Econometric Theory, **36**, 2020, 773–802.
doi:10.1017/S0266466619000276

REPRESENTATION OF I(1) AND I(2) AUTOREGRESSIVE HILBERTIAN PROCESSES

BRENDAN K. BEARE
University of Sydney
WON-KI SEO
Queen's University

We develop versions of the Granger–Johansen representation theorems for I(1) and I(2) vector autoregressive processes that apply to processes taking values in an arbitrary complex separable Hilbert space. This more general setting is of central relevance for statistical applications involving functional time series. An I(1) or I(2) solution to an autoregressive law of motion is obtained when the inverse of the autoregressive operator pencil has a pole of first or second order at one. We obtain a range of necessary and sufficient conditions for such a pole to be of first or second order. Cointegrating and attractor subspaces are characterized in terms of the behavior of the autoregressive operator pencil in a neighborhood of one.

1. INTRODUCTION

Results on the existence and representation of integrated solutions to vector autoregressive laws of motion are among the most important and subtle contributions of econometricians to time series analysis, yet also among the most widely misunderstood. The best known such result is the so-called Granger representation theorem, which first appeared in an unpublished UC San Diego working paper of Granger (1983). In this working paper, Granger, having recently introduced the concept of cointegration (Granger, 1981) sought to connect statistical models of time series based on linear process representations to regression-based models involving equilibrium correction terms, which had appeared in work by Sargan (1964) and Davidson, Hendry, Srba, and Yeo (1978). The main result of Granger (1983) first emerged in published form in Granger (1986) without proof, but more prominently in the widely cited *Econometrica* article by Engle and Granger (1987), where it is labeled the "Granger representation theorem,"

We thank Massimo Franchi, Peter Phillips and seminar participants at the Einaudi Institute for Economics and Finance, UC San Diego, the Université libre de Bruxelles, and the 2019 NBER-NSF Time Series Conference in Hong Kong for helpful discussions. Beare also thanks Phil Roberts for his feedback and encouragement. An earlier version of this article titled "Representation of I(1) autoregressive Hilbertian processes" was posted on the arXiv.org preprint repository in January 2017. Address correspondence to Brendan K. Beare, School of Economics, University of Sydney, Sydney, Australia; e-mail: brendan.beare@sydney.edu.au.

© Cambridge University Press 2019

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

with the exclusion of the first author presumably due to the article having resulted from the merger of previous independent contributions.

The proof of the Granger representation theorem in Engle and Granger (1987) is incorrect. Moreover, the error can be traced back to the original working paper of Granger (1983). A counterexample to Lemma A1 of Engle and Granger (1987), which is also Theorem 1 of Granger (1983), may be found buried in a footnote of Johansen (2009). Johansen was familiar with Granger's work on representation theory at an early stage, visiting UC San Diego and authoring a closely related Johns Hopkins working paper in 1985 that was eventually published as Johansen (1988). At around the same time, the doctoral thesis of Yoo (1987) at UC San Diego established the connection to Smith–McMillan forms. Johansen (1991) provided what appears to be the first correct statement and proof of a modified version of the Granger representation theorem, which we will call the Granger–Johansen representation theorem. This contribution did not merely correct a technical error of Granger; it reoriented attention toward a central issue: when does a given vector autoregressive law of motion admit an I(1) solution? The answer to this question is given by the Johansen I(1) condition, which is a necessary and sufficient condition on the autoregressive polynomial and its first derivative at one for a vector autoregressive law of motion to admit an I(1) solution. Johansen (1992) subsequently obtained analogous results for I(2) processes, and La Cour (1998) provided some results for higher order integrated processes.

An early contribution of Schumacher (1991) contained a striking observation on the Johansen I(1) condition: it corresponds to a necessary and sufficient condition for the inverse of a matrix pencil to have a simple pole at a given point in the complex plane. Various authors later exploited this insight, including Faliva and Zoia (2002, 2009, 2011), Franchi (2007), Johansen (2009), and Franchi and Paruolo (2011, 2016). A nice aspect of the connection to analytic function theory is that it extends naturally to the development of I(d) representation theory with integer-valued $d \geq 2$: just as the Johansen I(1) condition can be reformulated as a necessary and sufficient condition for a simple pole, analogous I(d) conditions can be reformulated as necessary and sufficient conditions for poles of order d. Franchi and Paruolo (2019b) have recently taken precisely this approach to develop a general I(d) representation theory. The introduction to their article contains a detailed discussion of the history of research on the Granger–Johansen representation theorem.

