SAPIENZA
UNIVERSITÀ DI ROMA

Dipartimento di Scienze Statistiche
Sezione di Statistica Economica ed Econometria

Massimo Franchi Paolo Paruolo

A general inversion theorem for cointegration

DSS Empirical Economics and Econometrics Working Papers Series

DSS-E3 WP 2017/3

Dipartimento di Scienze Statistiche
Sezione di Statistica Economica ed Econometria
"Sapienza" Università di Roma
P.le A. Moro 5 – 00185 Roma - Italia

http://www.dss.uniroma1.it

A GENERAL INVERSION THEOREM FOR COINTEGRATION[1]

Massimo Franchi and Paolo Paruolo

June 7, 2017

A generalization of the Granger and the Johansen Representation Theorems valid for any (possibly fractional) order of integration is presented. This is based on an inversion theorem that characterizes the order of the pole and the coefficients of the Laurent series representation of the inverse of a matrix function around a singular point. Explicit expressions of the matrix coefficients of the (polynomial) cointegrating relations, of the common trends and of the triangular representations are provided, either starting from the Moving Average or the Auto Regressive form. This unifies the different approaches in the literature, and extends them to an arbitrary order of integration.

Keywords: Cointegration, Common Trends, Triangular representation, Local Smith form, Moving Average representation, Autoregressive representation.

1. INTRODUCTION

The inversion of Moving Average (MA) forms into Auto Regressive (AR) forms (and viceversa) plays a central role in the representation theory of linear processes, both in the stationary case, see e.g. Brockwell and Davis (1991), and in the nonstationary case; see Johansen (1996) for the classes of integrated processes of order 1, $I(1)$, and 2, $I(2)$. This inversion is relevant both for vector ARMA processes, as well as for any (difference-) stationary process satisfying the Wold representation theorem, and hence possessing an MA representation.

The first result of the kind is the celebrated Granger Representation Theorem, see Granger (1981) and Engle and Granger (1987). Starting from the definition of an $I(1)$ process in terms of its MA representation, $\Delta X_t = F(L)\varepsilon_t$, Engle and Granger (1987) considered the inversion of $F(z)$ when $F(1)$ is singular, in order to derive its infinite lag Error Correction form. In this way, Granger's Representation Theorem linked the Common Trends representation for X_t to the cointegrating relations and to the adjustment towards equilibrium.

The proof of the duality (complementarity) between the (number of) common trends and the (number of) cointegrating relations lies at the heart of the interpretation of cointegrating

Massimo Franchi. Sapienza University of Rome, P.le A. Moro 5, 00185 Rome, Italy; tel. +39 06 49910784; e-mail: Email: massimo.franchi@uniroma1.it

Corresponding author: Paolo Paruolo. European Commission, Joint Research Centre (JRC), Via E.Fermi 2749, I-3027 Ispra (VA), Italy; Phone: +39 0332785642; Fax : +39 0332785733. Email: paolo.paruolo@ec.europa.eu

[1]JEL classification: C32.

relations as deviations from equilibria to which variables adjust to, and of common trends as drivers of the system. Moreover, Granger's Representation Theorem clarified that the error correction form and cointegration were not competing concepts, but rather different representations of the same system, see Hendry (2004).

Starting from the MA form of an $I(1)$ system, Phillips (1991) introduced the Triangular Representation; this was subsequently generalised to $I(d)$ systems, $\Delta^d X_t = F(L)\varepsilon_t$, by Stock and Watson (1993). The Triangular Representation summarizes the cointegration properties of the system; it does so by providing the MA representation for a set of (polynomial) linear combinations of the variables, whose number equals the dimension of the system. This set of (polynomial) linear combinations contains the cointegrating relations in the system plus some complementary linear combination of the differences of order d.

The Triangular Representation formed the basis of a semi-parametric inference approach on cointegration, in which the cointegrating relations are estimated parametrically, while the MA form – representing a stationary coloured process – is estimated non-parametrically; see Phillips and Hansen (1990), Sims et al. (1990), Stock and Watson (1993).

An alternative derivation of Granger's Representation Theorem was presented in Yoo (1986) and Engle and Yoo (1991), which made use of the Smith form of the matrix function $F(z)$ in the MA representation $\Delta X_t = F(L)\varepsilon_t$. The approach based on the Smith form was further extended to the case of $I(2)$ systems in Engle and Yoo (1991) and Haldrup and Salmon (1998), who described the (polynomial) cointegrating relations of the system exploiting the existence results of the Smith form to invert the MA form for the first or the second differences of the process.

In the state space framework, Bauer and Wagner (2012) provided a canonical representation of processes with unit roots with integer integration orders at arbitrary frequencies. In this approach, the order of integration is established by the maximal size of the Jordan blocks of the state matrix corresponding to the eigenvalue of unit modulus, and the cointegration properties are expressed via orthogonality conditions.

In a parallel strand of literature, the so-called cointegrated VAR literature, Johansen (1988a,b), see also Johansen (1991), considered the dual problem of inverting the AR representation $F(L)X_t = \varepsilon_t$ and derived conditions under which Granger's Representation Theorem holds for VAR processes. These conditions consist of a reduced rank restriction on $F(1)$ and a full rank condition that involves the first derivative of $F(z)$ at $z = 1$.[1]

The reduced rank condition corresponds to the existence of a pole of some order $m > 0$ in $F(z)^{-1}$ at $z = 1$, while the full rank condition establishes that the order of the pole m is exactly equal to one. This is here called the POLE(1) condition. Under the POLE(1) condition, X_t is $I(1)$ and Johansen (1988a,b) derived the Common Trends representation of a VAR.

He obtained in particular the explicit expression of the matrix that loads the random walk component in the Common Trends representation, G_0 say, the so-called MA impact matrix. Johansen (1994) used it to derive hypotheses on the constant and on deterministic

[1] The same condition can be found in the engineering literature, see Howlett (1982).

terms and this led to the 'star' class of models, see Chapter 4 in Johansen (1996), and Hansen (2005). The explicit form of G_0 was crucial in proving the mixed normality of the asymptotic distribution of the estimator of the cointegrating vectors in Johansen (1991).

The explicit form of the MA impact matrix G_0 was also exploited to derive maximum likelihood estimation and inference on it, see Paruolo (1997a) and Phillips (1998). Counterfactual thought experiments on the long-run behaviour of cointegrated systems also lead to long-run impact multipliers that are functions of MA impact matrix G_0, see Johansen (2005), and Omtzigt and Paruolo (2005) derived maximum likelihood estimation and inference on related long-run multipliers in cointegrated systems. The MA impact matrix plays also a central role in the estimation of the long-run variance matrix, see Phillips (1998), Paruolo (1997b) and Müller (2007) for the bivariate case.

Still starting from the AR form, another derivation of Granger's Representation Theorem was given by Archontakis (1998) employing the Jordan decomposition of the AR companion matrix and using the results by D'Autume (1992) who linked the order of integration to the size of the largest Jordan block of the AR companion matrix. Archontakis (1998) showed that the POLE(1) condition can be stated as the absence of a Jordan block of size greater than 1.

A generalization of Granger's Representation Theorem to $I(2)$ AR processes $F(L)X_t = \varepsilon_t$ was given in Johansen (1992), who stated the POLE(2) condition, under which X_t is $I(2)$, and derived the Common Trends representation. The POLE(2) condition consists of two reduced rank restrictions and one full rank condition on $F(z)$ at $z = 1$: the reduced rank conditions correspond to the existence of a pole in $F(z)^{-1}$ at $z = 1$ of some order $m > 1$, while the full rank condition establishes that the order of the pole m is exactly equal to two.

He derived the explicit form of the matrices G_0 and G_1 of the inverse that load the cumulated random walk and the random walk components in the $I(2)$ Common Trends representation and the explicit expressions for G_0 and G_1 clarified in which directions the process X_t is $I(d)$, for $d = 0, 1, 2$. The explicit expression of G_0 was instrumental in Paruolo (2002) to derive inference on it via likelihood methods; Omtzigt and Paruolo (2005) showed that G_0 enters the long-run impact multipliers, and discussed inference on them.

