Econometric Reviews

ISSN: 0747-4938 (Print) 1532-4168 (Online) Journal homepage: https://www.tandfonline.com/loi/lecr20

A general inversion theorem for cointegration

Massimo Franchi & Paolo Paruolo

To cite this article: Massimo Franchi & Paolo Paruolo (2019) A general inversion theorem for cointegration, Econometric Reviews, 38:10, 1176-1201, DOI: 10.1080/07474938.2018.1536100

To link to this article: https://doi.org/10.1080/07474938.2018.1536100

Published online: 20 Dec 2018.

ECONOMETRIC REVIEWS
2019, VOL. 38, NO. 10, 1176–1201
https://doi.org/10.1080/07474938.2018.1536100

A general inversion theorem for cointegration

Massimo Franchi[a] and Paolo Paruolo[b]

[a]Sapienza University of Rome, Rome, Italy; [b]European Commission Joint Research Centre, Ispra, Italy

ABSTRACT
A generalization of the Granger and the Johansen Representation Theorems valid for any (possibly fractional) order of integration is presented. This Representation Theorem is based on inversion results that characterize 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 contribution unifies different approaches in the literature and extends them to an arbitrary order of integration. The role of deterministic terms is discussed in detail.

KEYWORDS
Cointegration; common trends; triangular representation; moving average representation; auto regressive representation

JEL CLASSIFICATION
C32

1. Introduction

The inversion of Moving Average (MA) forms into Auto Regressive (AR) forms (and vice versa) plays a central role in the representation theory of linear processes; see for instance Brockwell and Davis (1991, Chapter 3) for the case of ARMA stationary processes. This is also true for non-stationary integrated processes of order d, $I(d)$, i.e. processes X_t possessing a MA representation in dth differences $\Delta^d X_t = F(L)\varepsilon_t$ with $F(z)$ analytic for all $|z|<1+\delta, \delta>0, F(1)\neq 0$, and ε_t a white noise process; see Johansen (1996, Chapter 4) for the cases of $d=1$, 2.

The first result of this kind for $I(1)$ processes is the celebrated Granger Representation Theorem, see Granger (1981) and Engle and Granger (1987). Starting from $\Delta X_t = F(L)\varepsilon_t$, with $F(1)\neq 0$ singular, Engle and Granger (1987) considered the inversion of $F(z)$ in order to derive the (infinite order) Error-Correction form; their proof was completed by Johansen (1996, Theorem 4.5), using inversion results from Johansen (1991).

The Granger Representation Theorem also linked the Common Trends representation, derived by summation of the MA form, to the Error-Correction form, containing the cointegrating relations—associated with equilibrium in the system—and the adjustment toward it. This proved that error-correction, common trends, and cointegration were different characteristics of the same system and not competing concepts, see Granger (2004) and Hendry (2004).

The Granger Representation Theorem also established that there is complementarity between the (number of) common trends and the (number of) cointegrating relations, and paved the way to the interpretation of cointegrating relations as (deviations from) equilibrium and of common trends as drivers of the system.

The Granger Representation Theorem initiated a literature on representations for $I(d)$ systems, to which many authors have contributed. Starting from the MA form of an $I(1)$ system, Phillips

CONTACT Paolo Paruolo paolo.paruolo@ec.europa.eu European Commission Joint Research Centre, Ispra, Italy.

(1991) introduced the Triangular representation, which was subsequently generalized by Stock and Watson (1993) to the general $I(d)$ case.

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 colored process—is estimated nonparametrically; see Phillips and Hansen (1990), Sims et al. (1990), and Stock and Watson (1993).

An alternative derivation of the Granger Representation Theorem was presented in Yoo (1986), 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).[1]

In a parallel strand of literature, the cointegrated VAR literature, Johansen (1988a,b, 1991) considered the dual problem of inverting the AR representation $F(L)X_t = \varepsilon_t$ with $F(z)$ a matrix polynomial and $F(1) \neq 0$ singular; he derived conditions under which the Granger 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 $z = 1$, see also Schumacher (1991).[2] The reduced rank condition corresponds to the existence of a pole of some order $m \geq 1$ 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 pair of conditions 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, C_0 say—the MA impact matrix—as a function of the AR coefficients. Johansen (1994) used it to derive hypotheses on the constant and on deterministic terms; this led to cointegrated VAR models with restricted deterministic components, see Johansen (1996, Chapter 5.7) and Hansen (2005). Moreover, the explicit form of C_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 C_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 behavior of cointegrated systems also lead to long-run impact multipliers that are functions of the MA impact matrix C_0, see Johansen (2005). Omtzigt and Paruolo (2005) derived maximum likelihood estimation and inference on related long-run impact multipliers in cointegrated systems. The MA impact matrix plays also a central role in the estimation of the long-run variance matrix, see Paruolo (1997b) and Phillips (1998).

Still starting from the AR form, another derivation of the Granger 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 showed that the POLE(1) condition can be stated as the absence of a Jordan block of size > 1 in the Jordan representation of the AR companion matrix; see also Neusser (2000) for an approach based on the Drazin inverse.

A generalization of the Granger Representation Theorem to $I(2)$ AR processes was given in Johansen (1992), who stated the POLE(2) condition, under which X_t is $I(2)$, and he derived the corresponding Common Trends representation. The POLE(2) condition consists of two reduced rank

[1]The Smith form is also a standard tool in the treatment of vector ARMA processes, see e.g. Hannan and Deistler (1988, Section 1.2).

[2]The same condition can be found in the engineering literature, see Howlett (1982), Lancaster (1966, eq. (4.4.7)), Schumacher (1986).

restrictions and one full rank condition. The two reduced rank conditions correspond to the existence of a pole in $F(z)^{-1}$ at $z=1$ of some order $m \geq 2$, while the full rank condition establishes that the order of the pole m is exactly equal to two, see Franchi (2007) and Faliva and Zoia (2009).

Johansen (1992, 2008b) derived the explicit form of the first two matrices in the Laurent expansion of the inverse, C_0 and C_1 say, which load the cumulated random walk and the random walk components in the $I(2)$ Common Trends representation. The form of C_0 and C_1 shows in which directions the process X_t is $I(d)$, for $d=0$, 1, 2. The explicit expression of C_0 was instrumental in Paruolo (2002) to derive inference on it via likelihood methods.

In the AR framework, the case of generic $I(d)$ processes was considered by several authors. D'Autume (1992) showed that the maximal dimension of a Jordan block of the AR companion matrix identifies the order of integration for generic d. la Cour (1998) extended recursively the algebraic necessary and sufficient conditions of Johansen (1992) to the case of AR process integrated of any order d, and she described the associated cointegration properties of the system.

In the state space framework, Bauer and Wagner (2012) provided a canonical representation of processes with unit roots at arbitrary frequencies and arbitrary integer integration orders. In this approach, the order of integration is characterized as the maximal size of the Jordan blocks of the state matrix corresponding to the eigenvalue of unit modulus, in line with the results by D'Autume (1992) cited above on the companion matrix.

The main contribution of the present paper is to show that all these derivations can be unified via local spectral theory, see Gohberg et al. (1993), making use of the results in Franchi and Paruolo (2011b, 2016). In particular, this paper employs a general inversion theorem which (i) provides explicit expressions for (polynomial) cointegrating relations and (ii) common trend loading matrices; (iii) applies to processes integrated of any order, (iv) starting either from a MA or AR forms, (v) possibly in the presence of polynomial deterministic trends. This general inversion theorem offers a unified treatment of the different representations of cointegrated systems, irrespectively of the chosen starting point, extending them (when appropriate) to any order of integration.

These tools provide a constructive approach to compute the relevant matrices of each representation in terms of alternative ones. This is useful for the interpretation of cointegrated systems in terms of adjustment, equilibrium relations, common trends identification and loadings. Moreover, these results provide a way to specify the deterministic polynomial trends so as to bound the overall trend degree in the data. All these developments are key for the derivation of properties of cointegration processes, that are, e.g. useful in deriving asymptotics for estimators and tests.

For a given matrix function $F(z)$, the order m of the pole of $F(z)^{-1}$ at $z=1$ is shown to play a central role in the representation theory. When starting from the MA form $\Delta^d X_t = F(L)\varepsilon_t$, the order m, which is generally different from d, characterizes the cointegration properties of X_t. A generalization of the Triangular representation in Stock and Watson (1993)—which assumes $m = d$—is given; it is shown that the cointegrating relations involve cumulations (and possibly differences) of X_t when $m > d$, while they involve only differences of X_t when $m < d$. On the other hand, when starting from the AR form $F(L)X_t = \varepsilon_t$, the order m of the pole of the inverse gives the order of integration of the process and characterizes its cointegration structure. For $m=1$, 2 the representation results in Johansen (1996) are obtained.

The present results also apply to fractionally integrated processes, both in the case of ARFIMA and for the class introduced by Johansen (2008a,b), and further studied in Franchi (2010) and Johansen and Nielsen (2010, 2012). Furthermore, they can be applied to any stationary, unit or explosive root with minor modifications, thus covering also the case of seasonal cointegration, see Hylleberg et al. (1990) and Johansen and Schaumburg (1999), and of common cyclical features, see Engle and Kozicki (1993), Vahid and Engle (1993), and Franchi and Paruolo (2011a).

Finally, the Granger–Johansen Representation Theorems have recently been shown to hold also for infinite dimensional AR processes in Hilbert spaces, see Chang et al. (2016), Hu and

Park (2016), and Beare et al. (2017) for the $I(1)$ case and Beare and Seo (2018) for the $I(2)$ case. Franchi and Paruolo (2017) provide an extension of the present results to the generic $I(d)$ case for infinite dimensional AR processes in Hilbert spaces.

The rest of the paper is organized as follows: the remaining part of this introduction reports notational conventions and preliminaries; Section 2 introduces basic definitions; Section 3 contains the general inversion theorem; Section 4 presents a characterization of common trends, cointegration, and the Triangular representation of MA and AR processes based on the inversion results in Section 3, including a discussion of deterministic terms. Section 5 reports conclusions and Appendix A contains proofs.

1.1. Notation and preliminaries

In the following, $a := b$ or $b =: a$ indicates that a is defined equal to b; for any square matrix A, $|A|$ indicates its determinant, while for $z \in \mathbb{C}, |z|$ denotes the modulus of z. For any sequence $(v_t)_{t \in \mathbb{Z}}$, where $\mathbb{Z} := \{..., -2, -1, 0, 1, 1...\}$ is the set of integers, $\Delta := 1 - L$ indicates the difference operator and L is the lag operator, defined as $Lv_t := v_{t-1}$.

The paper considers the inversion of the $p \times p$ matrix function $F(z)$ with $F(1)$ singular in the MA form $\Delta^d X_t = F(L)\varepsilon_t$ or in the AR form $F(L)X_t = \varepsilon_t$. The matrix function $F(z)$ is assumed to be analytic for all $z \in \mathbb{C}$ satisfying $|z| < 1 + \delta$ for $\delta > 0$, so that the coefficients of its expansion around 0 are geometrically decreasing, and hence absolutely summable. This implies that $F(z)$ is infinitely differentiable and its derivatives are analytic in the same disc, see e.g. Lemma 3.2.10 in Greene and Krantz (1997). This includes finite order ARs or MAs, in which case $F(z)$ is a matrix polynomial, which is analytic for all $z \in \mathbb{C}$.

The process ε_t represents a $p \times 1$ white noise process with finite second moments; this is usually taken either as an i.i.d. process, see e.g. Johansen (1996) or as a martingale difference sequence, see e.g. Stock and Watson (1993). The choice of type of white noise is irrelevant for the representation results discussed in this paper, in the sense that each representation result holds for the specific chosen type of white noise.

In the invertible MA or causal AR cases, the point of interest for the expansion of $F(z)$ is 0, $F(z) = \sum_{n=0}^{\infty} F_n z^n$, and $F(0) = F_0 = I$ is nonsingular; the coefficients of the inverse $F(z)^{-1} =: C(z) = \sum_{n=0}^{\infty} C_n z^n$, which solves the system of equations $F(z)C(z) = C(z)F(z) = I$, are found using the following recursions, see e.g. Johansen (1996, Theorem 2.1),

$$C_0 = F_0^{-1}, \quad C_n = \sum_{k=1}^{n} K_k C_{n-k}, \quad K_k := -F_0^{-1} F_k, \quad n = 1, 2, ... \tag{1.1}$$

In the integrated case, the point of interest for the expansion of $F(z)$ is 1; at this point $F(z) = \sum_{n=0}^{\infty} F_n (1-z)^n$ is singular, i.e. $|F(1)| = 0$. This yields an inverse of $F(z)$ with a pole of some order $m = 1, 2, ...$ at $z = 1$.

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 to characterize the order of the pole. In the mathematical literature, a classical approach to characterize the relation between a matrix function and its inverse is the local spectral theory, see Gohberg et al. (1993), which is based on the concepts of root functions and partial multiplicities.

Within this literature, Franchi and Paruolo (2011b, 2016) introduced a procedure called "extended local rank factorization" (ELRF) which provides an explicit way to construct all the relevant quantities of the local spectral theory in Gohberg et al. (1993). Moreover, the ELRF was shown to provide an efficient way to compute the recursions in Avrachenkov et al. (2001) and Howlett et al. (2009), thus unifying these two different approaches. The results in Franchi and

Paruolo (2016) are reviewed in Section 3 below and act as building blocks for the representation results in Section 4, which are the novel contributions of this paper.

The paper makes repeated use of rank factorizations: 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, which respectively span the column space and the row space of φ; the negative sign is chosen for convenience in later 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 $\varphi = -\alpha\beta'$, i.e. the orthogonal complement of the column space of α.

