \Delta_+^{-d} \varepsilon_t + C_1 \Delta_+^{-(d-b)} \varepsilon_t + \Delta_+^{-(d-2b)} Y_t^+ + \mu_t. \quad (24)$$

The matrices  $C_1$  and  $C_2$  are

$$C_2 = \beta_2 (\alpha'_2 \theta \beta_2)^{-1} \alpha'_2, \quad (25)$$

$$\begin{aligned} C_1 = & -\bar{\beta}_1 \bar{\alpha}'_1 + [\bar{\beta}_1 \bar{\alpha}'_1 \theta + \bar{\beta} \bar{\alpha}' \Gamma] C_2 + C_2 [\theta \bar{\beta}_1 \bar{\alpha}'_1 + \Gamma \bar{\beta} \bar{\alpha}'] \\ & - C_2 [\Gamma \bar{\beta} \bar{\alpha}' \theta + \theta \bar{\beta}_1 \bar{\alpha}'_1 \theta + \theta \bar{\beta} \bar{\alpha}' \Gamma + \Gamma \bar{\beta} \bar{\alpha}' \Gamma + \Gamma \bar{\beta} \bar{\alpha}' \Gamma \bar{\beta} \bar{\alpha}' \Gamma + \sum_{i=1}^k i \Psi_i] C_2 \end{aligned} \quad (26)$$

and  $\mu_t = -\Pi_+(L)^{-1} \Pi_-(L) X_t$  depends on initial values. The process  $Y_t$  is stationary with continuous spectrum, and  $X_t$  is fractional of order  $d$ ,  $(\beta, \beta_1)' X_t$  is fractional of order  $d - b$ , and  $\beta' X_t - \bar{\alpha}' \Gamma \Delta^b X_t$  is fractional of order  $d - 2b$ .

**Proof.** From Theorem 5 of Johansen (in press) it follows that

$$|\alpha'_2 (\frac{d}{du} \Pi^*(u)|_{u=1} \bar{\beta} \bar{\alpha}' \frac{d}{du} \Pi^*(u)|_{u=1} + \frac{1}{2} \frac{d^2}{du^2} \Pi^*(u)|_{u=1}) \beta_2| = |\alpha'_2 \theta \beta_2| \neq 0$$

implies that

$$\Pi^*(u)^{-1} = C_2 \frac{1}{(1-u)^2} + C_1 \frac{1}{1-u} + C_0 + (1-u)H(u), \quad 0 < |u-1| < \delta, \quad (27)$$

where  $C_2$  and  $C_1$  are given by (25) and (26). No complete expression was found for  $C_0$ , but it was shown that

$$\beta' C_0 \alpha = -I_r + \bar{\alpha}' \Gamma C_2 \Gamma \bar{\beta}. \quad (28)$$

Furthermore, the function

$$H^*(u) = C_0 + (1-u)H(u)$$

is regular without a pole at  $u = 1$ . Under the condition that the remaining roots of  $|\Pi^*(u)| = 0$  are outside  $\mathbb{C}_b$ ,  $H^*(u)$  is regular without singularities inside  $\mathbb{C}_b$  and continuous on  $\mathbb{C}_b$ . We define

$$F(z) = H^*(1 - (1-z)^b)$$

for  $|z| \leq 1$ . By Lemma 10 this is regular for  $|z| < 1$  and continuous on the boundary when  $d > 0$ , so that  $Y_t = \sum_{n=0}^{\infty} F_n \varepsilon_{t-n}$  is a stationary process with continuous spectrum, where  $F(z) = \sum_{n=0}^{\infty} F_n z^n$ ,  $|z| < 1$ . We then find

$$\Pi^{-1}(z) = (1-z)^{-(d-2b)} \Pi^*(u)^{-1} = C_2(1-z)^{-d} + C_1(1-z)^{-(d-b)} + (1-z)^{-(d-2b)} F(z)$$