In the AR framework the case of generic $I(d)$ processes was considered by several authors. The conditions of D'Autume (1992) on the maximal dimension of a Jordan block apply to this general case. la Cour (1998) extended recursively the algebraic necessary and sufficient conditions of Johansen (1992) on the AR coefficients to the case of AR process integrated of any order d, and she described the associated cointegration properties of the system, see also Franchi (2010).

In the engineering literature, the inversion of a matrix function around a point of singularity is a well studied problem, see among others Avrachenkov et al. (2001) and Howlett et al. (2009), who used the approach in Howlett (1982) recursively. In the mathematical literature, a classical approach to characterize the relation between a matrix function and its inverse is via the local spectral theory, based on the concepts of root functions, Jordan chains and local Smith form, see Gohberg et al. (1993).

Franchi and Paruolo (2011, 2016) introduced a procedure called 'extended local rank factorization' (ELRF) which characterizes the structure of Jordan pairs, Jordan chains and the local Smith form and showed that the ELRF coincides with the 'complete reduction processes' in Avrachenkov et al. (2001) thus linking the two approaches.

While the two strands of literature starting from the MA or AR forms are apparently different, the results in the present paper offer a unified treatment of the different representations of cointegrated systems. In particular, the order m of the pole of $F(z)^{-1}$ at $z = 1$ is shown to play a central role in the analysis. When starting from the MA form $\Delta^d X_t = F(L)\varepsilon_t$, m characterizes the cointegration properties of X_t which satisfies a generalization of the triangular representation in Stock and Watson (1993).[2] On the other hand, when starting from the AR form $F(L)X_t = \varepsilon_t$, the order of the pole of the inverse m provides the order of integration of the process.

Building on Franchi and Paruolo (2011, 2016), the present paper provides a general inversion theorem that builds an explicit link between approaches. In addition to existence theorems, the present results provide a constructive approach that allows to compute each representation in terms of the alternative ones. These results extend (when appropriate) results in the literature to any order of integration. They are also shown to apply to any fractional integration order, both for ARFIMA processes and for the class of processes introduced by Johansen (2008) and further studied in Franchi (2010), Johansen and Nielsen (2010, 2012). Moreover, they apply to any stationary, unit or explosive root; this covers also the case of seasonal cointegration, see Hylleberg et al. (1990) and Johansen and Schaumburg (1998).

The rest of the paper is organized as follows: the remaining part of this introduction reports notational conventions and preliminaries; Section 2 contains the general inversion theorem; Section 3 presents a characterization of common trends, cointegration and Triangular Representation of MA and AR processes based on the inversion results in Section 2. Section 4 concludes and Appendix A contains proofs.

Notation and preliminaries

The techniques presented in the paper make repeated use of rank factorizations and projections, whose notation is introduced here. Given a $p \times p$ matrix φ of rank $0 < r < p$, its rank factorization is written as $\varphi = -\alpha\beta'$, where α and β are $p \times r$ full column rank matrices that respectively span the column space and the row space of φ; the negative sign is chosen for convenience in the calculations. The matrix $\varphi_\perp$ indicates a $p \times p - r$ full column rank matrix that spans the orthogonal complement of the column space of φ or α.

The orthogonal projection matrix on the column space of φ is indicated by $P_\alpha := \bar{\alpha}\alpha' = \alpha\bar{\alpha}'$, where $\bar{\alpha} := \alpha(\alpha'\alpha)^{-1}$, and has rank r; $P_{\alpha_\perp} := I - P_\alpha = \bar{\alpha}_\perp\alpha_\perp' = \alpha_\perp\bar{\alpha}_\perp'$ of rank

[2] A non-zero difference $d - m$ regulates if the relevant starting point for the analysis is given by some difference of the process or some cumulation of it.

$p-r$ is the orthogonal projection matrix on the orthogonal complement of the column space of φ. Similarly for φ', one can define P_β and $P_{\beta_\perp}$. When $r=0$, i.e. $\varphi=0$, one sets $\alpha=\beta=\bar{\alpha}=\bar{\beta}=0$ and $\alpha_\perp=\beta_\perp=\bar{\alpha}_\perp=\bar{\beta}_\perp=I$. When $r=p$, i.e. φ of full rank, set either α or β equal to I and $\alpha_\perp=\beta_\perp=\bar{\alpha}_\perp=\bar{\beta}_\perp=0$.

The present paper provides a general inversion theorem of a regular analytic matrix function $F(z)=\sum_{n=0}^\infty F_n(z-z_0)^n$, where F_n are $p\times p$ matrices, z is a complex variable and z_0 is the centre point of the series representation. When $F(z_0)$ is singular, $F(z)^{-1}$ has a pole of some order $m>0$ at z_0. The order of the pole m and the features of the inverse are found to characterize the order of integration and the cointegration properties of the associated process X_t, both in the MA and AR cases. The inversion results are applied to the MA case $\Delta^d X_t = F(L)\varepsilon_t$ and to the AR case $F(L)X_t=\varepsilon_t$, where $\Delta := 1-L$ and L are the difference and the lag operators and ε_t is a white noise sequence.

In the invertible MA or causal AR cases, the point of interest for the expansion is $z_0=0$ and $F(z_0)=F_0=I$ is nonsingular; the inverse $F(z)^{-1} =: G(z)=\sum_{n=0}^\infty G_n z^n$, which solves the system of equations $F(z)G(z)=G(z)F(z)=I$, is found using the recursions[3]

$$G_0=F_0^{-1}, \qquad G_n=\sum_{k=1}^n K_k G_{n-k}, \qquad K_k := -F_0^{-1}F_k, \qquad n=1,2,\ldots. \tag{1.1}$$

In the integrated case, the point of interest for the expansion is $z_0=1$ and this is a point of singularity of $F(z)=\sum_{n=0}^\infty F_n(1-z)^n$, i.e. a point for which $F(z_0)=F_0$ is singular; thus the inversion of $F(z)$ around the singular point $z_0=1$ yields an inverse with a pole of some order $m=1,2,\ldots$. This case is studied in the next section.

2. THE INVERSION THEOREM

This section contains two main results, presented in Theorems 2.1 and 2.3 below; the former provides explicit expressions for the coefficients of the inverse function, while the latter provides a construction of the local Smith factorization of original matrix function. These results are based on Franchi and Paruolo (2011, 2016).

Consider the problem of inversion of a matrix function

$$F(z)=\sum_{n=0}^\infty F_n(1-z)^n, \qquad F_n\in\mathbb{R}^{p\times p}, \qquad F_0\neq 0, \qquad |F_0|=0, \tag{2.1}$$

around a singular point, in this case $z=1$. This includes the case of matrix polynomials $F(z)$, in which the degree of $F(z)$ is finite, k say, and $F_n=0$ for $n>k$.

The inversion of $F(z)$ around the singular point $z=1$ yields an inverse with a pole of some order $m=1,2,\ldots$; an explicit condition on the coefficients $\{F_n\}_{n=0}^\infty$ for $F(z)^{-1}$ to have a pole of given order m is described in Theorem 2.1 below; this is indicated as the POLE(m) condition in the following. Under the POLE(m) condition, $F(z)^{-1}$ has Laurent expansion

[3] See for instance Johansen (1996) Theorem 2.1.

around $z = 1$ given by

$$F(z)^{-1} =: (1-z)^{-m} G(z) = \sum_{n=0}^{\infty} G_n (1-z)^{n-m}, \qquad G_0 \neq 0, \qquad |G_0| = 0. \tag{2.2}$$

Note that $G(1) \neq 0$ is finite by construction and $G(z)$ is expanded around $z = 1$. In the following, the coefficients $\{G_n\}_{n=0}^{\infty}$ are called the Laurent coefficients. The first m of them, $\{G_n\}_{n=0}^{m-1}$, make up the principal part and characterize the singularity of $F(z)^{-1}$ at $z = 1$.