The orthogonal projection matrix on the column space of $\varphi = -\alpha\beta'$ is indicated by $P_\alpha := \bar{\alpha}\alpha' = \alpha\bar{\alpha}'$, where $\bar{\alpha} := \alpha(\alpha'\alpha)^{-1}$, with rank r; the orthogonal projection matrix on the orthogonal complement of the column space of $\varphi = -\alpha\beta'$ is $P_{\alpha_\perp} := I - P_\alpha = \bar{\alpha}_\perp\alpha'_\perp = \alpha_\perp\bar{\alpha}'_\perp$, of rank $p-r$. Similarly, one defines P_β and $P_{\beta_\perp}$ replacing α with β.

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, one can set either (α, β) equal to (I, φ) or to (φ, I), with $\alpha_\perp = \beta_\perp = \bar{\alpha}_\perp = \bar{\beta}_\perp = 0$. The rank factorization is not unique, because all previous assignments of α, β can be replaced by αQ, $\beta Q'^{-1}$ with Q a generic nonsingular square matrix. Similarly, $\alpha_\perp$ and $\beta_\perp$ can be replaced by $\alpha_\perp H, \beta_\perp K$ with H, K generic nonsingular square matrices.

As a last piece of notation, $\mathcal{P}_n(t)$ indicates the set of scalar polynomials $p_n(t) = \sum_{i=0}^n c_i t^i$ in t or order n, with $c_i \in \mathbb{R}$; when $c_i \in \mathbb{R}^p$, $p > 1$, the class of vector polynomials $p_n(t) = \sum_{i=0}^n c_i t^i$ in t or order n is indicated $\mathcal{P}_{n,p}(t)$. The truncation of order q of a generic function $a(z) = \sum_{n=0}^\infty a_n(1-z)^n$ is denoted as $a^{(q)}(z) := \sum_{n=0}^q a_n(1-z)^n$, i.e. $a(z) = a^{(q)}(z) + (1-z)^{q+1}a^\star(z)$, where $a^\star(z) := \sum_{n=0}^\infty a_{n+q+1}(1-z)^n$ is the remainder.

2. Integrated processes

This section introduces the definitions of difference and integral operators, following Gregoir (1999) and Gregoir and Laroque (1994), and of integrated and cointegrated processes of any integer order (including negative ones), following Johansen (1996, Chapter 3).

Definition 2.1 (Difference operator Δ and integral operator $\mathcal{S}$). *For a generic process v_t, $t \in \mathbb{Z}$, the difference operator Δ is defined as $\Delta v_t := v_t - v_{t-1}$ and the integral operator $\mathcal{S}$ is defined as*[3]

$$\mathcal{S}v_t := 1_{(t\geq 1)} \cdot \sum_{i=1}^t v_i - 1_{(t\leq -1)} \cdot \sum_{i=t+1}^0 v_i, \tag{2.1}$$

where $1_{(\cdot)}$ is the indicator function.

Remark that by definition $\mathcal{S}$ assigns value 0 to the cumulated process at time 0. In fact, applying the definition, see Properties 2.1, 2.2 in Gregoir (1999) and Lemma A.2 in Appendix A, one can verify that, for $t \in \mathbb{Z}$, one has

$$\Delta\mathcal{S}v_t = v_t, \quad \mathcal{S}\Delta v_t = v_t - v_0, \quad \mathcal{S}1 = t. \tag{2.2}$$

Equation (2.2) shows that $\mathcal{S}$ applied to Δv_t regenerates the level of the process v_t, up to a constant; this parallels the constant of integration in indefinite integrals. The integral operator $\mathcal{S}$ is hence the inverse of the difference operator Δ up a constant; Definition 2.1 chooses this constant so as to make any cumulated process equal 0 at time $t=0$.

[3]In Gregoir (1999) $\mathcal{S}$ is denoted S_ω, for $\omega = 0$, where ω is the frequency.

When $v_t = \varepsilon_t$ is white noise, Eq. (2.1) shows that $\mathcal{S}\varepsilon_t$ is a bilateral random walk for $t \in \mathbb{Z}$. In fact for $t > 0$ one has $\mathcal{S}\varepsilon_t = \sum_{i=1}^{t} \varepsilon_i$ while for $t < 0$ one finds $\mathcal{S}\varepsilon_t = -\sum_{i=t+1}^{0} \varepsilon_i$, i.e. on both sides of $t=0$ a random walk is generated with increment ε_t for positive time t and $-\varepsilon_{t+1}$ for negative t.

The notion of integration of order 0 is presented next.

Definition 2.2 ($I(0)$ and $I_{nc}(0)$ processes). *Let $V(z)$ be a (rectangular) matrix function, analytic for all $|z|<1+\delta, \delta>0$, and let ε_t be a white noise process; if v_t, $t \in \mathbb{Z}$, satisfies*

$$v_t - \mathrm{E}(v_t) = V(L)\varepsilon_t, \quad V(1) \neq 0, \tag{2.3}$$

then v_t is said to be 'integrated of order zero', indicated $v_t \sim I(0)$, and 'integrated of order zero and non-cointegrated', indicated $v_t \sim I_{nc}(0)$, if in addition $V(1)$ has full row rank; in symbols:

$$v_t \sim I_{nc}(0): \qquad v_t - \mathrm{E}(v_t) = V(L)\varepsilon_t, \quad V(1) \neq 0 \textit{ has full row rank.} \tag{2.4}$$

The notation $I_{nc}(0)$ is introduced here to indicate explicitly the case in which v_t does not cointegrate (at frequency 0), see Remark 2.8. The next definition presents positive and negative orders of integration.

Definition 2.3 (Order of integration). *Let $v_t \sim I(0)$ as in (2.3) and let a, b be finite non-negative integers; if*

$$\Delta^a(z_t - \mathrm{E}(z_t)) = \Delta^b\big(v_t - \mathrm{E}(v_t)\big) = \Delta^b V(L)\varepsilon_t, \tag{2.5}$$

then z_t is said to be integrated of order $a-b$, indicated $z_t \sim I(a-b)$. Similarly, if $v_t \sim I_{nc}(0)$ as in (2.4), then z_t satisfying (2.5) is said to be integrated of order a–b and non-cointegrated, indicated $z_t \sim I_{nc}(a-b)$.

Definition 3.3 in Johansen (1996) of an $I(d)$ process is found by setting $b=0$ in (2.5). Note that $b>0$ allows to define also negative orders of integration. The order of integration is given by the difference between a and b, and can be thought of as "dividing both sides of (2.5) by Δ^b". In the following, expression of the type $\Delta^{-h}X_t \sim I(0)$ for positive h are understood to mean $X_t = \Delta^h v_t$ for some $v_t \sim I(0)$.

Some implications of Definition 2.3 on the simplification of Δ are discussed in Remark 2.4. The remarks in the rest of this section consider for simplicity the case of constant expectations $\theta_s := \mathrm{E}(s_t), s_t = z_t, v_t$, but can be modified for general $\mathrm{E}(s_t)$ in a straightforward way.

Remark 2.4. (Cancellations of Δ). Take $a = b = 1$ in (2.5), which in this case reads $\Delta z_t = \Delta v_t$ with $v_t \sim I(0)$. Applying the $\mathcal{S}$ operator on both sides one obtains $z_t - z_0 = v_t - v_0$, see (2.2).[4] If one assigns the initial value of z_0 equal to v_0, one obtains $z_t = v_t$, which corresponds to the cancelation of Δ from both sides of (2.5). The same reasoning applies for generic $a, b > 0$ to the cancelation of $\Delta^{\min(a,b)}$ from both sides of (2.5).

Remark 2.4. shows that one can simplify powers of Δ from both sides of (2.5) by properly assigning initial values; this observation is implicitly incorporated in Definition 2.3 of $I(d)$ processes, which is next specialized for $I(1)$ and $I(-1)$ processes in Remarks 2.5 and 2.6.

Remark 2.5 ($I(1)$ process). Set $a = 1, b = 0$ in (2.5); one finds that z_t is by definition $I(1)$ if it satisfies $\Delta z_t = \theta_v + V(L)\varepsilon_t$ with $V(1) \neq 0$. Expanding $V(z)$ around 1, $V(z) = V(1) + (1-z)V^\star(z)$ and applying the $\mathcal{S}$ operator to both sides of the equation, one finds thanks to (2.2), that

[4]This result is usually stated as $z_t = v_t - a_0$ where $a_0 := z_0 - v_0$ is a generic constant, see e.g. Hannan and Deistler (1988) eq. (1.2.15).

$$z_t - z_0 = \theta_v t + V(1)\mathcal{S}\varepsilon_t + V^\star(L)\mathcal{S}\Delta\varepsilon_t = \theta_v t + V(1)\mathcal{S}\varepsilon_t + V^\star(L)(\varepsilon_t - \varepsilon_0), \quad \text{i.e.}$$
$$z_t = \theta_v t + V(1)\mathcal{S}\varepsilon_t + y_t - (z_0 - y_0),$$

where $y_t := V^\star(L)\varepsilon_t$ is a stationary component, $z_0 - y_0$ depends on initial values[5] of z and y and $\mathcal{S}\varepsilon_t$ is a bilateral random walk. Note that $V(1) \neq 0$ guarantees that the random walk component does not vanish.

Remark 2.6 ($I(-1)$ process). Take $a = 0$ and $b = 1$ in Definition 2.3; Eq. (2.5) takes the form $z_t - \theta_z = \Delta v_t$ and applying the $\mathcal{S}$ operator one obtains $Z_t := \mathcal{S}z_t = v_t - v_0 + \theta_z t$, where $v_t := V(L)\varepsilon_t \sim I(0)$. Hence the cumulated process Z_t is the sum of an $I(0)$ process, a constant and a linear trend.

Remark 2.5 and 2.6 show that the $\mathcal{S}$ operator generates deterministic components (constants and trends in the cases above) whose coefficients depend on the initial values of the processes.

Definition 2.7 (Cointegrating relations). *Let $z_t \sim I(d)$ and let $b(L) = b_0 + b_1\Delta + b_2\Delta^2 + \cdots + b_n\Delta^n$ be a p $\times$ s matrix polynomial of order $n \geq 0$ in Δ, with b_0 of full column rank; then $b(L)$ is called a cointegrating matrix polynomial (of order n) if $b(L)'z_t \sim I_{nc}(d - j)$ for $j > 0$.*

Observe that Definition 2.7 applies to any order of integration d, including negative orders.

Remark 2.8 (Non-cointegration). Let $z_t \sim I(d)$ with positive d and $\theta_z = 0, \Delta^d z_t = V(L)\varepsilon_t$, with $V(L) = \sum_{i=0}^{\infty} V_i\Delta^i$. Observe that $b(L)'\Delta^d z_t = b(L)'V(L)\varepsilon_t$ and defines $H(L) := b(L)'V(L)$, with expansion $H(L) = \sum_{i=0}^{\infty} H_i\Delta^i$, where both $V(L)$ and $H(L)$ are expanded around $z = 1$.

One can see that $b(L)$ is a (polynomial) cointegration matrix if and only if one can factor Δ^j from $H(L)$ for some positive j. In particular for the case $j = 1$ one can factor Δ from $H(L)$ if and only if $H_0 = b_0'V_0 = 0$, where $V_0 = V(1)$. Note that one can have $b_0'V_0 = 0$ with nonzero b_0 if and only if $V_0 = V(1)$ has reduced row rank. This justifies the condition in (2.4). A similar situation applies to the case of negative d.

Remark 2.9 (Normalization of cointegrating relations). Definition 2.7 requires that $b_0 \neq 0$, which can be shown not to be a restriction. In fact, assume by contradiction that $b(L) = \Delta^q b^\star(L)$ with $b^\star(1) \neq 0$ and $q > 0$; in this case $b(L)'\Delta^d z_t = b(L)'V(L)\varepsilon_t$ would read $b^\star(L)'\Delta^{q+d}z_t = \Delta^q b^\star(L)'V(L)\varepsilon_t$, which can be simplified by Remark 2.4 as $b^\star(L)'\Delta^d z_t = b^\star(L)'V(L)\varepsilon_t$. This shows that $b_0 \neq 0$ is not a restriction, but a (convenient) normalization of a cointegrating relation.

Definition 2.7 also requires b_0 to be of full column rank. Again, this is not restrictive; in fact, in case b_0 is not of full column rank s but of rank $r < s$ say, one can rotate b_0 so that its first r columns are nonzero and of full rank, and all the remaining columns are equal to 0; then one can redefine b_0 as the set of these first r columns. This shows that requiring b_0 to be of full column rank is a (convenient) normalization of a cointegrating relation.

3. The inversion theorem

This section reports the main technical results on inversion, presented in Theorems 3.3 and 3.5; the former provides explicit expressions for the coefficients of the inverse function, while the latter provides a construction of the local Smith factorization. These theorems are restatements of results in Franchi and Paruolo (2011b, 2016) and are reported here because they are instrumental in obtaining the representation results in Section 4, which are the novel contributions of this paper.

[5]If one could choose the initial value z_0 of the process z_t equal to y_0, this would set the last term $z_0 - y_0$ to 0. Johansen (1996, Chapter 4) chooses z_0, z_{-1} so as to make $\beta' z_t$ and $\beta_\perp' \Delta z_t$ stationary, where β are the cointegrating linear combinations. This amount to requiring $\beta'(z_0 - y_0) = 0$ and $\beta_\perp'(\Delta z_0 - V(L)\varepsilon_0) = 0$. Any of these approaches on initial values can be applied to the more general case studied in Section 4.