We construct the solution of the equations  $\Pi(L)X_t = \varepsilon_t$  by applying  $\Pi_+^{-1}(L)$  and find

$$X_t^+ = C_2 \Delta_+^{-d} \varepsilon_t + C_1 \Delta_+^{-(d-b)} \varepsilon_t + \Delta_+^{-(d-2b)} Y_t^+ - \Pi_+(L)^{-1} \Pi_-(L) X_t,$$

which proves (24). It is seen that  $X_t$  is  $\mathcal{F}(d)$  because  $C_2 \neq 0$ , that  $(\beta, \beta_1)' X_t$  is  $\mathcal{F}(d-b)$  because  $(\beta, \beta_1)' C_2 = 0$ , and  $(\beta, \beta_1)' C_1 \alpha_1 = (0, -I_s) \neq 0$ , see (25) and (26).

We next consider polynomial cofractionality and want to show that  $\beta' X_t - \bar{\alpha}' \Gamma \Delta^b X_t \in \mathcal{F}(d-2b)$ , and have to show that the process  $\Delta_+^{d-2b} (\beta' X_t - \bar{\alpha}' \Gamma \Delta^b X_t) = \Delta_+^{d-2b} (\beta' X_t^+ - \bar{\alpha}' \Gamma \Delta_+^b X_t^+ - \bar{\alpha}' \Gamma \Delta_-^b X_t)$  is in  $\mathcal{F}(0)_+$ . We find from (24) that

$$\begin{aligned} & \Delta_+^{d-2b} (\beta' X_t^+ - \bar{\alpha}' \Gamma \Delta_+^b X_t^+ - \bar{\alpha}' \Gamma \Delta_-^b X_t) \\ &= (\beta' C_1 - \bar{\alpha}' \Gamma C_2) \Delta_+^b \varepsilon_t + (\beta' Y_t^+ - \bar{\alpha}' \Gamma C_1 \varepsilon_t^+ - \bar{\alpha}' \Gamma \Delta_+^b Y_t^+) \\ &+ (\Delta_+^{d-2b} \beta' - \bar{\alpha}' \Gamma \Delta_+^{d-b}) \mu_t - \Delta_+^{d-2b} \bar{\alpha}' \Gamma \Delta_-^b X_t. \end{aligned}$$

From (25) and (26) it follows that  $\beta' C_1 - \bar{\alpha}' \Gamma C_2 = 0$ , so that the first parenthesis is zero. In order to see that  $\Delta_+^{d-2b} (\beta' X_t - \bar{\alpha}' \Gamma \Delta^b X_t)$  is in  $\mathcal{F}(0)_+$ , we find the spectrum at zero frequency for the process  $\beta' Y_t - \bar{\alpha}' \Gamma C_1 \varepsilon_t - \bar{\alpha}' \Gamma \Delta^b Y_t$ , which is proportional to

$$(\beta' C_0 - \bar{\alpha}' \Gamma C_1) \Omega (\beta' C_0 - \bar{\alpha}' \Gamma C_1)'$$

By the inequality (20), this is bounded below by

$$(\beta' C_0 - \bar{\alpha}' \Gamma C_1) \alpha (\alpha' \Omega^{-1} \alpha)^{-1} \alpha' (C_0' \beta - C_1' \Gamma' \bar{\alpha}).$$

From (28) and (26) it follows that

$$\beta' C_0 \alpha - \bar{\alpha}' \Gamma C_1 \alpha = -I_r + \bar{\alpha}' \Gamma C_2 \Gamma \bar{\beta} - \bar{\alpha}' \Gamma C_2 \Gamma \bar{\beta} = -I_r,$$

so that  $\beta' Y_t - \bar{\alpha}' \Gamma C_1 \varepsilon_t - \bar{\alpha}' \Gamma \Delta^b Y_t$  is  $\mathcal{F}(0)$ . Note that  $Y_t$  is stationary with spectrum at zero frequency given by  $C_0 \Omega C_0' / 2\pi$ , and hence probably non-zero so that  $Y_t$  is probably  $\mathcal{F}(0)$ . We do not have a complete expression for  $C_0$ , only an expression for  $\beta' C_0 \alpha$  which is enough to prove what we need, namely that  $\beta' X_t - \bar{\alpha}' \Gamma \Delta^b X_t$  is  $\mathcal{F}(d-2b)$ . ■