Parallel to the development of representation theorems for cointegrated systems in the 1980s and early 1990s was the development of asymptotic distribution theory for the statistical estimation of such systems, obtained by applying central limit theory on function spaces and associated results. This research was led by Phillips and his students at Yale; see, in particular, Phillips and Durlauf (1986), Phillips (1986, 1988, 1991), Phillips and Park (1988), Park and Phillips (1988, 1989), and Phillips and Hansen (1990), among many other contributions. Johansen (1991) used limit theorems developed in this body of work, applicable to

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

integrated systems exhibiting general serial dependence, to derive the asymptotic distribution of the maximum likelihood estimator for I(1) Gaussian vector autoregressive systems. Complementary work by Chan and Wei (1987, 1988) and Kurtz and Protter (1991) on asymptotic theory for regression with integrated processes has also been influential.

In this article, we provide representation theorems for I(1) and I(2) vector autoregressive processes taking values in an arbitrary complex separable Hilbert space. This more general setting is of central relevance for statistical applications involving functional time series (Hörmann and Kokoszka, 2012), and was first studied by Chang, Kim, and Park (2016) in the case of I(1) probability density-valued time series; see also Beare (2017) and Seo and Beare (2019). Our results here build on those we obtained in an earlier article with J. Seo (Beare, Seo, and Seo, 2017) establishing a representation theorem for the I(1) case. They differ from our earlier results in their explicit use of analytic function theory. In Theorems 3.1 and 4.1, our I(1) and I(2) representation theorems, we directly assume that the inverse of the autoregressive operator pencil has a pole of first or second order at one. We elaborate upon the meaning of these pole conditions in Theorems 3.2 and 4.2, which provide necessary and sufficient conditions to have a pole of first or second order. These results also provide explicit formulas for the coefficients in the principal part of the corresponding Laurent series.

Our article supersedes an earlier manuscript posted on the arXiv.org preprint repository in January 2017 (Beare and Seo, 2017) that dealt only with the I(1) case. During its preparation, several working papers have emerged that deliver related results. In particular, Franchi and Paruolo (2019a) study I(d) solutions to autoregressive laws of motion in complex separable Hilbert space, for integer-valued $d \geq 1$. Their necessary and sufficient condition for an I(d) solution involves an orthogonal direct sum decomposition of the Hilbert space into d closed subspaces. This contrasts with the direct sum conditions given by Beare, Seo, and Seo (2017) for the I(1) case, and here for the I(1) and I(2) cases, which involve nonorthogonal direct sums. We also provide a range of alternative formulations of our necessary and sufficient conditions, some of which may be easier to verify than others. Also relevant is recent work by Hu and Park (2016), who established an alternative I(d) condition for first-order autoregressive Hilbertian processes: the restriction of the autoregressive operator to the image of the Riesz projection associated with its unit eigenvalue differs from the identity by an operator nilpotent of degree d. Finally, Chang, Hu, and Park (2016) have developed I(1) representation theory for autoregressive Hilbertian processes under the assumption that the impact operator in the error correction representation is compact. Under this condition, the dimension of the cointegrating space must be finite, which contrasts with the setting of this article and the others cited in this paragraph, where the codimension of the cointegrating space must be finite. Finite codimensionality of the cointegrating space implies that the I(1) stochastic trend in the Beveridge–Nelson representation of our cointegrated process is confined to a finite dimensional linear subspace. It is a consequence of a compactness con-

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

dition we impose on the autoregressive operators. Franchi and Paruolo (2019a) have observed that finite codimensionality of the cointegrating space holds more generally if the autoregressive operator pencil has an eigenvalue of finite type at one.

We structure the remainder of the article as follows. Section 2 sets the scene with notation and essential mathematics. Our results on I(1) and I(2) processes are contained in Sections 3 and 4, respectively. We provide a brief discussion of directions for future research in Section 5. Appendix A.1 contains background material on the spectral properties of operator-valued functions, including a statement of the analytic Fredholm theorem, which is a key input to our results. The proofs of our results are collected in Appendix A.2.

2. PRELIMINARIES

The setting for our analysis is a separable complex Hilbert space $\mathcal{H}$ with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. If $\mathcal{H}'$ is another such space, we let $\mathcal{L}_{\mathcal{H},\mathcal{H}'}$ denote the Banach space of continuous linear operators from $\mathcal{H}$ to $\mathcal{H}'$ equipped with the operator norm. We are mostly concerned with the case $\mathcal{H} = \mathcal{H}'$ and write $\mathcal{L}_{\mathcal{H}}$ in place of $\mathcal{L}_{\mathcal{H},\mathcal{H}}$. To each operator $A \in \mathcal{L}_{\mathcal{H}}$, we associate two linear subspaces of $\mathcal{H}$: the kernel and range of A, given by

$$\ker A = \{x \in \mathcal{H} : A(x) = 0\}, \quad \text{ran } A = \{A(x) : x \in \mathcal{H}\}.$$

We let $\mathrm{I} \in \mathcal{L}_{\mathcal{H}}$ denote the identity map on $\mathcal{H}$.