Consider the problem of inversion of a matrix function

$$F(z)=\sum_{n=0}^{\infty}F_n(1-z)^n,\quad F_n\in\mathbb{R}^{p\times p},\quad F_0\neq 0,\quad |F_0|=0, \tag{3.1}$$

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

The inversion of $F(z)$ around the singular point $z=1$ yields an inverse with a pole of some order $m=1,2,\ldots$ at $z=1$; an explicit condition on the coefficients $\{F_n\}_{n=0}^{\infty}$ in (3.1) for $F(z)^{-1}$ to have a pole of given order m is described in Theorem 3.3; this is indicated as the $\text{POLE}(m)$ condition in the following. Under the $\text{POLE}(m)$ condition, $F(z)^{-1}$ has Laurent expansion around $z=1$ given by

$$F(z)^{-1}=:(1-z)^{-m}C(z)=\sum_{n=0}^{\infty}C_n(1-z)^{n-m},\quad C_0\neq 0,\quad |C_0|=0. \tag{3.2}$$

Note that $C(1)=C_0\neq 0$ is finite by construction and $C(z)$ is expanded around $z=1$. In the following, the coefficients $\{C_n\}_{n=0}^{\infty}$ are called the Laurent coefficients. The first m of them, $\{C_n\}_{n=0}^{m-1}$, make up the principal part and characterize the singularity of $F(z)^{-1}$ at $z=1$.

The next definition introduces quantities that are subsequently employed in the statements of the results.[6]

Definition 3.1 (Extended local rank factorization [ELRF]). *Let $0<r_0:=\text{rank}\ F_0<p$, $r_0^{\max}:=p-r_0$ and define α_0, β_0 by the rank factorization $F_0=-\alpha_0\beta_0'$. Moreover, for $j=1,2,\ldots$ define α_j, β_j by the rank factorization*

$$P_{a_{j\perp}}F_{j,1}P_{b_{j\perp}}=-\alpha_j\beta_j',\quad a_j:=(\alpha_0,\ldots,\alpha_{j-1}),\quad b_j:=(\beta_0,\ldots,\beta_{j-1}), \tag{3.3}$$

where P_x denotes the orthogonal projection on the space spanned by the columns of x and

$$F_{h+1,n}:=\begin{cases}F_n & \text{for } h=0\\ F_{h,n+1}+F_{h,1}\sum_{i=0}^{h-1}\bar{\beta}_i\bar{\alpha}_i'F_{i+1,n} & \text{for } h=1,2,\ldots\end{cases},\quad n=0,1,\ldots. \tag{3.4}$$

Finally, let

$$r_j:=\text{rank}\left(P_{a_{j\perp}}F_{j,1}P_{b_{j\perp}}\right),\quad r_j^{\max}:=p-\sum_{i=0}^{j-1}r_i, \tag{3.5}$$

and define

$$H_{j+1,n}:=\begin{cases}-1_{(n=m)}I & \text{for } j=0\\ H_{j,n+1}+F_{j,1}\sum_{i=0}^{j-1}\bar{\beta}_i\bar{\alpha}_i'H_{i+1,n} & \text{for } j=1,\ldots,m\end{cases},\quad n=0,1,\ldots, \tag{3.6}$$

where $1_{(\cdot)}$ is the indicator function.

[6]The case $r_0=0$ is excluded because otherwise one could re-define $F(z)$ factorizing $(1-z)^s$ from (3.1) for some positive s. The case $r_0=p$ is also excluded because it would imply $F(z_0)$ nonsingular, in which case the inversion formula (1.1) would apply.

The sequence of calculations in Definition 3.1 is called "extended local rank factorization" (ELRF). The coefficients F_n in (3.1) and

$$\{\alpha_j, \beta_j, r_j, F_{j+1,n}, H_{j+1,n}\}_{j=0,1,\dots,n=0,1,\dots} \tag{3.7}$$

are respectively the input and the output of the ELRF, performed at $z=1$.

Remark 3.2 (Bases with orthogonal blocks). Observe that α_h and α_j, $h \neq j$, are orthogonal matrices, $\alpha_h' \alpha_j = 0$; similarly this holds for β_h and β_j, $h \neq j$. Moreover, for some $j = 1, 2, \dots$, it is possible that $P_{a_{j\perp}} F_{j,1} P_{b_{j\perp}} = 0$, i.e. $r_j = 0$ and $\alpha_j = \beta_j = 0$. In this case, one needs to exclude α_j, β_j from a_{j+1}, b_{j+1} in (3.3). Note also that, as j increases, the space spanned by a_j and the space spanned by b_j are nondecreasing, and eventually coincide with $\mathbb{R}^p$ for some $j = s$; for all subsequent values of j, $j > s$, $P_{a_{j\perp}} F_{j,1} P_{b_{j\perp}}$ is equal to 0, because the orthogonal complements $a_{j\perp}$ and $b_{j\perp}$ have dimension 0, and hence all subsequent α_j, β_j are equal to 0. Thus there exists an integer s such that $r_s > 0$ and $r_j = 0$, $j > s$, and $a_{s+1} = (\alpha_0, \dots, \alpha_s)$ and $b_{s+1} = (\beta_0, \dots, \beta_s)$ are $p \times p$ nonsingular matrices with (nonzero) orthogonal blocks.

The next theorem states that the integer s such that $r_s > 0$ and $r_j = 0$ for all $j > s$ in Remark 3.2 is precisely the order m of the pole of $F(z)^{-1}$ at $z=1$; moreover, it provides a recursion for the Laurent coefficients, see (3.8), which generalizes formula (1.1) to the singular case.

Theorem 3.3 (POLE(m) condition and Laurent coefficients). *A necessary and sufficient condition for F(z) to have an inverse with pole of order $m = 1, 2, \dots$ at $z = 1$—called* POLE(m) *condition—is that*

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

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

$$C_n = \begin{cases} -\bar{\beta}_m \bar{\alpha}_m' & for\ n = 0 \\ H_n + \sum_{k=1}^{n} K_k C_{n-k} & for\ n = 1, \dots, m \\ \sum_{k=1}^{n} K_k C_{n-k} & for\ n = m+1, m+2, \dots \end{cases}, \qquad \begin{aligned} H_n &:= \sum_{j=0}^{m} \bar{\beta}_j \bar{\alpha}_j' H_{j+1,n} \\ K_k &:= \sum_{j=0}^{m} \bar{\beta}_j \bar{\alpha}_j' F_{j+1,k} \end{aligned}. \tag{3.8}$$

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 Howlett (1982, Theorem 3) and to the $I(1)$ condition in Johansen (1991, Theorem 4.1). Similarly, one has $m=2$ if and only if $r_1 < r_1^{\max}$,

$$r_2 = r_2^{\max}, \quad \text{where} \quad r_2 = \operatorname{rank} a_{2\perp}' \left(F_2 + F_1 \bar{\beta}_0 \bar{\alpha}_0' F_1\right) b_{2\perp} \quad \text{and} \quad r_2^{\max} = p - r_0 - r_1,$$

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

Theorem 3.3 is thus a generalization of the Johansen's $I(1)$ and $I(2)$ conditions and 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 $j = 0, \dots, m-1$ are reduced rank conditions, $r_j < r_j^{\max}$, which establish that the order of the pole is greater than j; the last one is a full rank condition, $r_m = r_m^{\max}$, which establishes that the order of the pole is exactly equal to m. These requirements make up the POLE(m) condition.

Theorem 3.3 provides in (3.8) a generalization of formula (1.1) to the singular case by giving a recursive expression of C_n in (3.2) in terms of the output of the ELRF. Equation (A.18) in the proof, see Appendix A, shows that H_n can be simplified as $H_n = -\bar{\beta}_{m-n} \bar{\alpha}_{m-n}' +$

$\sum_{j=m-n+1}^{m} \bar{\beta}_j \bar{\alpha}_j' H_{j+1,n}$ for $n = 0, 1, \ldots, m$. The additive term H_n in (3.8), which is absent in the nonsingular case, see (1.1), is present only for the steps $j = 1, \ldots, m$ in (3.8) and then disappears. 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 and the Moore-Penrose inverse of $\alpha_j \beta_j', \bar{\beta}_j \bar{\alpha}_j'$, in the singular case; see e.g. Theorem 5, p. 48, in Ben-Israel and Greville (2003) on Moore–Penrose inverses.

Finally, consider the local Smith factorization of $F(z)$ at $z=1$, see Gohberg et al. (1993), i.e. the factorization $F(z) = E(z)D(z)H(z)$, where $D(z) = \mathrm{diag}((1-z)^{\kappa_h})_{h=1,\ldots,p}$ is uniquely defined and contains the partial multiplicities $\kappa_1 \le \cdots \le \kappa_p$ of $F(z)$ at 1 and $E(z), H(z)$ are analytic and invertible in a neighborhood of $z=1$ and are nonunique. $D(z)$ and $E(z), H(z)$ are respectively called the local Smith form and extended canonical system of root functions of $F(z)$ at 1. Theorem 3.5 provides two constructions of the local Smith factorization in terms of the output of the ELRF.

The next definition introduces quantities that are employed in the statements of Theorem 3.5.[7]

Definition 3.4 (Root functions). *For $j = 0, \ldots, m$ and $n = 1, 2, \ldots,$ let*

$$\phi_{j,0} := -\bar{\alpha}_j, \quad \phi_{j,n} := \left(\bar{\alpha}_j' H_{j+1,m-j+n}\right)', \quad \gamma_{j,0} := \beta_j, \quad \gamma_{j,n} := -\left(\bar{\alpha}_j' F_{j+1,n}\right)',$$

and define the $p \times r_j$ matrix functions $\phi_j(z), \gamma_j(z)$ from

$$\phi_j(z) := \sum_{n=0}^{\infty} \phi_{j,n}(1-z)^n, \quad \gamma_j(z) := \sum_{n=0}^{\infty} \gamma_{j,n}(1-z)^n, \tag{3.9}$$

and the $p \times p$ matrix functions $\Phi(z), \Gamma(z)$ from

$$\Phi(z) := (\phi_0(z), \ldots, \phi_m(z))', \quad \Gamma(z) := (\gamma_0(z), \ldots, \gamma_m(z))'. \tag{3.10}$$

Similarly, for $j = 0, \ldots, m$ and $n = 1, 2, \ldots,$ let

$$\psi_{j,0} := -\bar{\beta}_j, \quad \psi_{j,n} := H_{j+1,m-j+n}\bar{\beta}_j, \quad \pi_{j,0} := \alpha_j, \quad \pi_{j,n} := -F_{j+1,n}\bar{\beta}_j$$

and define the $p \times r_j$ matrix functions $\psi_j(z), \pi_j(z)$ from

$$\psi_j(z) := \sum_{n=0}^{\infty} \psi_{j,n}(1-z)^n, \quad \pi_j(z) := \sum_{n=0}^{\infty} \pi_{j,n}(1-z)^n \tag{3.11}$$

and the $p \times p$ matrix functions $\Psi(z), \Pi(z)$ from

$$\Psi(z) := (\psi_0(z), \ldots, \psi_m(z)), \quad \Pi(z) := (\pi_0(z), \ldots, \pi_m(z)). \tag{3.12}$$

Finally, define the $p \times p$ matrix function $\Lambda(z)$ from

$$\Lambda(z) := \begin{pmatrix} (1-z)^0 I_{r_0} & & \\ & \ddots & \\ & & (1-z)^m I_{r_m} \end{pmatrix}. \tag{3.13}$$

The second result of this section is stated next.

Theorem 3.5 (Local Smith factorization). *One has that*

$$\Phi(z)F(z) = \Lambda(z)\Gamma(z), \quad |\Phi(1)| \neq 0, \quad |\Gamma(1)| \neq 0, \tag{3.14}$$

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

[7] In what follows, every statement concerning α_j or β_j implicitly assumes that they are nonzero, i.e. that $r_j > 0$. The modifications required in the case $r_j = 0$ are straightforward.

$$F(z)\Psi(z) = \Pi(z)\Lambda(z), \quad |\Psi(1)| \neq 0, \quad |\Pi(1)| \neq 0, \tag{3.15}$$

i.e. $\Psi(z)$ *and* $\Pi(z)$ *are extended canonical system of right root functions.*

Theorem 3.5 shows that the ELRF fully characterizes the elements of the local Smith factorization of $F(z)$ at 1. In fact, the values of j with $r_j > 0$ in the ELRF provide the distinct partial multiplicities of $F(z)$ at 1 and r_j gives the number of partial multiplicities that are equal to a given j; this characterizes the local Smith form $\Lambda(z)$. Moreover, it also provides two constructions of extended canonical system of root functions.

Remark that the jth block of rows in (3.14) can be written as

$$\phi_j(z)'F(z) = (1-z)^j\gamma_j(z)', \quad \gamma_j(z)'C(z) = (1-z)^{m-j}\phi_j(z)', \quad j = 0, \ldots, m, \tag{3.16}$$

where $\gamma_j(1)' = \beta_j'$ and $\phi_j(1)' = -\bar{\alpha}_j'$ have full row rank. This shows that $\phi_j(z)'$ are r_j left root functions of order j of $F(z)$ and that $\gamma_j(z)'$ are r_j left root functions of order $m-j$ of $C(z)$. As shown in Theorems 4.1 and 4.3, the concept of cointegrating relation coincides with that of left root function and its order of integration is equal to the corresponding entry in the local Smith form $\Lambda(z)$, i.e. to the corresponding partial multiplicity.