## 5 CONCLUSION

We have defined a class of vector autoregressive models,  $VAR_{d,b}(k)$ , that generate fractional, cofractional, and polynomial cofractional processes. The models have the advantage that one can give simple criteria in terms of the parameters for fractionality, cofractionality, and polynomial cofractionality. In this way we have a class of

models for cofractional processes and their adjustment to equilibrium relations, so that economic hypotheses can be formulated within the model, which can be tested against the data. Thus we have a platform for making model based inference on coefficients, relations, and fractional order. Inference for these model has still to be worked out.

## A APPENDIX

In this Appendix we first discuss regular functions in particular the function  $f(z) = 1 - (1 - z)^d$ , and the image of the unit circle under the mapping  $f(z)$ . We introduce some useful operators for solving difference equations involving fractional differences. Finally we give a result from Grenander and Rosenblatt (1956), which is the key to understanding the structure solution of the model equations.

### A.1 Regular functions

A regular (or holomorphic) function, see Phillips (1958) is defined as a complex valued differentiable function on an open (and arc connected) set  $\mathbb{D}$  of  $\mathbb{C}$ . We give two examples.

Consider first the function  $g(z) = 1/(1-3z)$ . This is defined for  $z \in \mathbb{C} \setminus \{1/3\}$ , and  $|g(z)| \rightarrow \infty$ ,  $z \rightarrow 1/3$ . Note that  $(1-3z)g(z)$ ,  $z \neq 1/3$ , can be extended by continuity to the point  $z = 1/3$ , and the result is a regular function defined on  $\mathbb{C}$ . Such a point is called an isolated singularity and  $(1 - 3z)g(z)$  has a removable singularity at  $z = 1/3$ , whereas  $g(z)$  has a pole of order one. The regular function  $g(z)$  can be expanded in a Taylor's series at all other points except  $z = 1/3$ . For instance, around  $z = 0$  we get the series  $g(z) = \sum_{n=0}^{\infty} 3^n z^n$ ,  $|z| < 1/3$ , with exponentially increasing coefficients, whereas around  $z = 2$ , the function can be represented by the series  $g(z) = -\frac{1}{5} \sum_{n=0}^{\infty} (\frac{3}{5})^n (2 - z)^n$ ,  $|2 - z| < \frac{5}{3}$ , where the coefficients are exponentially decreasing. Thus  $g$  is regular on  $\mathbb{C} \setminus \{1/3\}$ .

Another example is the function  $f(z) = 1 - (1 - z)^{0.8}$ . In order to get a single valued function we remove from the complex plane the line  $\mathbb{L} = \{\text{Im}(z) = 0, \text{Re}(z) \geq 1\}$ . Then  $f(z)$  is well defined and regular on  $\mathbb{D} = \mathbb{C} \setminus \mathbb{L}$ . Moreover it can be extended by continuity to the point  $z = 1$ , but the extended function is not differentiable at  $z = 1$ . The point is not an isolated singularity.

The main result about a regular function is that at any point of the domain of definition it can be expanded in a Taylor's series which converges in the largest open disk that does not contain any singularity. Notice that a polynomial  $p(z) = \sum_{i=0}^k a_i z^i$  is a regular function on  $\mathbb{C}$ , so that in a sense the regular functions correspond to infinite order polynomials. Another fundamental result is that for two regular functions the composition  $f(g(z))$  is again regular provided the range of  $g$  is in the domain of  $f$ .