A recursive formula for the Laurent coefficients is provided in (2.3) below. This generalizes (1.1) to the singular case.

Theorem 2.1 (pole(m) condition and Laurent coefficients) *Let α_0, β_0 be full column rank matrices with $0 < r_0 < p$ columns, defined by the rank factorization $F_0 = -\alpha_0 \beta_0'$.*[4] *A necessary and sufficient condition for $F(z)$ to have an inverse with pole of order $m = 1, 2, \ldots$ at $z = 1$ – called* pole(m) *condition – is that*

$$\begin{cases} r_j < r_j^{\max} & (\textit{reduced rank condition}) \textit{ for } j = 1, \ldots, m-1 \\ r_m = r_m^{\max} & (\textit{full rank condition}) \textit{ for } j = m \end{cases},$$

where $r_j^{\max} := p - \sum_{i=0}^{j-1} r_i$, r_j is the rank of α_j and β_j in the rank factorization

$$P_{a_{j\perp}} F_{j,1} P_{b_{j\perp}} = -\alpha_j \beta_j', \qquad a_j := (\alpha_0, \ldots, \alpha_{j-1}), \qquad b_j := (\beta_0, \ldots, \beta_{j-1}),$$

a_j and b_j cointain α_i and β_i only if $r_i > 0$, and

$$F_{j+1,n} := \begin{cases} F_n & \textit{for } j = 0 \\ F_{j,n+1} + F_{j,1} \sum_{i=0}^{j-1} \bar{\beta}_i \bar{\alpha}_i' F_{i+1,n} & \textit{for } j = 1, \ldots, m \end{cases}, \qquad n = 0, 1, \ldots.$$

Moreover, the Laurent coefficients $\{G_n\}_{n=0}^{\infty}$ satisfy

$$G_n = \begin{cases} -\bar{\beta}_m \bar{\alpha}_m' & \textit{for } n = 0 \\ H_n + \sum_{k=1}^{n} K_k G_{n-k} & \textit{for } n = 1, \ldots, m \\ \sum_{k=1}^{n} K_k G_{n-k} & \textit{for } n = m+1, m+2, \ldots \end{cases}, \qquad \begin{array}{l} H_n := \sum_{j=0}^{m} \bar{\beta}_j \bar{\alpha}_j' H_{j+1,n} \\ K_k := \sum_{j=0}^{m} \bar{\beta}_j \bar{\alpha}_j' F_{j+1,k} \end{array}, \tag{2.3}$$

where

$$H_{j+1,n} := \begin{cases} -1_{n=m} I & \textit{for } j = 0 \\ H_{j,n+1} + F_{j,1} \sum_{i=0}^{j-1} \bar{\beta}_i \bar{\alpha}_i' H_{i+1,n} & \textit{for } j = 1, \ldots, m \end{cases}, \qquad n = 0, 1, \ldots, \tag{2.4}$$

and $1_{.}$ is the indicator function.

[4] Here the case $r_0 = 0$ is excluded because otherwise one could re-define $F(z)$ factorizing $(1-z)^s$ from (2.1) for some positive s. The case $r = p$ is also excluded because it would imply $F(z_0)$ nonsingular, in which case the inversion for nonsingular $F(z)$ would apply.

Proof. See Lemma 3.1, Theorems 3.4, 3.5 and Corollary 3.6 in Franchi and Paruolo (2016). ■

Remark 2.2 One has $H_{j+1,n} = 0$ for $n > m$ and for $n + j < m$.

The extended local rank factorization (ELRF), see Franchi and Paruolo (2016), is given by the sequence of calculations in Theorem 2.1 and delivers the output

$$(2.5) \qquad m \qquad \text{and} \qquad \{\alpha_j, \beta_j, r_j, F_{j+1,n}, H_{j+1,n}\}_{j=0,\ldots,m,\, n=0,1,\ldots}.$$

Observe that because $\operatorname{rank} P_{a_{j\perp}} F_{j,1} P_{b_{j\perp}} = \operatorname{rank} a'_{j\perp} F_{j,1} b_{j\perp}$, one has $r_j = \operatorname{rank} a'_{j\perp} F_{j,1} b_{j\perp}$; hence $m = 1$ if and only if

$$r_1 = r_1^{\max}, \quad \text{where} \quad r_1 := \operatorname{rank} \alpha'_{0\perp} F_1 \beta_{0\perp} \quad \text{and} \quad r_1^{\max} := p - r_0.$$

This corresponds to the condition in Theorem 3 of Howlett (1982) and to the $I(1)$ condition in Theorem 4.1 in Johansen (1991).

Similarly, one has $m = 2$ if and only if $r_1 < r_1^{\max}$,

$$r_2 = r_2^{\max}, \quad \text{where} \quad r_2 := \operatorname{rank} a'_{2\perp} F_{2,1} b_{2\perp}, \quad F_{2,1} = F_2 + F_1 \bar{\beta}_0 \bar{\alpha}'_0 F_1, \quad \text{and} \quad r_2^{\max} := p - r_0 - r_1,$$

which corresponds to the $I(2)$ condition in Theorem 3 in Johansen (1992).

Theorem 2.1 shows that, in order to have a pole of order m in the inverse, one needs $m+1$ rank conditions on $F(z)$: the first $j = 1, \ldots, m$ are reduced rank conditions $r_j < r_j^{\max}$, that establish that the order of the pole is greater than $j-1$; the last one is the full rank condition $r_m = r_m^{\max}$ that establishes that the order of the pole is exactly m. These requirements make up the POLE(m) condition.

Also note that by construction the rank factorizations in Theorem 2.1 deliver mutually orthogonal components, namely $\alpha'_h \alpha_j = \beta'_h \beta_j = 0$, $h \neq j$. Thus $a := (\alpha_0, \ldots, \alpha_m)$ and $b := (\beta_0, \ldots, \beta_m)$ are bases of $\mathbb{R}^p$. When j is different from 0 or m, r_j can also be equal to 0; in this case, the corresponding α_j and β_j are equal to 0 and they do not appear in a and b. In what follows, every statement concerning α_j or β_j implicitly assumes that they are nonzero, i.e. that $r_j > 0$, because the modifications required in the case $r_j = 0$ are straightforward.

Eq. (2.3) gives a recursive expression of G_n in (2.2) in terms of the output of the ELRF. The additive term H_n in (2.3), which is absent in the nonsingular case, see (1.1), is present only for the first $m+1$ steps and then disappears, see Remark 2.2. After $m+1$ steps, the two formulae are identical, except for the definition of K_k, which involves the inverse of F_0 in the nonsingular case, while in the singular case it involves $\bar{\beta}_j \bar{\alpha}'_j$, which is the Moore-Penrose inverse of $\alpha_j \beta'_j$, see e.g. Theorem 5, p. 48, in Ben-Israel and Greville (2003).

The next results shows that the ELRF in Theorem 2.1 leads to the construction of the local Smith form and extended canonical system of root functions of $F(z)$.