Similarly, observe that the jth block of columns in (3.15) can be written as

$$F(z)\psi_j(z) = (1-z)^j\pi_j(z), \quad C(z)\pi_j(z) = (1-z)^{m-j}\psi_j(z), \quad j = 0, \ldots, m, \tag{3.17}$$

where $\psi_j(1) = -\bar{\beta}_j$ and $\pi_j(1) = \alpha_j$ have full column rank. That is, $\psi_j(z)$ are r_j right root functions of order j of $F(z)$ and $\pi_j(z)$ are r_j right root functions of order m–j of $C(z)$. This fact will be used when discussing deterministic terms in Theorem 4.8.

4. Common Trends, cointegration, and Triangular representations

This section contains the novel representation results; these include the explicit expressions of the matrix coefficients of the (polynomial) cointegrating relations, of the Common Trends and Triangular representations, either starting from the MA or the AR form of an $I(d)$ process. In particular, Section 4.1 (respectively Section 4.2) considers a generic MA (respectively AR) form and describes its cointegration properties in Theorem 4.1 (respectively Theorem 4.3) and its Triangular representation in Corollary 4.2 (respectively Corollary 4.6). This includes the Triangular representation in Stock and Watson (1993) as a special case.

Moreover, Corollaries 4.4 and 4.5 in Section 4.2 present the Granger Representation Theorem and the Johansen Representation Theorem for AR forms as special cases of Theorem 4.3. Section 4.3 considers the case with deterministic terms, Section 4.4 describes the explicit connection between the local Smith form and the Jordan structure and Section 4.5 discusses the case of non-integer d.

All the results in this section follow from Theorems 3.3 and 3.5, which thus prove to be unifying and useful tools in the representation theory of cointegrated processes.

4.1. MA forms

Consider a generic $I(d)$ process

$$\Delta^d X_t = F(L)\varepsilon_t, \quad F_0 \neq 0, \quad |F_0| = 0, \tag{4.1}$$

with $F(z)$ analytic for all $|z| < 1+\delta, \delta > 0$, having roots at $z = 1$ and at $|z| > 1$. This includes finite order MAs, in which case $F(z)$ is a matrix polynomial. Applying Theorems 3.3 and 3.5 to $F(z)$ in (4.1), one obtains the following result.

Theorem 4.1 (Cointegration properties of MA processes). *Write (3.1) as $F(z) = \sum_{n=0}^{d-1} F_n(1-z)^n + (1-z)^d F_d(z)$, let $Y_t := F_d(L)\varepsilon_t$ and for $h \in \mathbb{N}$ define the h-fold cumulated bilateral random walk $S_{h,t} := \mathcal{S}^h \varepsilon_t \sim I_{nc}(h)$; then the I(d) process X_t in (4.1) admits the following Common Trends representation for $t \in \mathbb{Z}$:*

$$X_t = \sum_{n=0}^{d-1} F_n S_{d-n,t} + Y_t + v(t), \tag{4.2}$$

where Y_t is stationary, $v(t) := \sum_{n=0}^{d-1} v_n t^n \in \mathcal{P}_{d-1,p}(t)$ where $v_0, \ldots, v_{d-1}$ depend on initial values of X_t, Y_t, ε_t for $t = -d, \ldots, 0$.

Next assume that F(z) in (4.1) satisfies the POLE*(m) condition; then the cointegration properties of X_t are fully described by the cointegrating relations*

$$\phi_j^{(j-1)}(L)' X_t \sim I_{nc}(d-j), \quad j = 1, \ldots, m, \tag{4.3}$$

where $\phi_j^{(j-1)}(z) = \sum_{k=0}^{j-1} \phi_{j,k}(1-z)^k$ is the truncation of order $j-1$ of the left root functions $\phi_j(z)$ in (3.9). Additionally, defining $\Phi_\circ(z) := (\bar{\alpha}_0, \phi_1^{(0)}(z), \ldots, \phi_m^{(m-1)}(z))'$, one has

$$\Lambda(L)^{-1} \Phi_\circ(L) \Delta^d X_t \sim I_{nc}(0), \quad |\Phi_\circ(1)| \neq 0, \tag{4.4}$$

where $\Lambda(z)$ is the Local Smith form of F(z), see (3.13), *and $\Phi_\circ(z)$ is a truncation of the extended canonical system of left root functions $\Phi(z)$ in* (3.10). *Moreover, the initial values can be chosen so that v(t) does not appear in* (4.4).

Note that the cointegrating relations coincide with the truncated left root functions of *F*(*z*), in this case chosen as $\phi_j(z)$, that the order of integration of a cointegrating relation is equal to the corresponding partial multiplicity and that the cointegration structure of X_t coincides with the truncation of an extended canonical system of left root functions of *F*(*z*), in this case chosen as $\Phi(z)$.

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

Corollary 4.2 (Triangular representation of MA processes). *Let X_t in (4.1) satisfy the* POLE*(m) condition on F(z); then X_t admits the Generalized Triangular representation*

$$\begin{pmatrix} \bar{\alpha}_0' \Delta^d X_t \\ \bar{\alpha}_1' \Delta^{d-1} X_t \\ \bar{\alpha}_2' \Delta^{d-2} X_t - \phi_{2,1}' \Delta^{d-1} X_t \\ \vdots \\ \bar{\alpha}_m' \Delta^{d-m} X_t - \sum_{k=1}^{m-1} \phi_{m,k}{}' \Delta^k X_t \end{pmatrix} \sim I_{nc}(0),$$

which reduces to the Triangular representation in Eq. (3.2) *of* Stock and Watson (1993) *in the special case $m = d$.*

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 $F(z)^{-1}$. On the other hand, the cointegration properties of X_t do not depend on the order of integration d but on the order m of the pole, which is associated with the structure of *F*(*z*).

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

$$\begin{pmatrix} \bar{\alpha}_0' \Delta^d X_t \\ \bar{\alpha}_1' \Delta^{d-1} X_t \end{pmatrix} \sim I_{nc}(0).$$

In this case, (α_0, α_1) is a block orthogonal basis of $\mathbb{R}^p$ and one has that $\alpha_0' X_t \sim I_{nc}(d)$, $\alpha_1' X_t \sim I_{nc}(d-1)$; this fully describes the cointegration properties of $X_t \sim I(d)$ and shows that no polynomial cointegration arises even though the order of integration is greater than one.

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

$$\begin{pmatrix} \bar{\alpha}_0' \Delta X_t \\ \bar{\alpha}_1' X_t \\ \bar{\alpha}_2' \Delta^{-1} X_t - \phi_{2,1}' X_t \\ \\ \bar{\alpha}_m' \Delta^{1-m} X_t - \sum_{k=1}^{m-1} \phi_{m,k}{}' \Delta^{1-m+k} X_t \end{pmatrix} \sim I_{nc}(0).$$

In this case, cointegrating relations occur in the direction of α_j, $j \neq 0$, and, if $j > 1$, they require cumulation of X_t in order to obtain an $I_{nc}(0)$ on the r.h.s. In fact, $(\alpha_0, \ldots, \alpha_m)$ is a block orthogonal basis of $\mathbb{R}^p$ and one has that $\alpha_0' X_t \sim I_{nc}(1)$, $\alpha_1' X_t \sim I_{nc}(0)$, $\bar{\alpha}_2' X_t - \phi_{2,1}' \Delta X_t \sim I_{nc}(-1)$, and so on until $\bar{\alpha}_m' X_t - \sum_{k=1}^{m-1} \phi_{m,k}{}' \Delta^k X_t \sim I_{nc}(1-m)$.

In general, Corollary 4.2 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 due to the fact that m can be greater than d. In this case $\Delta^j X_t$ corresponds to cumulations of X_t, see Definition 2.3 and Remark 2.6. While m does not influence the order of integration of X_t, it does impact the number of differences or cumulations that enter the cointegrating relations of X_t and thus determines the Generalized Triangular representation of the process.

4.2. AR forms

Consider a generic AR process

$$F(L) X_t = \varepsilon_t, \quad F_0 \neq 0, \quad |F_0| = 0, \tag{4.5}$$

with $F(z)$ analytic for all $|z| < 1 + \delta, \delta > 0$, having roots at $z = 1$ and at $|z| > 1$. This includes finite order ARs, in which case $F(z)$ is a matrix polynomial and hence it is analytic for all $z \in \mathbb{C}$. One can then apply Theorems 3.3 and 3.5 to $F(z)$ in (4.5), obtaining the following result.

Theorem 4.3 (Cointegration properties of AR processes). *The AR process X_t in (4.5) is I(d), $d = m$, if and only if the* POLE*(m) condition applies to $F(z)$. Write $C(z)$ in (3.2) as $C(z) = \sum_{n=0}^{d-1} C_n (1-z)^n + (1-z)^d C_d(z)$ and let $Y_t := C_d(L)\varepsilon_t$; then X_t admits the following Common Trends representation:*

$$X_t = \sum_{n=0}^{d-1} C_n S_{d-n,t} + Y_t + v(t) \tag{4.6}$$

where $S_{h,t} := \mathcal{S}^h \varepsilon_t$, Y_t is I(0), $v(t) := \sum_{n=0}^{d-1} v_n t^n \in \mathcal{P}_{d-1,p}(t)$ where $v_0, \ldots, v_{d-1}$ depend on the initial values of X_t, Y_t, ε_t for $t = -d, \ldots, 0$. The cointegration properties of X_t are fully described by the cointegrating relations

$$\gamma_j^{(m-j-1)}(L)' X_t \sim I_{nc}(j), \quad j = 0, 1, \ldots, m-1, \tag{4.7}$$

where $\gamma_j^{(m-j-1)}(z) = \sum_{k=0}^{m-j-1} \gamma_{j,k}(1-z)^k$ is the truncation of order $m-j-1$ of the left root functions $\gamma_j(z)$ in (3.9). Additionally, defining $\Gamma_\circ(z) := (\gamma_0^{(m-1)}(z), \ldots, \gamma_{m-1}^{(0)}(z), \beta_m)'$, one has

$$\Lambda(L)\Gamma_\circ(L)X_t \sim I_{nc}(0), \quad |\Gamma_\circ(1)| \neq 0, \tag{4.8}$$

where $\Lambda(z)$ is the Local Smith form of F(z), see (3.13), *and $\Gamma_\circ(z)$ is a truncation of the extended canonical system of left root functions $\Gamma(z)$ in* (3.10). *Moreover the initial values can be chosen so that v(t) does not appear in* (4.8).

Note that (i) the cointegrating relations coincide with the truncated left root functions of $C(z)$, in this case chosen as $\gamma_j(z)$, (ii) the order of integration of a cointegrating relation is equal to the corresponding partial multiplicity, and (iii) that the cointegration structure of X_t coincides with the truncation of an extended canonical system of left root functions of $C(z)$, in this case chosen as $\Gamma(z)$.

Setting $m = 1$ in Theorem 4.3 one finds Theorem 4.2 in Johansen (1996), as reported in the following corollary.

Corollary 4.4 (Cointegration properties of $I(1)$ AR processes). *The AR process X_t in (4.5) is I(1) if and only if the* POLE*(1) condition applies to F(z). Write C(z) in (3.2) as $C(z) = C_0 + (1-z)C_1(z)$ and let $Y_t := C_1(L)\varepsilon_t$; then X_t admits the following Common Trends representation:*

$$X_t = -\bar{\beta}_1\bar{\alpha}_1' S_{1,t} + Y_t + v_0,$$

where $S_{1,t} := \mathcal{S}\varepsilon_t$, Y_t is I(0) and v_0 depends on the initial values of X_t, Y_t, ε_t for $t = -1, 0$. The cointegration properties of X_t are fully described by the cointegrating relations

$$\begin{pmatrix} \beta_0' X_t \\ \beta_1' \Delta X_t \end{pmatrix} \sim I_{nc}(0) \tag{4.9}$$

and the initial values can be chosen so that v_0 does not appear in (4.9).

Similarly, setting $m = 2$ in Theorem 4.3 one finds Theorem 4.6 in Johansen (1996), as reported in the following corollary.

Corollary 4.5 (Cointegration properties of $I(2)$ AR processes). *The AR process X_t in (4.5) is I(2) if and only if the* POLE*(2) condition applies to F(z). Write C(z) in (3.2) as $C(z) = C_0 + C_1(1-z) + (1-z)^2 C_2(z)$ and let $Y_t := C_2(L)\varepsilon_t$; then X_t admits the following Common Trends representation:*

$$X_t = -\bar{\beta}_2\bar{\alpha}_2' S_{2,t} + C_1 S_{1,t} + Y_t + v_0 + v_1 t,$$

where $S_{h,t} := \mathcal{S}^h\varepsilon_t$,

$$C_1 = -\bar{b}\begin{pmatrix} 0 & 0 & \bar{\alpha}_0' F_{1,1}\bar{\beta}_2 \\ 0 & I_{r_1} & \bar{\alpha}_1' F_{2,1}\bar{\beta}_2 \\ \bar{\alpha}_2' F_{1,1}\bar{\beta}_0 & \bar{\alpha}_2' F_{2,1}\bar{\beta}_1 & \bar{\alpha}_2' F_{3,1}\bar{\beta}_2 \end{pmatrix}\bar{a}', \quad a = (\alpha_0, \alpha_1, \alpha_2), \quad b = (\beta_0, \beta_1, \beta_2),$$

Y_t is I(0) and v_0, v_1 depend on initial values of X_t, Y_t, ε_t for $t = -2, -1, 0$. The cointegration properties of X_t are fully described by the cointegrating relations

$$\begin{pmatrix} \beta_0' X_t - \bar{\alpha}_0' F_1 \Delta X_t \\ \beta_1' \Delta X_t \\ \beta_2' \Delta^2 X_t \end{pmatrix} \sim I_{nc}(0) \tag{4.10}$$

and the initial values can be chosen so that v_0, v_1 do not appear in (4.10).