If  $\Pi(z)$  is a matrix polynomial, then  $\Pi^{-1}(z)$  is a regular function with isolated singularities (poles) at the roots of  $|\Pi(z)| = 0$ . The representation (18) shows that  $(1-u)\Pi^*(u)^{-1}$  has a removable singularity at  $u = 1$ , or that by subtracting  $C(1-u)^{-1}$

from  $\Pi^*(u)^{-1}$  we get a function with removable singularity. If there are no further poles in the unit disk, then  $\Pi^{-1}(z) - C(1-u)^{-d}$  has an expansion around zero with exponentially decreasing coefficients. For fractional processes, see (6), the function  $\Pi^{-1}(z)$  is regular but with a different type of singularity at  $z = 1$ , because we cannot move around the pole inside the domain of definition. The main trick in this paper is to define  $\Pi(z)$  as  $\Pi^*(1 - (1-z)^d)$ , so that by removing the poles from  $\Pi^{*-1}(u)$  we remove the cause for the singularity from  $\Pi^{-1}(z)$ . If we remove these singularities, the function is regular without singularities in the open unit disk and continuous on the whole disk, the result given in Section A.6 shows that this is enough to discuss the processes in the sense of  $\mathcal{L}_2$ .

## A.2 The function $z^d$

For  $d > 0$ , the function  $h(z) = z^d$ , is multi-valued with a singularity at  $z = 0$ . For any  $n = 0, \pm 1, \dots$ , the representation  $z = re^{i\theta} = re^{i\theta + 2\pi ni}$ , shows that

$$z^d = r^d e^{id\theta + 2\pi dni}, n = 0, \pm 1, \pm 2, \dots$$

has many different values unless  $d$  is an integer. In order to get a unique value we associate with a complex number  $z$  the main argument  $\text{Arg}(z)$ , which is the argument  $\theta$  for which

$$z = re^{i\theta}, \quad -\pi < \theta \leq \pi.$$

Then  $\text{Arg}(z)$  is discontinuous at the negative real line, because for  $z \rightarrow -1$  with  $\text{Im}(z) > 0$ , we have  $\text{Arg}(z) \rightarrow \pi$ , but for  $z \rightarrow -1$ , with  $\text{Im}(z) < 0$ , we have  $\text{Arg}(z) \rightarrow -\pi$ . For  $d > 0$ , we can define the power function  $z^d$  uniquely if we avoid the negative real axis. Thus for  $z = re^{i\theta}$ ,  $-\pi < \theta < \pi$  we define

$$z^d = r^d e^{id\theta} = |z|^d e^{id\text{Arg}(z)}.$$

Notice that we have  $z^a z^b = z^{a+b}$ , but in general  $(z^a)^b \neq z^{ab}$ , as the following example shows.

**EXAMPLE 5.** Let  $z = e^{i\frac{3\pi}{4}}$ , then with the above definition  $z^2 = e^{-i\frac{\pi}{2}}$ , so that  $(z^2)^{1/2} = e^{-i\frac{\pi}{4}}$ , whereas  $z^{1/2} = e^{i\frac{3\pi}{8}}$  implies that  $(z^{1/2})^2 = e^{i\frac{3\pi}{4}} \neq (z^2)^{1/2}$ . ■

A consequence is that the function  $1 - (1-z)^{-d}$  has singularities at the line

$$\mathbb{L} = \{\text{Im}(z) = 0, \text{Re}(z) \geq 1\},$$

but it is regular on  $\mathbb{C} \setminus \mathbb{L}$  and we can expand it around  $z = 0$ , and find

$$1 - (1-z)^{-d} = 1 - \sum_{n=0}^{\infty} (-1)^n \binom{-d}{n} z^n, \quad |z| < 1.$$

### A.3 The image $\mathbb{C}_d$ of the unit disk under $f(z) = 1 - (1 - z)^d$

We investigate here the image of the unit disk under the mapping  $f(z) = 1 - (1 - z)^d$  where  $0 < d < \infty$ . We have for  $0 \leq \theta \leq \pi$  and  $z = e^{i\theta}$ ,

$$\begin{aligned} 1 - z &= 1 - e^{i\theta} = 1 - \cos(\theta) - i \sin(\theta) \\ &= 2 \sin^2 \frac{\theta}{2} - 2i \sin \frac{\theta}{2} \cos \frac{\theta}{2} = 2 \sin \frac{\theta}{2} \left( \sin \frac{\theta}{2} - i \cos \frac{\theta}{2} \right) \\ &= 2 \left( \cos \frac{\pi - \theta}{2} \right) \left( \cos \frac{\pi - \theta}{2} - i \sin \frac{\pi - \theta}{2} \right) = 2(\cos \psi)(\cos \psi - i \sin \psi) \\ &= 2(\cos \psi)e^{-i\psi}, \quad \psi = \frac{1}{2}(\pi - \theta) \in [0, \frac{1}{2}\pi]. \end{aligned}$$