THEOREM 2.3 (Local Smith factorization) *Given the output of the* ELRF *in* (2.5)*, define the* $p \times r_j$ *matrix functions*

(2.6)
$$\phi_j(z)' := -\bar{\alpha}_j' + \sum_{k=1}^{j} \delta_{j,k}'(1-z)^k, \quad \gamma_j(z)' := \beta_j' - \sum_{k=1}^{\infty} \xi_{j,k}'(1-z)^k, \quad \begin{array}{l} \delta_{j,k}' := \bar{\alpha}_j' H_{j+1,m-j+k} \\ \xi_{j,k}' := \bar{\alpha}_j' F_{j+1,k} \end{array},$$

and the $p \times p$ *matrix functions*

(2.7)
$$\Phi(z) := \begin{pmatrix} \phi_0(z)' \\ \vdots \\ \phi_m(z)' \end{pmatrix}, \quad \Lambda(z) := \begin{pmatrix} (1-z)^0 I_{r_0} & & \\ & \ddots & \\ & & (1-z)^m I_{r_m} \end{pmatrix}, \quad \Gamma(z) := \begin{pmatrix} \gamma_0(z)' \\ \vdots \\ \gamma_m(z)' \end{pmatrix}.$$

Then

$$(2.8) \qquad \Phi(z)F(z) = \Lambda(z)\Gamma(z), \qquad |\Phi(1)| \neq 0, \qquad |\Gamma(1)| \neq 0,$$

i.e. $\Lambda(z)$ *is the local Smith form of* $F(z)$ *at* 1 *and* $\Phi(z), \Gamma(z)$ *are extended canonical systems of root functions.*

Proof. See Appendix A. ■

It is well known, see Gohberg et al. (1993), that every matrix function admits a (unique) diagonal form $\Lambda(z)$ of the type given by (2.8). Theorem 2.3 shows that the ELRF provides a construction of both the canonical form $\Lambda(z)$ and of two (non-unique) extended canonical systems of root functions $\Phi(z), \Gamma(z)$. In particular one has that the values of j with $r_j > 0$ in the ELRF provide the distinct partial multiplicities of $F(z)$ at 1 and r_j is the number of partial multiplicities that are equal to a given j; this characterizes the local Smith form $\Lambda(z)$.

Moreover, $\Phi(z), \Gamma(z)$ are extended canonical system of root functions because they are nonsingular at $z = 1$ and because the j-th block of rows in (2.8) can be written as

$$(2.9) \qquad \phi_j(z)'F(z) = (1-z)^j \gamma_j(z)', \qquad j = 0, 1, \ldots, m,$$

which shows that $\phi_j(z)'$ are r_j left root functions of order j of $F(z)$. Equivalently, using (2.2), one finds

$$(2.10) \qquad \gamma_j(z)'G(z) = (1-z)^{m-j}\phi_j(z)', \qquad j = 0, 1, \ldots, m,$$

which shows that $\gamma_j(z)'$ are r_j left root functions of order $m - j$ of $G(z)$. As shown in Theorems 3.3 and 3.1 below, the concept of cointegrating relation coincides with that of root function and their order of integration is given by the corresponding entry in the local Smith form.

3. COMMON TRENDS, COINTEGRATION AND TRIANGULAR REPRESENTATIONS

This section discusses the application of Theorems 2.1 and 2.3 in the derivation of explicit expressions of the matrix coefficients of the (polynomial) cointegrating relations, of the common trends and triangular representations, either starting from the Moving Average or the Auto Regressive form of a stochastic process.

In particular, Section 3.1 (3.2) considers a generic MA (AR) form and describes its cointegration properties in Theorem 3.1 (3.3) and its Triangular Representation in Corollary 3.2 (3.6). Moreover, Corollaries 3.4 and 3.5 in Section 3.2 present Granger's Representation Theorem and Johansen's Representation Theorem for AR forms as special cases of Theorem 3.3. Finally, Section 3.3 describes the explicit connection between the local Smith form and the Jordan structure in Theorem 3.7 and Section 3.4 considers the case of non-integer d.

The following notation is employed: for a generic process u_t,

$$\begin{aligned} u_t \sim I(d): &\qquad \Delta^d u_t = U(L)\varepsilon_t, \qquad U(1) \neq 0,\\ u_t \sim I_{nc}(d): &\qquad \Delta^d u_t = U(L)\varepsilon_t, \qquad U(1) \text{ has full row rank,} \end{aligned}$$

where $U(z)$ is convergent for all $z \in \mathbb{C} : |z| < 1 + c$, with $c > 0$.

It is well known that cointegration (at frequency 0) is associated with the existence of a nonzero linear combination that is integrated of lower order, which corresponds to the fact that $U(1)$ has not full row rank. Hence $u_t \sim I(d)$ is integrated and it can be cointegrated, while $u_t \sim I_{nc}(d)$ is integrated and it is not cointegrated.

Further note that d in the previous definition can be positive, 0 or negative; in the last case this leads to cumulation of the process, because Δ^{-1} is defined by $\Delta^{-1}u_t := \sum_{i=1}^t u_j + u_0$, where $t > 0$ and u_0 is the initial value. As a last piece of notation, $a^{(q)}(z) := \sum_{n=0}^q a_n(1-z)^n$ denotes the truncation of order q of for a generic function $a(z) := \sum_{n=0}^\infty a_n(1-z)^n$, i.e. $a(z) - a^{(q)}(z) =: (1-z)^{q+1}a^\star(z)$, where $a^\star(z) = \sum_{n=0}^\infty a_{n+q+1}(1-z)^n$.

3.1. *MA forms*

Consider a generic $I(d)$ process

$$\Delta^d X_t = F(L)\varepsilon_t, \qquad F_0 \neq 0, \qquad |F_0| = 0, \tag{3.1}$$

with characteristic roots of $F(z)$ at $z = 1$ and at $|z| > 1$. Applying Theorems 2.1 and 2.3 to $F(z)$ in (3.1), one obtains the following result.

THEOREM 3.1 (Conitegration properties of MA processes) *The $I(d)$ process X_t in (3.1) admits the following Common Trends representation:*

$$X_t = \sum_{n=0}^{d-1} F_n S_{d-n,t} + F^\star(L)\varepsilon_t + v_0,$$

where $S_{h,t} := \sum_{i=1}^t S_{h-1,i} \sim I(h)$ for $h \geq 1$, $S_{0,t} := \varepsilon_t$, $F^\star(L)\varepsilon_t := \sum_{n=d}^\infty F_n\Delta^{n-d}\varepsilon_t$ is stationary, v_0 collects initial values, and the coefficients F_n are given in (2.1).

Next assume that X_t satisfies the POLE(m) *condition on $F(z)$; in this case the cointegration properties of X_t are fully described by the cointegrating relations*

$$\phi_j^{(j-1)}(L)'X_t \sim I_{nc}(d-j), \qquad j = 1, \ldots, m, \tag{3.2}$$

where $\phi_j^{(j-1)}(z)' = \alpha_j' - \sum_{k=1}^{j-1} \delta_{j,k}'(1-z)^k$ is obtained as the truncation of order $j-1$ of the root functions $\phi_j(z)$ in (2.6) in Theorem 2.3. Finally, defining $\Phi_c(z) := (\bar{\alpha}_0, \phi_1^{(0)}(z), \ldots, \phi_m^{(m-1)}(z))'$, one has

$$\Lambda(L)^{-1}\Phi_c(L)\Delta^d X_t = C(L)\varepsilon_t \sim I_{nc}(0), \qquad |\Phi_c(1)| \neq 0, \tag{3.3}$$

where $\Lambda(z)$ is the Local Smith form of $F(z)$, see (2.7).

Proof. See Appendix A. ■

Note that the cointegrating relations of X_t coincide with the truncated $\Phi(z)$ root functions of $F(z)$ while their order of integration is given by the corresponding entry in the local Smith form $\Lambda(z)$ of $F(z)$.

The previous theorem leads to a Generalized Triangular Representation, as shown in the following corollary.

Corollary 3.2 (Triangular representation of MA processes) *Let X_t in (3.1) satisfy the* POLE(m) *condition on $F(z)$; then it admits the Generalized Triangular Representation*

$$\begin{pmatrix} \bar{\alpha}_0'\Delta^d X_t \\ \alpha_1'\Delta^{d-1}X_t \\ \alpha_2'\Delta^{d-2}X_t - \delta_{2,1}'\Delta^{d-1}X_t \\ \\ \alpha_m'\Delta^{d-m}X_t - \sum_{k=1}^{m-1}\delta_{m,k}'\Delta^{d-m+k}X_t \end{pmatrix} = C(L)\varepsilon_t \sim I_{nc}(0),$$

which reduces to the Triangular Representation in eq. (3.2) *of Stock and Watson (1993) for $m = d$.*

Proof. Use (3.3) in Theorem 3.1. ■

Observe that the order of integration d of X_t is not affected by the structure of $F(z)$, and hence by the order m of the pole of $F(z)^{-1}$. However, the cointegration properties of X_t do not depend on d but on m, which is dictated by the structure of the F coefficients.