Theorem 4.3 leads to a Triangular representation, as shown in the following corollary.

Corollary 4.6 (Triangular representation of AR processes). *Let X_t in (4.5) satisfy the* POLE*(m) condition on F(z); then X_t is I(d) with $d = m$ and it admits the Triangular representation*

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

which coincides with the one in Eq. (3.2) *of* Stock and Watson (1993).

Note that in the AR case, differently from the MA case, the cointegrating relations do not involve cumulations of X_t but exclusively differences.

Comparing the cointegration properties of MA and AR processes in Theorems 4.1 and 4.3, one sees that the two extended canonical system of left root functions $\Phi(z)$ and $\Gamma(z)$ in Definition 3.4 play a symmetric role; the first one is used when starting from a MA form, and the second one when starting from a AR form. Moreover, the order of integration of a cointegrating relation is equal to the corresponding entry in the local Smith form $\Lambda(z)$ in Definition 3.4.

Remark 4.7 (Left root functions and cointegrating relations). These results show that (i) the concept of cointegrating relation coincides with that of (truncated) left root function, (ii) that the order of integration of a cointegrating relation is equal to the corresponding partial multiplicity and (iii) that the cointegration structure is fully described by an extended canonical system of left root functions, see panels (a–c) in Table 1.

4.3. Deterministic terms

This section extends Theorems 4.1 and 4.3 to the case in which deterministic terms μ_t are added to (4.1) or (4.5) as in

$$\Delta^d X_t = F(L)(\varepsilon_t + \mu_t) \tag{4.11}$$

or in

$$F(L) X_t = \varepsilon_t + \mu_t, \tag{4.12}$$

where μ_t is in the class $\mathcal{P}_{u,p}(t)$ of p-vector polynomials of order u in t.

Table 1. A glossary of notation. Elements involved in the inversion of $F(z)$, their mathematical definition and their econometric meaning.

Symbol	Equation	Mathematical definition	Econometric meaning
(a) MA forms, Sections 4.1 & 4.3			
$\phi_j(z)$	(3.9)	Left root functions	$I(j)$ cointegrating relations
$\Phi(z)$	(3.10)	System of left root functions	Triangular representation
$\psi_j(z)$	(3.11)	Right root functions	Limited-cumulation deterministic terms
$\Psi(z)$	(3.12)	System of right root functions	Structure of deterministic terms
(b) AR forms, Sections 4.2 & 4.3			
$\gamma_j(z)$	(3.9)	Left root functions	$I(j)$ cointegrating relations
$\Gamma(z)$	(3.10)	System of left root functions	Triangular representation
$\pi_j(z)$	(3.11)	Right root functions	Limited-cumulation deterministic terms
$\Pi(z)$	(3.12)	System of right root functions	Structure of deterministic terms
(c) MA and AR forms, Sections 4.1, 4.2 & 4.3			
$\Lambda(z)$	(3.13)	Local Smith form	Integration orders in triangular forms
(d) ELRF, Section 3			
j	(3.7)	Partial multiplicities	Integration order of cointegrating relations
r_j	(3.5)	Number of partial multiplicities $= j$	Number of $I(j)$ cointegrating relations
α_j, β_j	(3.3)	Bases of subspaces	Linear combinations involved in CI, CT
$F_{j,n}, H_{j,n}$	(3.4) & (3.6)	Recursive coefficients from $F(z)$	Building blocks of CI, CT

The generic polynomial $\mu_t := \sum_{n=0}^{u} c_n t^n \in \mathcal{P}_{u,p}(t)$ can be represented as $\mu_t = a(L)t^u$, where $a(L) := \sum_{n=0}^{u} a_n (1-L)^n$ is a $p \times 1$ vector polynomial; this is because $\Delta t^u = \sum_{n=1}^{u} \binom{u}{n} (-1)^n t^{u-n} \in \mathcal{P}_{u-1}(t)$, $\Delta^j t^u \in \mathcal{P}_{u-j}(t)$ for $j \le u$ and $\Delta^{u+1} t^u = 0$, see Lemma A.2. Hence one has

$$\mu_t := \sum_{n=0}^{u} c_n t^n = a(L)t^u, \quad a(L) := \sum_{n=0}^{u} a_n (1-L)^n, \quad a_0 = c_u \neq 0. \tag{4.13}$$

In the MA case (4.11), applying the $\mathcal{S}^d$ operator to both sides of (4.11) as in Theorem 4.1, one finds that X_t includes the term $F(L)\mathcal{S}^d\mu_t$, which generally is an element of $\mathcal{P}_{u+d,p}(t)$, i.e. a deterministic $p \times 1$ vector polynomial of order u+d. In the AR case (4.12), the inverse of $F(z)$ is $(1-z)^{-m}C(z)$ and, setting $d = m$, one obtains the equation $\Delta^d X_t = C(L)(\varepsilon_t + \mu_t)$. By the same reasoning as in the MA case, X_t hence includes the term $C(L)\mathcal{S}^d\mu_t$, which generally is an element of $\mathcal{P}_{u+d,p}(t)$, i.e. a deterministic $p \times 1$ vector polynomial of order u+d.

This general rule applies *unless there are cancelations* in the leading terms of $F(L)\mathcal{S}^d\mu_t$ or $C(L)\mathcal{S}^d\mu_t$, i.e. in the coefficients of the highest powers of t. The highest order trend t^{d+u} is loaded into X_t by $F_0 a_0$ in the MA case and by $C_0 a_0$ in the AR case. Given that both F_0 and C_0 have reduced rank, see Theorems 4.1 and 4.3, one can have cancelations of these coefficients for appropriate choices of a_0. Similarly, one can have cancelations of more coefficients, and hence a deterministic polynomial of given order, for appropriate choices of the a_n coefficients in (4.13).

The contribution of this section is to describe the conditions on μ_t that give rise to reductions in the order $u + d$ of the polynomial trend. In particular, it is shown that the reduction in the order of the trend is at most equal to m, the order of the pole of $F(z)$ at 1. In the analysis, the extended canonical system of right root functions, chosen here as $\Psi(z)$ and $\Pi(z)$ in Definition 3.4, play a central role; the former is used when starting from a MA form and the latter when starting from a AR form. This is the dual of the role played by the extended canonical system of left root functions, chosen above as $\Phi(z)$ and $\Gamma(z)$, which were used in Sections 4.1 and 4.2 to characterize the cointegrating relations.

Theorem 4.8 (Cointegration properties with deterministic terms). *Let $X_t \sim I(d)$ satisfy (4.11) or (4.12), where μ_t is defined in (4.13), let $\psi_j(z), \pi_j(z)$ be as in (3.11) and define*

$$\psi_{j:m}(z) := \left(\psi_j(z), \ldots, \psi_m(z)\right), \quad \pi_{0:j}(z) := \left(\pi_0(z), \ldots, \pi_j(z)\right); \tag{4.14}$$

finally, let j be a fixed integer in the range $0 \le j \le m$. Then:

MA) A necessary and sufficient condition for

i.1) X_t in (4.11) to contain trends in the class $\mathcal{P}_{q,p}(t)$ of order $q := d + u - j$, and

i.2) $\phi_h(L)'X_t \sim I_{nc}(d-h)$ to contain trends in the class $\mathcal{P}_{s_h,r_h}(t)$ of order $s_h \le d + u - h$ for $j<h \le m$ and $s_j = d + u - j$ for $h = j$

is that:

$$a(z) = \psi_{j:m}^{(u)}(z)\varphi, \tag{4.15}$$

where $\varphi := (\varphi_j', \ldots, \varphi_m')'$, $\varphi_j \neq 0$, is partitioned conformably with the block of right root functions $\psi_{j:m}(z)$ in (4.14) and $\psi_{j:m}^{(u)}(z)$ is the truncation of order u of $\psi_{j:m}(z)$.

AR) Similarly, a necessary and sufficient condition for

ii.1) X_t in (4.12) to contain trends in the class $\mathcal{P}_{q,p}(t)$ of order $q := u + j$, and

ii.2) $\gamma_h(L)'X_t \sim I_{nc}(h)$ to contain trends in the class $\mathcal{P}_{s_h,r_h}(t)$ of order $s_h \le u + h$ for $0 \le h<j$ and $s_j = u + j$ for $h = j$

is that:

$$a(z) = \pi_{0:j}^{(u)}(z)\varphi, \tag{4.16}$$

where $\varphi := (\varphi_0', \ldots, \varphi_j')'$, $\varphi_j \neq 0$, is partitioned conformably with the block of right root functions $\pi_{0:j}(z)$ in (4.14) and $\pi_{0:j}^{(u)}(z)$ is the truncation of order u of $\pi_{0:j}(z)$.

As a particular case of this result, one finds the analysis in Section 5.7 in Johansen (1996). In fact, set $m=1$ and $j=0$ in Theorem 4.8*ii*); in this case (4.16) states that if $a(z) = \pi_0^{(u)}(z)\varphi = \alpha_0\varphi$, then $X_t \sim I(1)$, solution of $F(L)X_t = \varepsilon_t + \mu_t$, involves a deterministic trend of order $u+0=u$, and viceversa. For $j=1$, (4.16) states that if $a(z) = \pi_0^{(u)}(z)\varphi_0 + \pi_1^{(u)}(z)\varphi_1$ with $\varphi_1 \neq 0$ then $X_t \sim I(1)$, solution of $F(L)X_t = \varepsilon_t + \mu_t$, involves a deterministic trend of order $u+1$, and viceversa. This case for $u=1$ (respectively $u=0$) is analyzed in (5.13) and (5.14) (respectively in (5.15) and (5.16)) in Johansen (1996).

Remark 4.9 (Limited-cumulation deterministic terms and right root functions). Equations (4.15) and (4.16) characterize deterministic components that have a given controlled degree of cumulation; in Table 1, they are indicated as "limited-cumulation" deterministic terms. These results show that the structure of the deterministic terms is fully described by an extended canonical system of right root functions, see panels (a,b) in Table 1.

4.4. 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 $(j_1, j_2, \dots, j_{w+1})$ 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. $\mathcal{J}_+ := \mathcal{J} \setminus \{0\} = (j_1, j_2, \dots, j_w)$. Note that the index set $\mathcal{J}_+$ contains at least one element (equal to m), and at most m elements, $\mathcal{J}_+ = (m, m-1, \dots, 1)$, and hence $1 \le w \le m$. Finally let $\mathcal{K}$ be the ordered set that contains each $j \in \mathcal{J}_+$ repeated r_j times and indicate its elements by $(k_1, k_2, \dots, k_{p-r_0}) := \mathcal{K}$, i.e.

$$\mathcal{K} := (\underbrace{j_1, \dots, j_1}_{r_{j_1}\,\text{times}}, \underbrace{j_2, \dots, j_2}_{r_{j_2}\,\text{times}}, \dots, \underbrace{j_w, \dots, j_w}_{r_{j_w}\,\text{times}}) =$$

$$= (\underbrace{k_1, \dots, k_{r_{j_1}}}_{=j_1}, \underbrace{k_{r_{j_1}+1}, \dots, k_{r_{j_1}+r_{j_2}}}_{=j_2}, \dots \underbrace{k_{\sum_{i=1}^{w-1} r_{j_i}+1}, \dots, k_{p-r_0}}_{=j_w}).$$

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 $a_{j_1}, \dots, a_{j_w}$ on the main diagonal.

Given the extended canonical system of left root functions $\Phi(z)$ in (3.10) and the index set $\mathcal{K}$, one can construct a Jordan pair of $F(z)$ at $z=1$ as follows.[8]

Theorem 4.10 *(Jordan pair at $z = 1$). Let $u_{i,n}$ be the ith column of Φ_n in the extended canonical system of left root functions $\Phi(z) = \sum_{n=0}^{\infty} \Phi_n (1-z)^n$ in (3.10), and let k_i be the ith element in the index set $\mathcal{K}$; for $i = 1, \dots, p - r_0$, define*

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

respectively of dimension $p \times k_i$ and $k_i \times k_i$. Then the columns of U_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

[8] Similar results apply to $\Gamma(z)$, $\Psi(z)$ and $\Pi(z)$.

$$U := (U_i)_{i=1}^{p-r_0}, \quad J := \mathrm{diag}\big(I_{r_j} \otimes J_j\big)_{j\in\mathcal{J}_+},$$

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

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 collected in the matrix J in Theorem 4.10; 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.

4.5. Fractional integration orders

The present results also apply to the cases of noninteger d of the ARFIMA type. This can be seen by choosing $d \in \mathbb{R}$ in the MA form (4.1), or replacing X_t with $(1-L)^s X_t$, $s \in \mathbb{R}$, in the AR form (4.5), i.e. $F(L)(1-L)^s X_t = \varepsilon_t, s \in \mathbb{R}$. The present analysis applies as well to the class of fractionally integrated processes defined in Johansen (2008a,b), see Eq. (3.1) in Franchi (2010). In fact, one can replace L with $L_b := 1-(1-L)^b, b \in \mathbb{R}$ and consider the fractional version of (4.5) $F(L_b)X_t = \varepsilon_t$ with $b \in \mathbb{R}, m \in \mathbb{N}$.

5. Conclusions

The present results show that the concepts of root functions and partial multiplicities in the local spectral theory are central for the representation theory of cointegrated systems. In particular, the concept of cointegrating relation coincides with that of left root function and the order of integration of a cointegrating relation is equal to the corresponding partial multiplicity. Moreover, the impact of deterministic terms on the process is shown to be determined by the characteristics of right root functions and the corresponding partial multiplicities.