For  $-\pi \leq \theta \leq 0$  we get the same result but with  $\psi$  replaced by  $-\psi$ . We find for different  $d$  the curves given in Figure 1. A point on the curve is  $(x(\psi), y(\psi))$  given by

$$\begin{aligned} x(\psi) &= 1 - (2 \cos \psi)^d \cos(d\psi), \\ y(\psi) &= \pm (2 \cos \psi)^d \sin(d\psi). \end{aligned}$$

They intersect the real axis for  $\psi = 0$  and  $\psi = \min(\frac{\pi}{2}, \frac{\pi}{d})$ . Note that  $y(\psi)$  is positive for  $0 < \psi < \min(\frac{\pi}{d}, \frac{\pi}{2})$ .

### A.4 The operators $h_+(L)$ and $h_-(L)$

Let the complex function  $h$  be given by a convergent power series

$$h(z) = \sum_{n=0}^{\infty} h_n z^n, \quad |z| < 1,$$

for which  $h(0) = 1$ , so that  $h(z) \neq 0$ ,  $|z| < \delta$ . The function  $g(z) = 1/h(z)$  is given by a power series

$$g(z) = \sum_{n=0}^{\infty} g_n z^n, \quad |z| < \delta,$$

where the coefficients satisfy

$$\sum_{n+m=k} g_n h_m = 1_{\{k=0\}}. \quad (\text{A.1})$$

If  $\sum_{n=0}^{\infty} h_n x_{n-t}$  is well defined we define the operators

$$\begin{aligned} h_+(L)x_t &= h(L)(x_t 1_{\{t \geq 1\}}) = \begin{cases} \sum_{n=0}^{t-1} h_n x_{t-n}, & t = 1, 2, \dots \\ 0, & t = \dots, -1, 0 \end{cases}, \\ h_-(L) &= h(L) - h_+(L). \end{aligned}$$

Note that the operator  $h_+(L)$  is defined for any sequence because only finitely many terms occur. Similarly we define  $g_+(L)$  for any sequence and notice that it follows from (A.1) that

$$h_+(L)g_+(L)x_t = 1_+ x_t = g_+(L)h_+(L)x_t$$

for any sequence, whereas

$$h_-(L)g_+(L)x_t = 0. \quad (\text{A.2})$$

As an example we consider the function  $h(z) = 1 - z$  and  $h(L) = 1 - L = \Delta$ . The function is inverted as  $g(z) = 1/(1 - z) = \sum_{n=0}^{\infty} z^n$ ,  $|z| < 1$ , so that  $g_n = 1, n = 0, 1, \dots$ . Thus  $\Delta^{-1}x_t$  is only defined if  $\sum_{n=0}^{\infty} x_{t-n}$  is well defined, but  $\Delta_+^{-1}x_t = \sum_{i=1}^t x_i, t = 1, 2, \dots$  is always defined.

## A.5 Solving fractional difference equations

We can use the operators  $h_+$  and  $h_-$  to solve difference equations involving fractional differences. Thus for instance we can solve the difference equation

$$\Delta^d X_t = \varepsilon_t, \quad t = 1, \dots, T$$

as function of  $\varepsilon_1, \dots, \varepsilon_T$  and initial values  $\{X_s, s \leq 0\}$ . We cannot apply the operator  $\Delta^{-d}$  on both sides because it may not be defined on the sequence  $\varepsilon_t$ . We instead apply the operator  $\Delta_+^{-d}$  and use the relation  $\Delta_+^{-d}\Delta_+^d = 1_+ = 1\{t \geq 1\}$ , to get

$$\Delta_+^{-d}\Delta^d X_t = \Delta_+^{-d}(\Delta_+^d X_t + \Delta_-^d X_t) = X_t + \Delta_+^{-d}\Delta_-^d X_t = \Delta_+^{-d}\varepsilon_t, \quad t = 1, \dots, T,$$