For example, an $I(d)$ process with $m = 1$ admits Common Trends representation

$$X_t = \sum_{n=0}^{d-1} F_n S_{d-n,t} + F^\star(L)\varepsilon_t + v_0$$

and Generalized Triangular Representation

$$\begin{pmatrix} \bar{\alpha}_0'\Delta X_t \\ \alpha_1' X_t \end{pmatrix} \sim I_{nc}(d-1).$$

In this case, cointegration occurs only in the direction of α_1; no further decrements are possible and this fully describes its its cointegration properties.

On the other hand, an $I(1)$ process with generic m admits Common Trends representation

$$X_t = -\alpha_0 \beta_0' \sum_{i=1}^{t} \varepsilon_i + F^\star(L)\varepsilon_t + v_0$$

and Generalized Triangular Representation

$$\begin{pmatrix} \bar{\alpha}_0' \Delta X_t \\ \alpha_1' X_t \\ \alpha_2' \Delta^{-1} X_t - \delta_{2,1}' \Delta^{-1} X_t \\ \\ \alpha_m' \Delta^{1-m} X_t - \sum_{k=1}^{m-1} \delta_{m,k}' \Delta^{1-m+k} X_t \end{pmatrix} = C(L)\varepsilon_t \sim I_{nc}(0).$$

In this case, cointegrated relations occurs in the direction of α_j, $j = 1, \ldots, m$ and they involve cumulation of X_t if $m > 1$.

In general, the Generalized Triangular Representation in Corollary 3.2 shows that the cointegrating relations involve $\Delta^j X_t$ for $j = d - m, \ldots, d - 1$, and some of these powers may be negative. In this case $\Delta^j X_t$ corresponds to cumulations of X_t. While m does not influence the order of integration of X_t, it does impact the number of differences or cumulations of X_t that enter the cointegration structure of the system and determines its triangular representation.

3.2. *AR forms*

Next consider a generic AR process $A(L)X_t = \varepsilon_t$, with characteristic roots of $A(z)$ at $z = 1$ and at $|z| > 1$. Since $|A(1)| = 0$, the first s coefficients in the expansion of $A(z)$ around 1, $A(z) := \sum_{n=0}^{\infty} A_n(1-z)^n$, could be equal to 0 for some $s > 0$, leading to a factorization of the type $A(z) = (1-z)^s F(z)$. A generic AR process can then be written as

$$F(L)\Delta^s X_t = \varepsilon_t, \qquad F_0 = A_s \neq 0, \qquad |F_0| = 0. \tag{3.4}$$

One can then apply Theorems 2.1 and 2.3 to $F(z)$ in (3.4), obtaining the following result.

Theorem 3.3 (Cointegration properties of AR processes) *The process X_t in (3.4) is $I(d)$ with $d = m+s$ if and only if the* POLE(m) *condition applies to $F(z)$. In this case the following Common Trends representation holds:*

$$X_t = \sum_{n=0}^{d-1} G_n S_{d-n,t} + G^\star(L)\varepsilon_t + v_0,$$

where $S_{h,t} := \sum_{i=1}^{t} S_{h-1,i} \sim I(h)$ for $h \geq 1$, $S_{0,t} := \varepsilon_t$, $G^\star(L)\varepsilon_t := \sum_{n=d}^{\infty} G_n \Delta^{n-d}\varepsilon_t$ is stationary, v_0 collects initial values, and the Laurent coefficients G_n are given in (2.3) in

Theorem 2.1. If $d = m$, i.e. $s = 0$, $G^\star(L)\varepsilon_t$ is $I(0)$. The cointegration properties of X_t are fully described by the cointegrating relations

$$\gamma_j^{(m-j-1)}(L)'X_t \sim I_{nc}(d-m+j), \qquad j = 0, 1, \ldots, m-1, \tag{3.5}$$

where $\gamma_j^{(m-j-1)}(z) = \beta_j' - \sum_{k=1}^{m-j-1} \xi_{j,k}'(1-z)^k$ is obtained as the truncation of order $m-j-1$ of the root functions $\gamma_j(z)$ in (2.6) in Theorem 2.3. Finally, defining $\Gamma_c(z) := (\gamma_0^{(m-1)}(z), \ldots, \gamma_{m-1}^{(0)}(z), \beta_m)'$, one has

$$\Lambda(L)\Gamma_c(L)\Delta^s X_t = C(L)\varepsilon_t \sim I_{nc}(0), \qquad |\Gamma_c(1)| \neq 0, \tag{3.6}$$

where $\Lambda(z)$ is the Local Smith form of $F(z)$, see (2.7).

Proof. See Appendix A. ■

Note that the cointegrating relations of X_t coincide with the truncated $\Gamma(z)$ root functions of $G(z)$ while their order of integration is given by the corresponding entry in the local Smith form $\Lambda(z)$ of $F(z)$.

Setting $s = 0$ and $m = 1$ in Theorem 3.3 one finds Theorem 4.2 in Johansen (1996), i.e. the following result.

Corollary 3.4 (Cointegration properties of $I(1)$ AR processes) *The process X_t in (3.4) is $I(1)$ if and only if the POLE(1) condition applies to $F(z)$. In this case the following Common Trends representation holds:*

$$X_t = -\bar{\beta}_1\bar{\alpha}_1' \sum_{i=1}^{t} \varepsilon_i + G^\star(L)\varepsilon_t + v_0,$$

where $G^\star(L)\varepsilon_t := \sum_{n=1}^{\infty} G_n\Delta^{n-1}\varepsilon_t \sim I(0)$, v_0 is a constant which depends on the initial values of the process, the Laurent coefficients G_n are given in (2.3) in Theorem 2.1, and the cointegration properties of X_t are fully described by

$$\begin{pmatrix} \beta_0' X_t \\ \beta_1'\Delta X_t \end{pmatrix} \sim I_{nc}(0). \tag{3.7}$$

Proof. See Appendix A. ■

Similarly, setting $s = 0$ and $m = 2$ in Theorem 3.3 one finds Theorem 4.6 in Johansen (1996), i.e. the following result.

Corollary 3.5 (Cointegration properties of $I(2)$ AR processes) *The process X_t in (3.4) is $I(2)$ if and only if the POLE(2) condition applies to $F(z)$. In this case the following Common Trends representation holds:*

$$X_t = -\bar{\beta}_2\bar{\alpha}_2' \sum_{i=1}^{t}\sum_{h=1}^{i} \varepsilon_h + G_1 \sum_{i=1}^{t} \varepsilon_i + G^\star(L)\varepsilon_t + v_{0t},$$

where

$$G_1 = -\bar{b} \begin{pmatrix} 0 & 0 & \bar{\alpha}_0' F_{1,1} \bar{\beta}_2 \\ 0 & I_{r_1} & \bar{\alpha}_1' F_{2,1} \bar{\beta}_2 \\ \bar{\alpha}_2' F_{1,1} \bar{\beta}_0 & \bar{\alpha}_2' F_{2,1} \bar{\beta}_1 & \bar{\alpha}_2' F_{3,1} \bar{\beta}_2 \end{pmatrix} \bar{a}', \qquad a = (\alpha_0, \alpha_1, \alpha_2), \qquad b = (\beta_0, \beta_1, \beta_2),$$

$G^\star(L)\varepsilon_t := \sum_{n=2}^{\infty} G_n \Delta^{n-2} \varepsilon_t \sim I(0)$, *$v_{0t}$ is a polynomial in t of degree 1 which depends on the initial values of the process, the Laurent coefficients G_n are given in (2.3) in Theorem 2.1, and the cointegration properties of X_t are fully described by*

$$\begin{pmatrix} \beta_0' X_t - \bar{\alpha}_0' F_1 \Delta X_t \\ \beta_1' \Delta X_t \\ \beta_2' \Delta^2 X_t \end{pmatrix} \sim I_{nc}(0). \tag{3.8}$$

Proof. See Appendix *A*. ■

The previous theorem leads to a Generalized Triangular Representation, as shown in the following corollary.