The general inversion results deliver both left and right extended canonical system of root functions and the partial multiplicities as recursive expressions of the coefficients of the matrix function to be inverted. The inversion theorem is 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 inversion is performed. The present results unify and clarify existing representation results in the literature, and extend them to any integer order. The present derivations carry over to fractionally integrated processes and they can be applied to any (stationary, unit, explosive) root, which characterize seasonal cointegration and common cyclical features.

Appendix A: Proofs

Let $\xi_n(T) := \sum_{t=1}^T t^n$ for $n = 1, 2, 3, \ldots$ and $T = 1, 2, \ldots$ be defined as in Exercises 5, 6, 7 on page 83 of Anderson (1971), who used the symbol $\psi_n(T)$ for $\xi_n(T)$. The following lemma discusses properties of $\xi_n(T)$ as a polynomial in T. Recall that $\mathcal{P}_u(t)$ indicates the set of polynomials in t or order u.

Lemma A.1 (Sums of powers). *The following three recursive formulae hold for* $\xi_n(T)$ *for generic* $n = 1, 2, 3, \ldots$, *for* n *even* $(n = 2q)$, *and for* n *odd* $(n = 2q-1)$, *with* $q = 1, 2, 3, \ldots$:

$$\xi_n(T) = \frac{1}{n+1}\left((T+1)^{n+1} - (T+1) - \sum_{k=1}^{n-1}\binom{n+1}{k}\xi_k(T)\right), \tag{A.1}$$

$$\xi_{2q}(T) = \frac{1}{2q+1}\left(\frac{(T+1)^{2q+1} + T^{2q+1} - 1 - 2T}{2} - \sum_{i=1}^{q-1}\binom{2q+1}{2i}\xi_{2i}(T)\right), \tag{A.2}$$

$$\xi_{2q-1}(T) = \frac{1}{2q}\left(\frac{(T+1)^{2q} + T^{2q} - 1}{2} - \sum_{i=1}^{q-1}\binom{2q}{2i-1}\xi_{2i-1}(T)\right). \tag{A.3}$$

These three expressions show that $\xi_n(T)$ is a polynomial in T of order $n + 1$, $\xi_n(T) \in \mathcal{P}_{n+1}(T)$.

Next, use (A.1), or (A.2) and (A.3) as the definition of $\xi_n(T)$ as a polynomial in T of order $n + 1$ for $T \in \mathbb{Z}$; then, for all non-negative $s = 1, 2, 3, \ldots$, one has

$$\xi_n(-s) = (-1)^{n+1}\xi_n(s-1). \tag{A.4}$$

Proof. Equations (A.1) and (A.2) follow from Anderson (1971) Exercises 5 and 6, solving for $\xi_n(T)$ and $\xi_{2q}(T)$ using the fact that $\xi_0(T) := \sum_{t=1}^{T} 1 = T$. In order to prove (A.3), note that

$$C := \sum_{t=1}^{T}(t+1)^{2q} - \sum_{t=1}^{T}(t-1)^{2q} = \sum_{t=1}^{T}\left(\sum_{k=0}^{2q}\binom{2q}{k}t^k - \sum_{k=0}^{2q}\binom{2q}{k}t^k(-1)^{2q-k}\right)$$

$$= \sum_{t=1}^{T}\left(\sum_{k=0}^{2q}\binom{2q}{k}\left(1+(-1)^{k+1}\right)t^k\right) = 2\sum_{k=0}^{2q}\binom{2q}{k}1_{(k\ \text{odd})}\xi_k(T) = 2\sum_{i=1}^{q}\binom{2q}{2i-1}\xi_{2i-1}(T).$$

Hence one sees that C equals $\sum_{t=2}^{T+1} t^{2q} - \sum_{t=1}^{T-1} t^{2q} = (T+1)^{2q} + T^{2q} - 1$. This implies (A.3).

Formula (A.1) can be used to show that $\xi_n(T)$ is a polynomial in T of order $n+1$, by induction over n. Start from $n=1$, for which (A.1) gives $\xi_1(T) = ((T+1)^2 - (T+1))/2 = T(T+1)/2$. Next assume that $\xi_k(T)$ is a polynomial of order $k+1$ for $k = 1, 2, \ldots, n-1$, and observe that (A.1) for $\xi_n(T)$ begins with a polynomial of order $n+1$. This (or alternatively Eqs. (A.2) and (A.3) along similar lines) shows that $\xi_n(T)$ is a polynomial in T of order $n+1$.

Next property (A.4) is proved, separately for n even and n odd. In the case of n even, one has $(-1)^{n+1} = -1$ in (A.4), i.e. $\xi_n(-s) = -\xi_n(s-1)$. The proof is by induction over q for $n = 2q, q = 1, 2, \ldots$ using (A.2). First observe that property (A.4) holds for $n=2$ (i.e. $q=1$). In fact $\xi_2(T) = (2T+1)T(T+1)/6$, see e.g. Anderson (1971, p. 83), so that $\xi_2(-s) = (-2s+1)(-s)(-s+1)/6 = -(2s-1)s(s-1)/6$, and $\xi_2(s-1) = (2(s-1)+1)(s-1)s/6 = (2s-1)(s-1)s/6$.

Assume next that property (A.4) holds for $n = 2, \ldots, 2q-2$ and proceed to show that it holds also for $n = 2q$. Using (A.2) and the induction hypothesis, one finds

$$\xi_{2q}(s-1) = \frac{1}{2q+1}\left(\frac{s^{2q+1} + (s-1)^{2q+1} + 1 - 2s}{2} - \sum_{i=1}^{q-1}\binom{2q+1}{2i}\xi_{2i}(s-1)\right); \tag{A.5}$$

$$\xi_{2q}(-s) = \frac{1}{2q+1}\left(-\frac{(s-1)^{2q+1} + s^{2q+1} + 1 - 2s}{2} + \sum_{i=1}^{q-1}\binom{2q+1}{2i}\xi_k(s-1)\right). \tag{A.6}$$

From (A.5) and (A.6), one finds that $\xi_{2q}(-s) + \xi_{2q}(s-1) = 0$, hence proving (A.4) for $n = 2q$.

In the case of n odd, $n = 2q-1$, one has $(-1)^{n+1} = 1$ in (A.4), i.e. $\xi_n(-s) = \xi_n(s-1)$. The proof is by induction over $q = 1, 2, \ldots$, with $n = 2q-1$ using (A.3). First consider $n = q = 1$ where $\xi_1(T) = \frac{1}{2}T(T+1)$, for which $\xi_1(-s) = (-s)(-s+1)/2 = s(s-1)/2 = \xi_1(s-1)$, and hence (A.4) holds for $n = q = 1$.

Next assume that Eq. (A.4) holds for $n = 1, \ldots, 2q-3$ and proceed to prove it for $n = 2q-1$. From (A.3) one finds, using the induction hypothesis,

$$\xi_{2q-1}(-s) = \frac{1}{2q}\left(\frac{(s-1)^{2q} + s^{2q} - 1}{2} - \sum_{i=1}^{q-1}\binom{2q}{2i-1}\xi_{2i-1}(s-1)\right); \tag{A.7}$$

$$\xi_{2q-1}(s-1) = \frac{1}{2q}\left(\frac{s^{2q} + (s-1)^{2q} - 1}{2} - \sum_{i=1}^{q-1}\binom{2q}{2i-1}\xi_{2i-1}(s-1)\right). \tag{A.8}$$

From (A.7) and (A.8), one finds that $\xi_{2q-1}(-s) - \xi_{2q-1}(s-1) = 0$; this proves (A.4) for $n = 2q-1$ and completes the proof.

Further properties of Δ and $\mathcal{S}$ are stated in the following lemma. Let here $\varsigma_{n,t} := \mathcal{S}^n 1$ for $n = 0, 1, \ldots$, where 1 is the constant process. Note that this implies $\mathcal{S}^n u = \varsigma_{n,t} u$, when u is a constant vector.

Lemma A.2 (Properties of Δ, $\mathcal{S}$) *The operators Δ and $\mathcal{S}$ are linear, i.e. $\mathcal{O}(a_t + b_t) = \mathcal{O}a_t + \mathcal{O}b_t$ for $\mathcal{O} = \Delta, \mathcal{S}$. One has $\varsigma_{1,t} = t$ and $\varsigma_{n,t}$ for $n = 1, 2, \ldots$ is a polynomial in t of order n, $\varsigma_{n,t} \in \mathcal{P}_n(t)$. More in general, for $p_u(t) \in \mathcal{P}_u(t)$,*

$$\mathcal{S}^n p_u(t) \in \mathcal{P}_{u+n}(t) \tag{A.9}$$

$$\Delta^s p_u(t) \in \mathcal{P}_{u-s}(t), \quad 0 < s \le u \tag{A.10}$$

$$\Delta^{u+j} p_u(t) = 0, \qquad j > 0. \tag{A.11}$$

Moreover, for $t \in \mathbb{Z}$ one has

$$\mathcal{S}^s \Delta^h v_t = \mathcal{S}^{s-h} v_t - \sum_{n=s-h}^{s-1} \varsigma_{n,t} \Delta^{h-s+n} v_0, \quad 0<h \leq s. \tag{A.12}$$

Taking $h = s$ in (A.12), one finds as a special case

$$\mathcal{S}^s \Delta^s v_t = v_t - \sum_{n=0}^{s-1} \varsigma_{n,t} \Delta^n v_0. \tag{A.13}$$

Proof. Linearity of Δ, $\mathcal{S}$ follows by definition. Next consider $\varsigma_{n,t}$. For $n = 1$, using the definition (2.1) one finds that $\varsigma_{1,t} := \mathcal{S}1 = t$, a polynomial of order 1 in t, because $\mathcal{S}1 := 1_{(t\geq1)} \sum_{i=1}^{t} 1 - 1_{(t\leq-1)} \sum_{i=t+1}^{0} 1 = 1_{(t\geq1)}|t| - 1_{(t\leq-1)}|t| = \text{sign}(t)|t| = t$. Next proceed by induction over n, assuming that $\varsigma_{n-1,t} \in \mathcal{P}_{n-1}(t)$, with form $\varsigma_{n-1,t} = \sum_{i=0}^{n-1} a_i t^i$, and showing that $\varsigma_{n,t} \in \mathcal{P}_n(t)$, where $\varsigma_{n,t} = \mathcal{S}\varsigma_{n-1,t} = \sum_{i=0}^{n-1} a_i \mathcal{S}t^i$. The proof follows by showing that $\mathcal{S}t^i \in \mathcal{P}_{i+1}(t)$, where the order of $\varsigma_{n,t}$ comes from $\mathcal{S}t^{n-1} \in \mathcal{P}_n(t)$. For $t \geq 1, \mathcal{S}t^i = \sum_{k=1}^{t} k^i = \xi_i(t)$ which is in $\mathcal{P}_{i+1}(t)$ by Lemma A.1. For $t = 0$, one has $\mathcal{S}0 = 0$. Finally for $t \leq -1$ one has that $\mathcal{S}t^i := -\sum_{k=t+1}^{0} k^i = (-1)^{i+1} \sum_{h=1}^{|t|-1} h^i = (-1)^{i+1} \xi_i(|t|-1) = \xi_i(-|t|) = \xi_i(t)$ by (A.4). Hence $\mathcal{S}t^i = \xi_i(t) \in \mathcal{P}_{i+1}(t)$ for all values of $t \in \mathbb{Z}$. This completes the proof that $\varsigma_{n,t} := \mathcal{S}^n 1 \in \mathcal{P}_n(t)$. This derivation also shows (A.9). Direct application of the definitions imply (A.10) and (A.11). Next consider Eq. (A.12). For $0<h \leq s$ one finds:

$$\begin{aligned} \mathcal{S}^s \Delta^h v_t &= \mathcal{S}^{s-1}\left(\Delta^{h-1} v_t - \Delta^{h-1} v_0\right) = \mathcal{S}^{s-1} \Delta^{h-1} v_t - \varsigma_{s-1,t} \Delta^{h-1} v_0 = \\ &= \mathcal{S}^{s-2}\left(\Delta^{h-2} v_t - \Delta^{h-2} v_0\right) - \varsigma_{s-1,t} \Delta^{h-1} v_0 = \\ &= \mathcal{S}^{s-2} \Delta^{h-2} v_t - \sum_{n=s-2}^{s-1} \varsigma_{n,t} \Delta^{h-s+n} v_0 = \cdots = \mathcal{S}^{s-h} v_t - \sum_{n=s-h}^{s-1} \varsigma_{n,t} \Delta^{h-s+n} v_0; \end{aligned}$$

this proves (A.12) and (A.13).

Proof of Theorem 3.3. This is a restatement of Lemma 3.1, Theorems 3.4, 3.5 and Corollary 3.6 in Franchi and Paruolo (2016). Hence the proof is omitted.