which gives

$$X_t = \Delta_+^{-d}\varepsilon_t + \mu_t = \sum_{n=0}^{t-1} (-1)^n \binom{-d}{n} \varepsilon_{t-n} + \mu_t, \quad t = 1, \dots, T,$$

where

$$\mu_t = -\Delta_+^{-d}\Delta_-^d X_t = -\sum_{n=0}^{t-1} (-1)^n \binom{-d}{n} \sum_{k=t}^{\infty} (-1)^k \binom{d}{k} X_{t-n-k}$$

is a function of the infinitely many initial values  $X_0, X_{-1}, \dots$

In the special case  $d = 1$ , we find

$$\Delta_+ x_t = \begin{cases} x_t - x_{t-1}, & t = 2, 3, \dots \\ x_1, & t = 1 \\ 0, & t = \dots, -1, 0 \end{cases}, \quad \Delta_- x_t = \begin{cases} 0, & t = 2, 3, \dots \\ -x_0, & t = 1 \\ x_t - x_{t-1}, & t = \dots, -1, 0 \end{cases}$$

so that  $\mu_t = -\Delta_+^{-1}\Delta_- x_t = x_0 1_+$ , which gives the well known solution

$$x_t = x_0 + \sum_{i=1}^t \varepsilon_i.$$

## A.6 A result from spectral theory

We can infer the properties of the coefficients of a linear process by analyzing the transfer function, as the next lemma shows. The result is taken from Grenander and Rosenblatt (1956, page 288) and shows how continuity of the transfer function on the boundary of the unit disk can be used to prove the existence of a stationary process with finite variance, see also Brockwell and Davis (1991, Theorem 4.10.1).

LEMMA 10 *Let  $\mu$  be Lebesgue measure on the interval  $[0, 2\pi]$ . Consider the regular matrix valued function  $F(z)$  which has no singularities in the open unit disk, so that  $F(z) = \sum_{n=0}^{\infty} F_n z^n$ ,  $|z| < 1$ . Then  $F$  can be extended by  $\mathcal{L}_2(\mu)$  continuity to the boundary  $|z| = 1$  if and only if  $\sum_{n=0}^{\infty} \|F_n\|^2 < \infty$ .*

*In this case  $X_t = \sum_{k=0}^{\infty} F_k \varepsilon_{t-k}$  is a stationary process with mean zero and variance  $\sum_{k=0}^{\infty} F_k \Omega F_k'$ .*

*If in particular  $F(z)$  is continuous for  $|z| \leq 1$ , it is  $\mathcal{L}_2(\mu)$  continuous, and the spectral density is continuous and given by*

$$f(\lambda) = \frac{1}{2\pi} F(e^{-i\lambda}) \Omega F(e^{i\lambda})'.$$

**Proof.** For  $r < 1$ , we find, using the regular function  $F$ ,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} F(re^{-i\lambda}) F(re^{i\lambda})' d\lambda = \sum_{n,m=0}^{\infty} r^{n+m} \frac{1}{2\pi} \int_{-\pi}^{\pi} F_n F_m' e^{i\lambda(m-n)} d\lambda = \sum_{n=0}^{\infty} F_n F_n' r^{2n} \quad (\text{A.3})$$

First assume that  $F$  has been extended by  $\mathcal{L}_2(\mu)$  continuity to the boundary, so that  $F(e^{i*}) = \lim_{r \rightarrow 1} F(re^{i*})$ , where the limit is in the sense of  $\mathcal{L}_2(\mu)$ , that is,  $\int_{-\pi}^{\pi} |F(e^{i\theta}) - F(re^{i\theta})|^2 d\theta \rightarrow 0$ ,  $r \rightarrow 1$ . Then the left hand side of (A.3) converges towards the finite limit  $\frac{1}{2\pi} \int_{-\pi}^{\pi} F(e^{-i\lambda}) F(e^{i\lambda})' d\lambda$ . Hence the limit of the right hand side,  $\sum_{n=0}^{\infty} F_n F_n'$ , is finite.