COROLLARY 3.6 (Triangular representation of AR processes) *Let X_t in (3.4) satisfy the* POLE(m) *condition on $F(z)$; then it admits the Generalized Triangular Representation*

$$\begin{pmatrix} \beta_0' \Delta^s X_t - \sum_{k=1}^{m-1} \xi_{0,k}' \Delta^{s+k} X_t \\ \beta_1' \Delta^{s+1} X_t - \sum_{k=1}^{m-2} \xi_{1,k}' \Delta^{s+k+1} X_t \\ \vdots \\ \beta_{m-1}' \Delta^{s+m-1} X_t \\ \beta_m' \Delta^{s+m} X_t \end{pmatrix} = C(L)\varepsilon_t \sim I_{nc}(0)$$

which reduces to the Triangular Representation in eq. (3.2) *of Stock and Watson (1993) for $s = 0$, i.e. $m = d$.*

Proof. Use (3.6) in Theorem 3.3. ■

Note that special cases of the Triangular Representation for the AR form are given in (3.7) for the $I(1)$ case, and in (3.8) for the $I(2)$ case.

Comparing the cointegration properties of MA and AR processes in Theorems 3.1 and 3.3, one sees that the two extended canonical systems of root functions $\Phi(z)$ and $\Gamma(z)$ play a symmetric role; one of them is used when starting from a MA form, and the other one when starting from a AR form.

3.3. *Jordan forms*

This subsection deals with the connection with the Jordan form approach, in which the order of integration is given by the maximal size of the Jordan blocks corresponding to the eigenvalue at 1.

The following additional notation is needed here; let $\mathcal{J} := (j : r_j > 0)$ be the ordered set that contains the $w+1 := \#\mathcal{J}$ indexes j that correspond to nonzero ranks r_j. Indicate the elements of $\mathcal{J}$ by $(j_1, j_2, \ldots, j_{w+1})$ and fix the reverse ordering $m = j_1 > j_2 > \cdots > j_w > j_{w+1} = 0$. Next let $\mathcal{J}_+$ be the ordered set that contains only the positive elements of $\mathcal{J}$, i.e. $\mathcal{J}_+ := \mathcal{J} \setminus \{0\} = (j_1, j_2, \ldots, j_w)$. Note that the index set $\mathcal{J}_+$ contains at least one element (equal to m), and at most m elements, $\mathcal{J}_+ = (m, m-1, \ldots, 1)$, and hence $1 \leq w \leq m$. Finally let $\mathcal{K}$ be the ordered set that contains each $j \in \mathcal{J}_+$ repeated r_j times and indicate its elements by $(k_1, k_2, \ldots, k_{p-r_0}) := \mathcal{K}$, i.e.

$$\begin{aligned}\mathcal{K} :=&(\underbrace{j_1, \ldots, j_1}_{r_{j_1}\text{times}}, \underbrace{j_2, \ldots, j_2}_{r_{j_2}\text{times}}, \ldots, \underbrace{j_w, \ldots, j_w}_{r_{j_w}\text{times}}) = \\ =&(\underbrace{k_1, \ldots, k_{r_{j_1}}}_{=j_1}, \underbrace{k_{r_{j_1}+1}, \ldots, k_{r_{j_1}+r_{j_2}}}_{=j_2}, \ldots \underbrace{k_{\sum_{i=1}^{w-1} r_{j_i}+1}, \ldots, k_{p-r_0}}_{=j_w}).\end{aligned}$$

Note that the index set $\mathcal{K}$ contains $\sum_{j\in\mathcal{J}_+} r_j = p-r_0$ elements. In the following $\operatorname{diag}(a_j)_{j\in\mathcal{J}_+}$ indicates a block diagonal matrix with $a_{j_1}, \ldots, a_{j_w}$ on the main diagonal.

Given the extended canonical system of root functions $\Phi(z)$ in (2.7) and the index set $\mathcal{K}$, one can construct a Jordan pair of $F(z)$ at $z = 1$ as follows.[5]

THEOREM 3.7 (Jordan pair at $z = 1$) *Let $\phi_{i,n}$ be the i-th column of Φ_n in the extended canonical system of root functions $\Phi(z) = \sum_{n=0}^{\infty} \Phi_n (1-z)^n$ in (2.7), and let k_i be the i-th element in the index set $\mathcal{K}$; for $i = 1, \ldots, p-r_0$, define*

$$X_i := (\phi_{i,n})_{n=0}^{k_i-1}, \qquad J_{k_i} := \begin{pmatrix} z_0 & 1 & & \\ & \ddots & \ddots & \\ & & z_0 & 1 \\ & & & z_0 \end{pmatrix},$$

respectively of dimension $p \times k_i$ and $k_i \times k_i$. Then the columns of X_i form a Jordan chain of maximal length k_i and J_{k_i} is the corresponding Jordan block. Collecting the Jordan chains and the Jordan blocks respectively in

$$X := (X_i)_{i=1}^{p-r_0}, \qquad J := \operatorname{diag}(I_{r_j} \otimes J_j)_{j\in\mathcal{J}_+},$$

one has that (X, J) is a Jordan pair of $F(z)$ at $z = 1$.

Proof. Direct consequence of Theorem 2.3 and the definition of Jordan pairs in Gohberg et al. (1993). ■

This theorem contains the results in D'Autume (1992), Archontakis (1998), and Bauer and Wagner (2012) as special cases. In fact, take for example the companion matrix of an AR process; the Jordan blocks of this companion matrix corresponding to the eigenvalue at 1 are

[5] A similar result applies to $\Gamma(z)$.

collected in the matrix J in Theorem 3.7; this follows e.g. from Corollary 1.21 in Gohberg et al. (1982). Hence the characterization of the order of integration as the maximal size of the Jordan blocks of the companion matrix corresponding to the eigenvalue at 1 is easily obtained by the ELRF.

3.4. *Non-integer integration orders*

The present results also apply to the cases of non-integer d of the ARFIMA type; this can be seen by choosing $s \in \mathbb{R}$ in (3.4) or $d \in \mathbb{R}$ in (3.1). The analysis applies as well to the class of fractionally integrated processes defined in Johansen (2008, 2009), see eq. (3.1) in Franchi (2010). In fact, the results carry over by defining $L_b := 1 - (1 - L)^b$ and rewriting (3.4) as $F(L_b)\Delta^s X_t = \varepsilon_t$ with $s := d - mb$, $d, b \in \mathbb{R}$, $m \in \mathbb{N}$, $0 < mb \leq d$, and replacing L with L_b.

4. CONCLUSIONS

The general inversion results deliver both the Laurent coefficients of the inverse as well as a construction of the local Smith form and of the root functions, as recursive expressions of the coefficients of the matrix function to be inverted. These results are based on the ELRF, which consists in performing a finite sequence of rank factorizations of matrices that involve the derivatives of the matrix function evaluated at the point around which the inverse is conducted.

The general representation results unify and clarify existing representation results in the literature, and extend them to any integer order. The present results carry over to fractionally integrated processes; moreover they are not specific to the unit root case and can be applied to any (stationary, unit, explosive) root, including the seasonal case.