Proof of Theorem 3.5. Proof of (3.14). Write the identity $F(z)F(z)^{-1} = I$ as the following linear system in the F_n, C_n matrices

$$\begin{gathered} F_0 C_0 = 0 \\ F_0 C_1 + F_1 C_0 = 0 \\ \vdots \\ F_0 C_{m-1} + \cdots + F_{m-1} C_0 = 0 \\ F_0 C_m + F_1 C_{m-1} + \cdots + F_m C_0 = I \\ F_0 C_{m+1} + F_1 C_m + \cdots + F_{m+1} C_0 = 0 \\ \vdots \end{gathered} \tag{A.14}$$

In the following, equations in system (A.14) are indexed according to the highest value of the subscript of C_n; for instance $F_0 C_0 = 0$ is referred to as equation 0. Remark that the identity appears in equation m, which is the order of the pole. Lemma 3.1 in Franchi and Paruolo (2016) shows that equation $n \geq j = 0, \ldots, m$ in system (A.14) implies

$$\alpha_j \beta_j' C_{h-j} = P_{a_{j\perp}} \sum_{k=1}^{h-j} F_{j+1,k} C_{h-j-k} + P_{a_{j\perp}} H_{j+1,h-j}, \quad h \geq j = 0, \ldots, m, \tag{A.15}$$

where α_j, β_j, a_j, and $F_{j+1,k}$ are as in Definition 3.1 and

$$H_{j+1,h-j} = \begin{cases} 0 & \text{for } h<m \\ -I & \text{for } h = m \\ H_{j,h-j+1} + F_{j,1} \sum_{i=0}^{j-1} \bar{\beta}_i \bar{\alpha}_i' H_{i+1,h-j} & \text{for } h > m \end{cases} \tag{A.16}$$

follows by applying definition (3.6). Pre-multiplying (A.15) by $\bar{\alpha}_j'$ and rearranging one thus finds

$$\beta_j' C_{h-j} - \bar{\alpha}_j' \sum_{k=1}^{h-j} F_{j+1,k} C_{h-j-k} = \bar{\alpha}_j' H_{j+1,h-j}, \quad h \geq j = 0, \ldots, m. \tag{A.17}$$

Next define $\gamma_j(z)' := \sum_{n=0}^{\infty} \gamma_{j,n}'(1-z)^n$, where $\gamma_{j,0}' := \beta_j'$ and $\gamma_{j,n}' := -\bar{\alpha}_j' F_{j+1,n}$ for $n \geq 1$, and consider $C(z) = \sum_{n=0}^{\infty} C_n(1-z)^n$ in (3.2). Writing $\gamma_j(z)'C(z) = \sum_{n=0}^{\infty} \zeta_{j,n}'(1-z)^n$, where $\zeta_{j,n}' := \sum_{k=0}^{n} \gamma_{j,k}' C_{n-k}$ is found by convolution, one has

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

where the last equality follows by setting $n = h - j$ in (A.17). Moreover, setting $n = h - j$ in (A.16) one finds

$$H_{j+1,n} = \begin{cases} 0 & \text{for } n < m-j \\ -I & \text{for } n = m-j \\ H_{j,n+1} + F_{j,1} \sum_{i=0}^{j-1} \bar{\beta}_i \bar{\alpha}_i' H_{i+1,n} & \text{for } n > m-j \end{cases} \tag{A.18}$$

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 for $j = 0, \ldots, m$ one has

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

That is, for $j = 0, \ldots, m$ one has $\gamma_j(z)'F(z)^{-1} = (1-z)^{-j} \phi_j(z)'$ and hence

$$\phi_j(z)'F(z) = (1-z)^j \gamma_j(z)'. \tag{A.19}$$

Next consider $\phi_j(z)' := \sum_{n=0}^{\infty} \phi_{j,n}'(1-z)^n$, where $\phi_{j,0}' := -\bar{\alpha}_j'$ and $\phi_{j,n}' := \bar{\alpha}_j' H_{j+1,m-j+n}$ for $n \geq 1$, and $F(z) = \sum_{n=0}^{\infty} F_n(1-z)^n$ in (3.1). Writing $\phi_j(z)'F(z) = \sum_{n=0}^{\infty} \zeta_{j,n}'(1-z)^n$, where $\zeta_{j,n}' := \sum_{k=0}^{n} \phi_{j,k}' F_{n-k}$ is found by convolution, from (A.19) one has

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

Defining $\Phi(z), \Gamma(z)$ and $\Lambda(z)$ as in Definition 3.4, from (A.19) one finds (3.14), where $\Phi(1) = -(\bar{\alpha}_0, \ldots, \bar{\alpha}_m)'$ and $\Gamma(1) = (\beta_0, \ldots, \beta_m)'$ are nonsingular. The proof of (3.15) starts by transposing (A.14) and then proceeds along the same lines of that of (3.14). It is thus omitted.

Proof of Theorem 4.1. Write (3.1) as $F(z) = \sum_{n=0}^{d-1} F_n(1-z)^n + (1-z)^d F_d(z)$, where $F_d(z) := \sum_{n=d}^{\infty} F_n(1-z)^{n-d}$ is analytic for all $|z| < 1 + \delta, \delta > 0$. Hence the coefficients of the expansion $F_d(z) = \sum_{n=0}^{\infty} F_n^* z^n$ are geometrically decreasing and the process $Y_t := F_d(L)\varepsilon_t$ is stationary. Substituting in (4.1) one has

$$\Delta^d X_t = \sum_{j=0}^{d-1} F_j \Delta^j \varepsilon_t + \Delta^d Y_t. \tag{A.20}$$

Pre-multiply both sides of (A.20) by $\mathcal{S}^d$; by (A.13) one has $\mathcal{S}^d \Delta^d X_t = X_t - \nu_{x,t}$, $\nu_{x,t} := \sum_{n=0}^{d-1} \gamma_{n,t} \Delta^n X_0$, and $\mathcal{S}^d \Delta^d Y_t = Y_t - \nu_{y,t}$, $\nu_{y,t} := \sum_{n=0}^{d-1} \gamma_{n,t} \Delta^n Y_0$. Moreover, by (A.12) one has

$$\mathcal{S}^d \Delta^j \varepsilon_t = \mathcal{S}^{d-j} \varepsilon_t - \sum_{n=d-j}^{d-1} \varsigma_{n,t} \Delta^{j-d+n} \varepsilon_0, \quad 0 < j \leq d,$$

and hence $\sum_{j=0}^{d-1} F_j \mathcal{S}^d \Delta^j \varepsilon_t = \sum_{j=0}^{d-1} F_j S_{d-j,t} - \nu_{\varepsilon,t}$, $\nu_{\varepsilon,t} := \sum_{j=0}^{d-1} F_j \sum_{n=d-j}^{d-1} \varsigma_{n,t} \Delta^{j-d+n} \varepsilon_0$. Hence the solution of (A.20) is

$$X_t = \sum_{j=0}^{d-1} F_j S_{d-j,t} + Y_t + \nu_{d-1,t}, \quad \nu_{d-1,t} := \nu_{x,t} - \nu_{y,t} - \nu_{\varepsilon,t},$$

where $v_{d-1,t} =: \sum_{n=0}^{d-1} v_n t^n$ is a polynomial of order $d-1$ in t whose coefficients depend on initial values, see the definitions of $v_{x,t}, v_{y,t}$ and $v_{\varepsilon,t}$. This completes the proof of the first part of the statement. Pre-multiplying (4.1) by $\phi_j(L)'$, see (3.9), and using $\phi_j(L)'F(L) = \Delta^j \gamma_j(L)'$, see (3.16), one finds

$$\Delta^d \phi_j(L)' X_t = \Delta^j \gamma_j(L)' \varepsilon_t, \quad j = 0, ..., m. \tag{A.21}$$

Because $\gamma_j(1)' = \beta_j'$ has full row rank, this shows that for $j = 0, ..., m$ one has $\Delta^d \phi_j(L)' X_t \sim I_{nc}(-j)$, i.e. $\phi_j(L)' X_t \sim I_{nc}(d-j)$, see Definition 2.2. Next it is shown that the same holds for the truncated version $\phi_j^{(j-1)}(L)' X_t \sim I_{nc}(d-j), j = 1, ..., m$. Substituting $\phi_j(z)' = \phi_j^{(j-1)}(z)' + (1-z)^j \phi_j^\star(z)'$ in (A.21) and rearranging one finds

$$\Delta^d \phi_j^{(j-1)}(L)' X_t = \Delta^j \left(\gamma_j(L)' \varepsilon_t - \phi_j^\star(L)' \Delta^d X_t \right), \quad j = 1, ..., m,$$

and thus, substituting $\Delta^d X_t = F(L)\varepsilon_t$,

$$\Delta^d \phi_j^{(j-1)}(L)' X_t = \Delta^j \left(\gamma_j(L)' - \phi_j^\star(L)' F(L) \right) \varepsilon_t, \quad j = 1, ..., m. \tag{A.22}$$

Using $\gamma_j(1)' = \beta_j', \phi_j^\star(1)' = \phi_{j,j}' = \bar{\alpha}_j' H_{j+1,m}$ and $F(1) = \alpha_0 \beta_0'$, one finds $\gamma_j(1)' - \phi_j^\star(1)' F(1) = \beta_j' - \bar{\alpha}_j' H_{j+1,m} \alpha_0 \beta_0'$. Because $(\gamma_j(1)' - \phi_j^\star(1)' F(1)) \bar{\beta}_j = I_{r_j}, \gamma_j(1)' - \phi_j^\star(1)' F(1)$ has full row rank. This shows that for $j = 1, ..., m$ one has $\Delta^d \phi_j^{(j-1)}(L)' X_t \sim I_{nc}(-j)$, i.e. $\phi_j^{(j-1)}(L)' X_t \sim I_{nc}(d-j)$, and completes the proof of the second part of the statement. Grouping $\Delta^{d-j} \phi_j^{(j-1)}(L)' X_t \sim I_{nc}(0)$ together and using $\Lambda(z)$ defined in (3.13), one finds (4.4), where $\Phi_c(1) = -(\bar{\alpha}_0, ..., \bar{\alpha}_m)'$ is square and nonsingular. This completes the proof of the statement.

Proof. Proof of Corollary 4.2. Use (4.4) in Theorem 4.1.

Proof. Proof of Theorem 4.3. By Theorem 3.3, $F(z)^{-1} = (1-z)^{-m} C(z)$ with $C(1) \neq 0$ if and only if the pole(m) condition on $F(z)$ hold, i.e. one has $\Delta^m X_t = C(L)\varepsilon_t$ with $C(1) \neq 0$, which shows that $X_t \sim I(d)$, $d = m$. Proceeding along the lines of the proof of Theorem 4.1 one finds the statement.

Proof. Proof of Corollary 4.4. Setting $m = 1$ in Theorem 3.3 one has $C_0 = -\bar{\beta}_1 \bar{\alpha}_1'$ and setting $m = 1$ in Theorem 4.3 one has

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

and hence the statement.

Proof of Corollary 4.5. Setting $m = 2$ in Theorem 3.3 one has $C_0 = -\bar{\beta}_2 \bar{\alpha}_2'$ and $C_1 = H_1 + K_1 C_0$, where

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

Definition (3.6) 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

$$C_1 = -\bar{b} \begin{pmatrix} 0 & 0 & \bar{\alpha}_0' F_{1,1} \bar{\beta}_2 \\ 0 & I_{r_1} & \bar{\alpha}_1' F_{2,1} \bar{\beta}_2 \\ \bar{\alpha}_2' F_{1,1} \bar{\beta}_0 & \bar{\alpha}_2' F_{2,1} \bar{\beta}_1 & \bar{\alpha}_2' F_{3,1} \bar{\beta}_2 \end{pmatrix} \bar{a}'.$$

Setting $m = 2$ in Theorem 4.3 one has

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

and hence the statement.

Proof of Corollary 4.6. Use (4.8) in Theorem 4.3.

Proof of Theorem 4.8. First note that, because $\Delta^{u+s}t^u = 0$ see (A.11), one can omit the superscript (u) in (4.15) when substituting into μ_t, since $\mu_t = a(L)t^u = \psi^{(u)}_{j:m}(L)\varphi t^u = \psi_{j:m}(L)\varphi t^u$. Similarly for (4.16), $\mu_t = a(L)t^u = \pi^{(u)}_{0:j}(L)\varphi t^u = \pi_{0:j}(L)\varphi t^u$.

Second, recall that from (3.17) one has the right root functions $\psi_j(z) := \sum_{n=0}^{\infty}\psi_{j,n}(1-z)^n$ and $\pi_j(z) := \sum_{n=0}^{\infty}\pi_{j,n}(1-z)^n$, where $\psi_{j,0} := -\bar{\beta}_j, \pi_{j,0} := \alpha_j$ and $\psi_{j,n} := H_{j+1,m-j+n}\bar{\beta}_j, \pi_{j,n} := -F_{j+1,n}\beta_j, n = 1,2,\ldots,$ such that (3.17) holds. Combining $\Phi(z)F(z) = \Lambda(z)\Gamma(z)$ and $F(z)\Psi(z) = \Pi(z)\Lambda(z)$, see (3.14) and (3.15), one finds $\Phi(z)\Pi(z)\Lambda(z) = \Phi(z)F(z)\Psi(z) = \Lambda(z)\Gamma(z)\Psi(z)$. Hence, recalling that $F(z)^{-1} = (1-z)^{-m}C(z)$,

$$\Phi(z)\Pi(z)\Lambda(z) = \Lambda(z)\Gamma(z)\Psi(z), \tag{A.23}$$

$$C(z)\Pi(z) = (1-z)^m\Psi(z)\Lambda(z)^{-1}, \tag{A.24}$$

$$\Gamma(z)C(z) = (1-z)^m\Lambda(z)^{-1}\Phi(z). \tag{A.25}$$

MA, Sufficiency. First consider (4.11) under (4.15). One has, $\Delta^d X_t = F(L)\varepsilon_t + F(L)a(L)t^u$. Thanks to $F(z)\Psi(z) = \Pi(z)\Lambda(z)$, one has $F(L)a(L)t^u = F(L)\psi_{j:m}(L)\varphi t^u = \Delta^j\pi^\star_j(L)\varphi t^u = \pi^\star_j(L)\varphi p_{u-j}(t)$, i.e.

$$\Delta^d X_t = F(L)\varepsilon_t + \pi^\star_j(L)\varphi p_{u-j}(t), \tag{A.26}$$

where $\pi^\star_j(z) := (1-z)^{-j}\Pi(z)\Lambda(z)(0,I)' = \pi_{j:m}(z)\text{blkdiag}(I_{r_j}, I_{r_{j+1}}(1-z), \ldots, I_{r_m}(1-z)^{m-j})$, and $p_{u-j}(t) := \Delta^j t^u \in \mathcal{P}_{u-j}(t)$ by (A.10). Note that $\pi^\star_j(1)\varphi = \alpha_j\varphi_j \neq 0$ because α_j has full column rank and $\varphi_j \neq 0$ in (4.15). Applying $\mathcal{S}^d$ on both sides of (A.26), one finds that $X_t \sim I(d)$ involves the term $\pi^\star_j(L)\varphi\mathcal{S}^d p_{u-j}(t) \in \mathcal{P}_{d+u-j}(t)$ by (A.9). This proves sufficiency of (4.15) for *i*.1).