Next, assume that  $\sum_{n=0}^{\infty} F_n F_n'$  is finite, then

$$\phi_k(\lambda) = F(r_k e^{i\lambda}), \quad r_k \rightarrow 1$$

is a Cauchy sequence in  $\mathcal{L}_2(\mu)$ , because by (A.3)

$$\int_{-\pi}^{\pi} (\phi_k(\lambda) - \phi_m(\lambda)) (\phi_k(-\lambda) - \phi_m(-\lambda))' d\lambda = \sum_{n=0}^{\infty} F_n F_n' (r_k^n - r_m^n)^2 \rightarrow 0$$

as  $r_k$  and  $r_m$  tend to one.

By completeness of  $\mathcal{L}_2(\mu)$  there is an  $\mathcal{L}_2$  limit  $\phi(*)$ , and we therefore extend  $F$  by  $\mathcal{L}_2(\mu)$ -continuity by defining  $F(e^{i\lambda}) = \phi(\lambda)$ .

Let  $\mathbb{P}$  denote the probability measure underlying the random sequence  $\varepsilon_t$ . If  $\sum_{n=0}^{\infty} F_n F_n' < \infty$ , then  $\sum_{n=0}^{\infty} F_n r_k^n \varepsilon_{t-n}$ ,  $k = 1, 2, \dots$  is an  $\mathcal{L}_2(\mathbb{P})$  Cauchy sequence for  $r_k \rightarrow 1$ , because

$$E\left(\sum_{n=m}^{\infty} F_n (r_k^n - r_m^n) \varepsilon_{t-n}\right)^2 = \sum_{n=m}^{\infty} F_n \Omega F_n' (r_k^n - r_m^n)^2 \rightarrow 0, \quad r_k, r_m \rightarrow 1.$$

Thus  $X_t$  exists as an element of  $\mathcal{L}_2(\mathbb{P})$  and the variance is as given.

If finally  $F(z)$  is continuous for  $|z| \leq 1$ , it is bounded and therefore  $\mathcal{L}_2(\mu)$  continuous. This gives the value of the continuous spectral density. ■

## A.7 A trigonometric inequality

LEMMA 11

$$(\cos \psi)^{d+1} \geq \cos((1+d)\psi), \quad 0 \leq \psi \leq \min\left(\frac{\pi}{2}, \frac{\pi}{d}\right), \quad d > 1. \quad (\text{A.4})$$

**Proof.** We apply a product representation of  $\cos \psi$ , see Gradshteyn and Ryzhik (1971, #1.431.3),

$$\cos \psi = \prod_{k=0}^{\infty} \left(1 - \frac{4\psi^2}{(2k+1)^2\pi^2}\right),$$

and we therefore want to prove

$$\prod_{k=0}^{\infty} \left(1 - \frac{4\psi^2}{(2k+1)^2\pi^2}\right)^{d+1} > \prod_{k=0}^{\infty} \left(1 - \frac{4\psi^2(d+1)^2}{(2k+1)^2\pi^2}\right). \quad (\text{A.5})$$

The function  $f(z) = (1-u)^{d+1} - (1-(d+1)^2u)$  is convex on  $[0, 1[$ ,  $f(0) = 0$ , and derivative  $f'(0) = (d+1)d > 0$ , for  $d > 0$ . Therefore  $f$  is strictly increasing and positive on the interval  $u \in [0, 1[$ . For each of the factors we then have

$$\left(1 - \frac{4\psi^2}{(2k+1)^2\pi^2}\right)^{d+1} > 1 - \frac{4\psi^2(d+1)^2}{(2k+1)^2\pi^2}, \quad k = 0, 1, \dots$$

The inequality (A.5) would follow if the factors were positive.