APPENDIX A: PROOFS

Proof of Theorem 2.3. Equation (3.3) in Franchi and Paruolo (2016) gives

$$\alpha_j \beta_j' G_{h-j} = P_{a_{j\perp}} \sum_{k=1}^{h-j} F_{j+1,k} G_{h-j-k} + P_{a_{j\perp}} H_{j+1,h-j}, \qquad h \geq j = 0, 1, \ldots, m,$$

where

$$\text{(A.1)} \qquad H_{j+1,h-j} = \begin{cases} 0 & \text{for } h < m \\ -I & \text{for } h = m \\ H_{j,h-j+1} + F_{j,1} \sum_{i=0}^{j-1} \bar{\beta}_i \bar{\alpha}_i' H_{i+1,h-j} & \text{for } h > m \end{cases}$$

follows by applying definition (2.4). Pre-multiplying by $\bar{\alpha}_j'$ and rearraging one thus finds

$$\text{(A.2)} \qquad \beta_j' G_{h-j} - \bar{\alpha}_j' \sum_{k=1}^{h-j} F_{j+1,k} G_{h-j-k} = \bar{\alpha}_j' H_{j+1,h-j}, \qquad h \geq j = 0, 1, \ldots, m.$$

Next consider $\gamma_j(z)' =: \sum_{n=0}^{\infty} \gamma_{j,n}' (1-z)^n$ in (2.6), where $\gamma_{j,0}' = \beta_j'$ and $\gamma_{j,n}' = -\bar{\alpha}_j' F_{j+1,n}$ for $n \geq 1$, and $G(z) = \sum_{n=0}^{\infty} G_n (1-z)^n$ in (2.2). Writing $\gamma_j(z)' G(z) = \sum_{n=0}^{\infty} \zeta_{j,n}' (1-z)^n$, where $\zeta_{j,n}' := \sum_{k=0}^{n} \gamma_{j,k}' G_{n-k}$ is found by convolution, one has

$$\zeta_{j,n}' = \beta_j' G_n - \bar{\alpha}_j' \sum_{k=1}^{n} F_{j+1,k} G_{n-k} = \bar{\alpha}_j' H_{j+1,n}, \qquad n \geq 0, \qquad j = 0, 1, \ldots, m,$$

where the last equality follows by setting $n = h - j$ in (A.2). Setting $n = h - j$ in (A.1) one finds

$$H_{j+1,n} = \begin{cases} 0 & \text{for } n < m-j \\ -I & \text{for } n = m-j \\ H_{j,n+1} + F_{j,1}\sum_{i=0}^{j-1} \bar{\beta}_i \bar{\alpha}_i' H_{i+1,n} & \text{for } n > m-j \end{cases}$$

and hence one has

$$\zeta_{j,n}' = \begin{cases} 0 & \text{for } n < m-j \\ -\bar{\alpha}_j' & \text{for } n = m-j \\ \bar{\alpha}_j' H_{j+1,n} & \text{for } n > m-j \end{cases}.$$

This shows that

$$\gamma_j(z)' G(z) = (1-z)^{m-j}\phi_j(z)', \qquad \phi_j(z)' := -\alpha_j' + \bar{\alpha}_j' \sum_{k=1}^{\infty} H_{j+1,m-j+k}(1-z)^k,$$

i.e. $\gamma_j(z)'F(z)^{-1} = (1-z)^{-j}\phi_j(z)'$ or $\phi_j(z)'F(z) = (1-z)^j\gamma_j(z)'$, for $j = 0, 1, \ldots, m$. Because $H_{j+1,n} = 0$ for $n > m$, see Remark 2.2 one has $H_{j+1,m-j+k} = 0$ for $k > j$; hence $\phi_j(z)$ is a matrix polynomial of degree j. ∎

Proof of Theorem 3.1. Pre-multiplying $\Delta^d X_t = F(L)\varepsilon_t$ by $\phi_j(L)'$ and using eq. (2.9) one obtains $\phi_j(L)'\Delta^{d-j}X_t = \gamma_j(L)'\varepsilon_t$ with $\gamma_j(1)' = \beta_j'$ of full row rank. This shows that $\phi_j(L)'X_t \sim I_{nc}(d-j)$. Substituting $\phi_j(z)' = \phi_j^{(j-1)}(z)' + (1-z)^j\phi_j^\star(z)'$ one finds $\phi_j^{(j-1)}(L)'\Delta^{d-j}X_t = (\gamma_j(L)' - \phi_j^\star(L)'F(L))\varepsilon_t$. Using (2.6) and (2.1), one has $\gamma_j(1)' - \phi_j^\star(1)'F(1) = \beta_j' - \bar{\alpha}_j' H_{j+1,m}\alpha_0\beta_0'$ which has full row rank. This shows that also $\phi_j^{(j-1)}(L)'X_t \sim I_{nc}(d-j)$. Grouping (3.2) together and pre-multiplying by $\Lambda(L)^{-1}$ defined in (2.7), one finds (3.3), where $\Phi_c(1) = -a'$ is square and nonsingular. Moreover, $C(1) = \Gamma(1)$ so that $C(1)\bar{b}$ is lower triangular with identities on the main diagonal and this shows that $C(1)$ is square and nonsingular. ∎

Proof of Theorem 3.3. By Theorem 2.1, $F(z)^{-1} = G(z)(1-z)^{-m}$ with $G(1) \neq 0$ if and only if the POLE(m) condition on $F(z)$ hold, i.e. one has $\Delta^{m+s}X_t = G(L)\varepsilon_t$ with $G(1) \neq 0$, which shows that $X_t \sim I(d)$, $d = m + s$. Pre-multiplying $\Delta^d X_t = G(L)\varepsilon_t$ by $\gamma_j(L)'$ and using eq. (2.10) one obtains $\gamma_j(L)'\Delta^{s+j}X_t = \phi_j(L)'\varepsilon_t$ with $\phi_j(1)' = -\alpha_j'$ of full row rank. This shows that $\gamma_j(L)'X_t \sim I_{nc}(s+j)$. Substituting $\gamma_j(z)' = \gamma_j^{(m-j-1)}(z)' + (1-z)^{m-j}\gamma_j^\star(z)'$ one finds $\gamma_j^{(m-j-1)}(L)'\Delta^{s+j}X_t = (\phi_j(L)' - \gamma_j^\star(L)'G(L))\varepsilon_t$. Using (2.6) and (2.3), one has $\phi_j(1)' - \gamma_j^\star(1)'G(1) = -\alpha_j' - \bar{\alpha}_j' F_{j+1,m-j}\bar{\beta}_m\bar{\alpha}_m'$ which has full row rank. This shows that also $\gamma_j^{(m-j-1)}(L)'X_t \sim I_{nc}(s+j)$. Grouping (3.5) together and pre-multiplying by $\Lambda(L)$ defined in (2.7), one finds (3.6), where $\Gamma_c(1) = b'$ is square and nonsingular. Moreover, $C(1) = \Phi(1)$ so that $-C(1)a$ is upper triangular with identities on the main diagonal and this shows that $C(1)$ is square and nonsingular. ∎

Proof of Corollary 3.4. Setting $m = 1$ in Theorem 2.1 one has $G_0 = -\bar{\beta}_1\bar{\alpha}_1'$ and setting $s = 0$ and $m = 1$ in Theorem 3.3 one has

$$\Lambda(z) = \begin{pmatrix} I_{r_0} & 0 \\ 0 & (1-z)I_{r_1} \end{pmatrix}, \qquad \Gamma_c(z) = \begin{pmatrix} \beta_0' \\ \beta_1' \end{pmatrix}$$

and this completes the proof. ∎

Proof of Corollary 3.5. Setting $m = 2$ in Theorem 2.1 one has $G_0 = -\bar{\beta}_2\bar{\alpha}_2'$ and $G_1 = H_1 + K_1 G_0$, where

$$H_1 = \sum_{j=0}^{2} \bar{\beta}_j\bar{\alpha}_j' H_{j+1,1} = \bar{b}\begin{pmatrix} \bar{\alpha}_0' H_{1,1} \\ \bar{\alpha}_1' H_{2,1} \\ \bar{\alpha}_2' H_{3,1} \end{pmatrix}, \qquad K_1 = \sum_{j=0}^{2} \bar{\beta}_j\bar{\alpha}_j' F_{j+1,1} = \bar{b}\begin{pmatrix} \bar{\alpha}_0' F_{1,1} \\ \bar{\alpha}_1' F_{2,1} \\ \bar{\alpha}_2' F_{3,1} \end{pmatrix}.$$