Next pre-multiply (4.11) by $\Phi(L)$, see (3.9); using $\Phi(L)F(L) = \Lambda(L)\Gamma(L)$, see (3.16), one finds

$$\Delta^d\Phi(L)'X_t = \Lambda(L)\Gamma(L)\varepsilon_t + \Phi(L)F(L)\Psi(L)(0,\varphi')'t^u = \Lambda(L)\Gamma(L)\varepsilon_t + \Lambda(L)\Gamma(L)\Psi(L)(0,\varphi')'t^u,$$

where use is made of (A.23). Taking the hth block of rows, one finds that

$$\Delta^d\phi_h(L)'X_t = \Delta^h\gamma_h(L)'\varepsilon_t + \Delta^h\gamma_h(L)'\psi_{j:m}(L)\varphi t^u = \Delta^h\gamma_h(L)'\varepsilon_t + \gamma_h(L)'\psi_{j:m}(L)\varphi p_{u-h}(t),$$

where use is made of (A.10) and $\gamma_j(1)'\psi_j(1) = -\beta_j'\bar{\beta}_j = -I_{r_j}$, which implies that $\gamma_h(1)'\psi_{j:m}(1)$ contains a square nonsingular block, because $j \le h \le m$. Given that $\varphi_j \neq 0$ this implies that $s_j = d+u-j$, while for $j<h \le m$ the order of the trend s_h is at most $d+u-h$. This proves sufficiency of (4.15) for *i*.2).

MA, Necessity. Assume now that *i*.1) holds. This implies that in $\Delta^d X_t = F(L)\varepsilon_t + F(L)a(L)t^u$ one has $F(L)a(L) = \Delta^j b(L)$ for some $b(L)$ with $b(1) \neq 0$, i.e. that $a(L)$ is a right root function of $F(L)$ of order j. This implies that $a(L)$ can be expressed as linear combinations of the right root functions $\psi_s(L)$ in (3.16) of order equal to j or higher, i.e. that $a(z) = \psi^{(u)}_{j:m}(z)\varphi$, where for the order to be j, the coefficient of $\psi^{(u)}_j(z)$ needs to be nonzero. This implies (4.15). A similar derivation applies to show necessity assuming *i*.2) holds.

AR, Sufficiency. Consider (4.12) under (4.16), where $F(L)X_t = \varepsilon_t + \mu_t$, which implies

$$\Delta^m X_t = C(L)\varepsilon_t + C(L)\mu_t. \tag{A.27}$$

Because of (A.24), one has

$$C(L)\mu_t = C(L)a(L)t^u = \Delta^m\Psi(L)\Lambda^{-1}(L)(\varphi',0)'t^u = \Delta^{m-j}\psi^\star_j(L)\varphi t^u = \psi^\star_j(z)\varphi p_{u-m+j}(t),$$

where

$$\psi^\star_j(z) := (1-z)^j\Psi(z)\Lambda^{-1}(z)(I,0)' = \pi_{0:j}(z)\text{blkdiag}\left(I_{r_0}(1-z)^j, I_{r_1}(1-z)^{j-1}, \ldots, I_{r_j}\right)$$

so that

$$\Delta^m X_t = C(L)\varepsilon_t + \psi^\star_j(z)\varphi p_{u-m+j}(t),$$

where $\psi^\star_j(1)\varphi = -\bar{\beta}_j\varphi_j \neq 0$ because β_j has full column rank and $\varphi_j \neq 0$. Applying $\mathcal{S}^m$ to both sides of the equation, one proves *ii*.1). To show *ii*.2), pre-multiply (A.27) by $\Gamma(L)$ and use (A.25), (A.23) to obtain

$$\begin{aligned}\Gamma(L)\Delta^m X_t &= \Gamma(L)C(L)\varepsilon_t + \Gamma(L)C(L)\Pi(L)(\varphi',0)'t^u\\ &= (1-L)^m\Lambda(L)^{-1}\Phi(L)\varepsilon_t + (1-L)^m\Gamma(L)\Psi(L)\Lambda(L)^{-1}(\varphi',0)'t^u\\ &= (1-L)^m\Lambda(L)^{-1}\Phi(L)\varepsilon_t + (1-L)^m\Lambda(L)^{-1}\Phi(L)\Pi(L)(\varphi',0)'t^u.\end{aligned}$$

Taking the hth block of rows, one finds that

$$\Delta^m\gamma_h(L)X_t = \Delta^{m-h}\phi_h(L)'\varepsilon_t + \Delta^{m-h}\phi_h(L)'\pi_{0:j}(L)\varphi t^u = \Delta^{m-h}\phi_h(L)'\varepsilon_t + \phi_h(L)'\pi_{0:j}(L)\varphi p_{u-m+h}(t),$$

where $\phi_h(1)'\pi_j(1) = -\bar{\alpha}_j'\alpha_j = -I_{r_j}$. This implies that $\phi_h(1)'\pi_{0:j}(1)$ contains a square nonsingular block, because $0 \le h \le j$. Because $\varphi_j \neq 0$ this implies that $s_j = u+j$, while for $0 \le h<j$ the order of the trend s_h is at most $u + h$. This proves sufficiency of (4.15) for *ii*.2).

AR, Necessity. Assume now that *ii*.1) holds. This implies that in $\Delta^m X_t = C(L)\varepsilon_t + C(L)a(L)t^u$ one has $C(L)a(L) = \Delta^{m-j}b(L)$ for some $b(L)$ with $b(1) \neq 0$, i.e. that $a(L)$ is a right root function of $C(L)$ of order m–j. This implies that $a(L)$ can be expressed as linear combinations of the right root functions $\pi_s(L)$ in (3.17) of order equal to j or lower, i.e. that $a(z) = \pi_{0:j}^{(u)}(z)\varphi$, where for the order to be j, the coefficient of $\pi_j^{(u)}(z)$ needs to be non-zero. This implies (4.16). A similar derivation applies to show necessity assuming *ii*.2) holds.

Proof of Corollary 4.10. Direct consequence of Theorem 3.5 and the definition of Jordan pairs in Gohberg et al. (1993).

Acknowledgments

Comments and suggestions on a previous version from the editor and two anonymous referees are gratefully acknowledged. The work of the first author was partially supported by MIUR PRIN grant 2010J3LZEN.

References

Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. New York: Wiley.

Archontakis, F. (1998). An alternative proof of Granger's Representation Theorem for I(1) systems through Jordan matrices. *Journal of the Italian Statistical Society* 7(2):111–127.

Avrachenkov, K., Haviv, M., Howlett, P. (2001). Inversion of analytic matrix functions that are singular at the origin. *SIAM Journal on Matrix Analysis and Applications* 22(4):1175–1189.

Bauer, D., Wagner, M. (2012). A state space canonical form for unit root processes. *Econometric Theory* 28(06): 1313–1349.

Beare, B., Seo, J., Seo, W. (2017). Cointegrated linear processes in hilbert space. *Journal of Time Series Analysis* 38(6):1010–1027.

Beare, B., Seo, W. (2018). Representation of I(1) and I(2) autoregressive hilbertian processes. *Arxiv e-Print, arXiv* 1701:08149v2.

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

Brockwell, P., Davis, R. (1991). *Time Series: Theory and Methods*, 2nd ed. New York: Springer-Verlag.

Chang, Y., Kim, C., Park, J. (2016). Nonstationarity in time series of state densities. *Journal of Econometrics* 192(1): 152–167.

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

Engle, R., Granger, C. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica* 55(2):251–276.

Engle, R., Kozicki, S. (1993). Testing for common features. *Journal of Business and Economic Statistics* 11(4): 369–395.

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

Faliva, M., Zoia, M. (2009). *Topics in Dynamic Model Analysis*. Berlin: Springer-Verlag.

Franchi, M. (2007). The integration order of vector autoregressive processes. *Econometric Theory* 23:546–553.

Franchi, M. (2010). A representation theory for polynomial cofractionality in vector autoregressive models. *Econometric Theory* 26(04):1201–1217.

Franchi, M., Paruolo, P. (2011a). A characterization of vector autoregressive processes with common cyclical features. *Journal of Econometrics* 163(1):105–117.

Franchi, M., Paruolo, P. (2011b). Inversion of regular analytic matrix functions: Local smith form and subspace duality. *Linear Algebra and Its Applications* 435(11):2896–2912.

Franchi, M., Paruolo, P. (2016). Inverting a matrix function around a singularity via local rank factorization. *SIAM Journal on Matrix Analysis and Applications* 37(2):774–797.

Franchi, M., Paruolo, P. (2017). Cointegration in functional autoregressive processes. *Arxiv e-Print, arXiv* 1712: 07522.

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

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

Granger, C. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics* 16(1):121–130.

Granger, C. W. J. (2004). Time series analysis, cointegration, and applications. In T. Frängsmyr (Ed.), *The Nobel Prizes 2003*. Stockholm: Almqvist and Wiksell International, pp. 360–366. Available at https://www.nobelprize.org/prizes/economics/2003/granger/lecture/.

Greene, R., Krantz, S. (1997). *Function Theory of One Complex Variable*. New York: John Wiley & Sons.

Gregoir, S. (1999). Multivariate time series with various hidden unit roots, part I. *Econometric Theory* 15(4): 435–468.

Gregoir, S., Laroque, G. (1994). Polynomial cointegration: Estimation and test. *Journal of Econometrics* 63(1): 183–214.

Haldrup, N., Salmon, M. (1998). Representation of I(2) cointegrated systems using the Smith-McMillan form. *Journal of Econometrics* 84(2):303–325.

Hannan, E., Deistler, M. (1988). *The Statistical Theory of Linear Systems*. New York: John Wiley & Sons.

Hansen, P. (2005). Granger's Representation Theorem: A closed-form expression for I(1) processes. *The Econometrics Journal* 8(1):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. *The Journal of the Australian Mathematical Society. Series B. Applied MathematicsMathematics* 23(04):357–382.

Howlett, P., Avrachenkov, K., Pearce, C., Ejov, V. (2009). Inversion of analytically perturbed linear operators that are singular at the origin. *Journal of Mathematical Analysis and Applications* 353(1):68–84.

Hu, B., Park, J. (2016). *Econometric Analysis of Functional Dynamics in the Presence of Persistence*. Mimeo: Indiana University.

Hylleberg, S., Engle, R. F., Granger, C. W. J., Yoo, B. S. (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(2-3):231–254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* 59(6):1551–1580.

Johansen, S. (1992). A representation of vector autoregressive processes integrated of order 2. *Econometric Theory* 8(02):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*. 2nd printing. Oxford: Oxford University Press.

Johansen, S. (2005). Interpretation of cointegrating coefficients in the cointegrated vector autoregressive model. *Oxford Bulletin of Economics and Statistics* 67(1):93–104.

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

Johansen, S. (2008b). Representation of cointegrated autoregressive process with application to fractional process. *Econometric Reviews* 28(1-3):121–145.

Johansen, S., Nielsen, M. (2010). Likelihood inference for a nonstationary fractional autoregressive model. *Journal of Econometrics* 158(1):51–66.

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

Johansen, S., Schaumburg, E. (1999). Likelihood analysis of seasonal cointegration. *Journal of Econometrics* 88(2): 301–339.

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

Lancaster, P. (1966). *Lambda Matrices and Vibrating Systems*. Oxford: Pergamon Press.

Neusser, K. (2000). An algebraic interpretation of cointegration. *Economics Letters* 67(3):273–281.

Omtzigt, P., Paruolo, P. (2005). Impact factors. *Journal of Econometrics* 128(1):31–68.

Paruolo, P. (1997a). Asymptotic inference on the moving average impact matrix in cointegrated I(1) VAR systems. *Econometric Theory* 13(01):79–118.

Paruolo, P. (1997b). Standard errors for the long run variance matrix. *Econometric Theory* Problem: 13(02): 305–306. Solution: 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(2):283–306.

Phillips, P., Hansen, B. (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(1-2):21–56.

Schumacher, J. (1986). Residue formulas far meromorphic matrices. In C. Byrnes and A. Lindquist (Eds.), *Computational and Combinatorial Methods in System Theory*. Amsterdam: North-Holland, pp. 97–111.

Schumacher, J. (1991). System-theoretic trends in econometrics. In A. Antoulas (Ed.), *Mathematical System Theory: The Influence of R.E. Kalman*. Berlin: Springer, pp. 559–577.

Sims, C., Stock, J., Watson, M. (1990). Inference in linear time series models with some unit roots. *Econometrica* 58(1):113–144.

Stock, J., Watson, M. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61(4):783–820.

Vahid, F., Engle, R. (1993). Common trends and common cycles. *Journal of Applied Econometrics* 8:341–360.

Yoo, B. S. (1986). Multi-cointegrated time series and generalized error correction models. Economics Department, Working Paper. San Diego: University of California.