The right hand side is positive for  $k = 1, 2, \dots$  because

$$1 - \frac{4\psi^2(d+1)^2}{(2k+1)^2\pi^2} \geq 1 - \frac{4\psi^2(d+1)^2}{3^2\pi^2} \geq 1 - \frac{(d+1)^2}{3^2} \min\left(1, \frac{2}{d}\right)^2.$$

For  $0 < d \leq 2$  we find  $\min(1, \frac{2}{d}) = 1$ , and  $1 - \frac{(d+1)^2}{3^2} \geq 0$ , and for  $d \geq 2$ , we find  $\min(1, 2/d) = 2/d$  and  $1 - 4(d+1)^2/(3^2d^2) \geq 0$ .

Thus the result (A.4) holds when the factor with  $k = 0$  on the right hand side is positive, but when it is negative the inequality holds trivially. ■



## REFERENCES

- Anderson, T.W. (1971) *The Statistical Analysis of Time Series*. Wiley.
- Beran, J. (1994) *Statistics for Long-memory Processes*. Chapman & Hall.
- Breitung, J. & U. Hassler (2002) Inference on the Cointegration Rank in Fractionally Integrated Processes. *Journal of Econometrics* 110, 167–185.
- Brockwell, P.J. & R.A. Davis (1991) *Time Series: Theory and Methods*. (2. ed.) Springer.
- Davidson, J. (2002) A model of fractional cointegration, and tests for cointegration using the bootstrap. *Journal of Econometrics* 110, 187–212.
- Dittmann, I. (2004) Error correction models for fractionally cointegrated time series. *Journal of Time Series Analysis* 25, 27–32.
- Engle, R.F. & C.W.J. Granger (1987) Co-integration and error correction: representation, estimation and testing. *Econometrica* 55, 251–276.
- Engsted, T. & S. Johansen (1999) Granger’s representation theorem and multicointegration. In R.F. Engle & H. White (eds.) *Cointegration, Causality and Forecasting. A Festschrift in Honour of Clive W.J. Granger*. Oxford University Press, pp. 220–211.
- Gradshteyn, I.S., & I.M. Ryzhik (1971) *Tables of Integrals, Sums, Series and Products* (6th ed.). eds A. Jeffrey & D. Zwillinger. New York: Academic Press.
- Granger, C.W.J. (1986) Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics* 48, 213–228.
- Granger, C.W.J. & R. Joyeux (1980) An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis* 1, 15–29.
- Grenander, U. & M. Rosenblatt (1956) *Statistical Analysis of Stationary Time Series*. Almqvist and Wiksell.
- Hosking, J.R.M. (1981) Fractional differencing. *Biometrika* 165–176.
- Johansen, S. (1996) *Likelihood Based Inference on Cointegration in the Vector Autoregressive Model* (2. ed.). Oxford University Press.
- Johansen, S. (1997) Likelihood analysis of the  $I(2)$  model. *Scandinavian Journal of Statistics* 24, 433–462.
- Johansen, S. (in press) Representation of cointegrated autoregressive processes with application to fractional processes. *Econometric Review*.
- Lasak, K. (2005) Likelihood based testing for fractional cointegration. Discussion paper, University of Barcelona.
- Lyhagen, J. (1998) Maximum likelihood estimation of the multivariate fractional cointegrating model. Working paper, Stockholm School of Economics.
- Phillips, E.G. (1958) *Functions of a Complex Variable with Applications*. Oliver and Boyd.
- Robinson, P.M. & D. Marinucci (2001) Semiparametric fractional cointegration analysis. *Journal of Econometrics* 105, 225–247.

### Legend to Figure 1

The images  $\mathbb{C}_d$  of the unit disk under the mapping  $z \rightarrow 1 - (1 - z)^d$ , for four values of  $d$ . The curves are defined for  $0 \leq \psi \leq \pi/2$  by

$$\begin{aligned}x(\psi) &= 1 - (2 \cos \psi)^d \cos(d\psi), \\y(\psi) &= \pm(2 \cos \psi)^d \sin(d\psi),\end{aligned}$$

and intersect the real axis for  $\psi = 0$  and  $\psi = \min(\frac{\pi}{2}, \frac{\pi}{d})$ .

The image of the unit disk under the mapping  $1-(1-z)^d$   
for  $d = 0.7$  and  $d = 2.0$