Definition (2.4) implies $H_{2,1} = -I$, $H_{j+1,n} = 0$ for $j + 1 + n \leq m$ and $H_{1,n} = 0$ for $n > m$; hence one has $H_{1,1} = 0$, $H_{3,1} = -F_{1,1}\bar{\beta}_0\bar{\alpha}_0' - F_{2,1}\bar{\beta}_1\bar{\alpha}_1'$ and thus one finds

$$G_1 = -\bar{b}\begin{pmatrix} 0 & 0 & \bar{\alpha}_0' F_{1,1}\bar{\beta}_2 \\ 0 & I_{r_1} & \bar{\alpha}_1' F_{2,1}\bar{\beta}_2 \\ \bar{\alpha}_2' F_{1,1}\bar{\beta}_0 & \bar{\alpha}_2' F_{2,1}\bar{\beta}_1 & \bar{\alpha}_2' F_{3,1}\bar{\beta}_2 \end{pmatrix}\bar{a}'.$$

Setting $s = 0$ and $m = 2$ in Theorem 3.3 one has

$$\Lambda(z) = \begin{pmatrix} I_{r_0} & 0 & 0 \\ 0 & (1-z)I_{r_1} & 0 \\ 0 & 0 & (1-z)^2 I_{r_2} \end{pmatrix}, \qquad \Gamma_c(z) = \begin{pmatrix} \beta_0' - \bar{\alpha}_0' F_{1,1}(1-z) \\ \beta_1' \\ \beta_2' \end{pmatrix}$$

and this completes the proof. ■

REFERENCES

Archontakis, F. (1998). An alternative proof of Granger's representation theorem for I(1) systems through Jordan matrices. *Statistical Methods and Applications 7*, 111–127.

Avrachenkov, K., M. Haviv, and P. Howlett (2001). Inversion of analytic matrix functions that are singular at the origin. *SIAM Journal of Matrix Analysis and Applications 22*, 1175–89.

Bauer, D. and M. Wagner (2012). A State Space Canonical Form for Unit Root Processes. *Econometric Theory 28*, 1313–1349.

Ben-Israel, A. and T. Greville (2003). *Generalized Inverses: Theory and Applications, 2nd ed.* Springer.

Brockwell, P. and R. Davis (1991). *Time Series: Theory and Methods, second edition.* Springer-Verlag.

D'Autume, A. (1992). Deterministic dynamics and cointegration of higher orders. Working Papers 1992-43, Cashiers de reserche Econome Mathematiques et applications.

Engle, R. and C. Granger (1987). Co-integration and Error Correction: Representation, Estimation, and Testing. *Econometrica 55*, 251–276.

Engle, R. and S. Yoo (1991). Cointegrated economic time series: an overview with new results. *in Long-run Economic Relations: Readings in Cointegration Engle and Granger eds., Oxford University Press*, 237–266.

Franchi, M. (2010). A representation theory for polynomial cofractionality in vector autoregressive models. *Econometric Theory 26*, 1201–1217.

Franchi, M. and P. Paruolo (2011). Inversion of regular analytic matrix functions: local Smith form and subspace duality. *Linear Algebra and its Applications 435*, 2896–2912.

Franchi, M. and P. Paruolo (2016). Inverting a matrix function around a singularity via local rank factorization. *SIAM Journal of Matrix Analysis and Applications 37*, 774–797.

Gohberg, I., M. Kaashoek, and F. van Schagen (1993). On the local theory of regular analytic matrix functions. *Linear Algebra and its Applications 182*, 9–25.

Gohberg, I., P. Lancaster, and L. Rodman (1982). *Matrix polynomials.* Academic Press, New York.

Granger, C. (1981). Some Properties of Time Series Data and their use in Econometric Model Specification. *Journal of Econometrics 16*, 121–130.

Haldrup, N. and M. Salmon (1998). Representation of I(2) cointegrated systems using the Smith-McMillan form. *Journal of Econometrics 44*, 303–325.

Hansen, P. (2005). Granger's representation theorem: a closed-form expression for I(1) processes. *Econometrics Journal 8*, 23–38.

Hendry, D. F. (2004). The Nobel Memorial Prize for Clive W. J. Granger. *Scandinavian Journal of Economics 106*(2), 187–213.

Howlett, P. (1982). Input retrieval in finite dimensional linear systems. *Journal of Australian Mathematical Society (Series B) 23*, 357–82.

Howlett, P., K. Avrachenkov, C. Pearce, and V. Ejov (2009). Inversion of analytically perturbed linear operators that are singular at the origin. *Journal of Mathematical Analysis and Applications 353*, 68–84.

Hylleberg, S., R. F. Engle, C. W. J. Granger, and B. S. Yoo (1990). Seasonal integration and cointegration. *Journal of Econometrics 44*(1-2), 215–238.

Johansen, S. (1988a). The mathematical structure of error correction models. *Contemporary Mathematics 80*, 359–386.

Johansen, S. (1988b). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control 12*, 231–254.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica 59*, 1551–1580.

Johansen, S. (1992). A representation of vector autoregressive processes integrated of order 2. *Econometric Theory 8*, 188–202.

Johansen, S. (1994). The role of the constant and linear terms in cointegration analysis of nonstationary variables. *Econometric Reviews 13*(2), 205–229.

Johansen, S. (1996). *Likelihood-based Inference in Cointegrated Vector Auto-Regressive Models*. Oxford University Press.

Johansen, S. (2005). Interpretation of cointegrating coefficients in the cointegrated vector autoregressive model. *Oxford Bulletin of Economics and Statistics 67*, 93–104.

Johansen, S. (2008). A representation theory for a class of vector autoregressive models for fractional processes. *Econometric Theory 24*, 651–676.

Johansen, S. (2009). Representation of cointegrated autoregressive process with application to fractional process. *Econometric Reviews 28*, 121–145.

Johansen, S. and M. Nielsen (2010). Likelihood inference for a nonstationary fractional autoregressive model. *Journal of Econometrics 158*, 51–66.

Johansen, S. and M. Nielsen (2012). Likelihood inference for a fractionally cointegrated vector autoregressive model. *Econometrica 80*, 2667 – 2732.

Johansen, S. and E. Schaumburg (1998). Likelihood analysis of seasonal cointegration. *Journal of Econometrics 88*, 301–339.

la Cour, L. (1998). A parametric characterization of integrated vector autoregressive (VAR) processes. *Econometric Theory 14*, 187–199.

Müller, U. K. (2007). A theory of robust long-run variance estimation. *Journal of Econometrics 141*, 1331–1352.

Omtzigt, P. and P. Paruolo (2005). Impact factors. *Journal of Econometrics 128*, 31–68.

Paruolo, P. (1997a). Asymptotic inference on the moving average impact matrix in cointegrated I(1) VAR systems. *Econometric Theory 13*(1), 79–118.

Paruolo, P. (1997b). Standard errors for the long run variance matrix. *Econometric Theory 13*, *Problem*: Vol. 13 (1), 305–306, *Solution*: Vol. 14 (1), 152–153.

Paruolo, P. (2002). Asymptotic inference on the moving average impact matrix in cointegrated I(2) VAR systems. *Econometric Theory 18*(3), 673–690.

Phillips, P. (1991). Optimal inference in cointegrated systems. *Econometrica 59*, 283–306.

Phillips, P. and B. Hansen (1990). Statistical inference in Instrumental Variables regression with I(1) processes. *Review of Economic Studies 57*(1), 99–125.

Phillips, P. C. (1998). Impulse response and forecast error variance asymptotics in nonstationary VARs. *Journal of Econometrics 83*(1172), 21 – 56.

Sims, C., J. Stock, and M. Watson (1990). Inference in linear time series models with some unit roots. *Econometrica 58*, 113–144.

Stock, J. and M. Watson (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica 61*, 783–820.

Yoo, B. S. (1986). Multi-cointegrated time series and generalized error correction models. Economics department, working paper, University of California, San Diego.