A central concern of our analysis will be the decomposability of $\mathcal{H}$ into sums of certain linear subspaces of $\mathcal{H}$. Given linear subspaces V and W of $\mathcal{H}$, we write $V + W$ for the linear subspace of all $x \in \mathcal{H}$ such that $x = v + w$ for some $v \in V$ and $w \in W$. When V and W are linear subspaces of $\mathcal{H}$ with $V \cap W = \{0\}$, we may instead write $V \oplus W$ for their sum, and call it a direct sum. When we write $\mathcal{H} = V \oplus W$, we are asserting that $\mathcal{H} = V + W$ and that $V \cap W = \{0\}$. In this case, any $x \in \mathcal{H}$ may be uniquely decomposed as $x = v + w$ with $v \in V$ and $w \in W$.

Orthogonal complements and projections play a key role in our analysis. Given a linear subspace V of $\mathcal{H}$, we define its orthogonal complement by

$$V^{\perp} = \{x \in \mathcal{H} : \langle x, v \rangle = 0 \text{ for all } v \in V\}.$$

The orthogonal complement to a linear subspace of $\mathcal{H}$ is always a closed linear subspace of $\mathcal{H}$. Given a closed linear subspace V of $\mathcal{H}$, it is always the case that $\mathcal{H} = V \oplus V^{\perp}$. Thus, any $x \in \mathcal{H}$ may be uniquely decomposed as $x = v + v'$ with $v \in V$ and $v' \in V^{\perp}$. We denote by $\mathrm{P}_V \in \mathcal{L}_{\mathcal{H}}$ the orthogonal projection on V, which maps a point $x = v + v'$ to v.

Our main results concern the representation of time series taking values in $\mathcal{H}$, but only the most basic understanding of probability on $\mathcal{H}$ is required. As in Beare, Seo, and Seo (2017), we let $L^2_{\mathcal{H}}$ denote the Banach space of random elements

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

Z of $\mathcal{H}$ (identifying random elements that are equal with probability one) that satisfy $E\|Z\|^2 < \infty$ and $EZ = 0$, equipped with the norm $\|Z\|_{L^2_{\mathcal{H}}} = (E\|Z\|^2)^{1/2}$. Refer to that article for the definition of EZ and of the covariance operator of an element of $L^2_{\mathcal{H}}$. For further details, the monograph of Bosq (2000) provides a comprehensive treatment of linear processes taking values in a real Hilbert or Banach space. A complex Hilbert space setting was studied more recently by Cerovecki and Hörmann (2017).

3. I(1) AUTOREGRESSIVE HILBERTIAN PROCESSES

In this section, we state our results for I(1) autoregressive processes. Let $p \in \mathbb{N}$, and consider the following AR(p) law of motion in $\mathcal{H}$:

$$X_t = \sum_{j=1}^{p} \Phi_j(X_{t-j}) + \varepsilon_t. \tag{3.1}$$

Here, the X_t's and ε_t's are random elements of $\mathcal{H}$, and the Φ_j's are continuous linear operators from $\mathcal{H}$ to $\mathcal{H}$. We say that the AR(p) law of motion (3.1) is engendered by the map $\Phi : \mathbb{C} \mapsto \mathcal{L}_{\mathcal{H}}$ given by

$$\Phi(z) = \mathrm{I} - \sum_{j=1}^{p} z^j \Phi_j. \tag{3.2}$$

We will refer to an operator-valued polynomial function of a complex variable as an operator pencil; note that some authors reserve this term for linear polynomials. We impose the following conditions on the objects just introduced.

Assumption 3.1. (i) $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ is an iid sequence in $L^2_{\mathcal{H}}$ with positive definite covariance operator $\Sigma \in \mathcal{L}_{\mathcal{H}}$. (ii) $\Phi_1, \ldots, \Phi_p$ are compact operators in $\mathcal{L}_{\mathcal{H}}$ such that $\Phi : \mathbb{C} \mapsto \mathcal{L}_{\mathcal{H}}$ is noninvertible at $z = 1$ and invertible at every other z in the closed unit disk.

Remark 3.1. The innovations ε_t are referred to as strong white noise due to their being centered (i.e., zero expected value) and iid. We impose these conditions for simplicity, but the results to be developed remain valid if the iid condition is replaced with the weaker requirement that the cross-covariance operators for the ε_t's are all zero, as in Franchi and Paruolo (2019a). In the latter case, the ε_t's are merely said to be white noise. More generally, one might consider allowing the ε_t's to be a general I(0) process as in Johansen (1992) and Cheng and Phillips (2009, 2012), or even a subexponential process as in Al Sadoon (2018), but we do not pursue this route here.

Remark 3.2. The results to be developed remain valid if the assumption that $\Phi_1, \ldots, \Phi_p$ are compact is replaced with the weaker but less easily interpretable requirement that $\Phi(z)$ has an eigenvalue of finite type at $z = 1$. See Franchi and Paruolo (2019a) for details.

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

The approach we will take to developing representation theory for I(1) and I(2) autoregressive processes in $\mathcal{H}$ essentially boils down to studying the behavior of $\Phi(z)^{-1}$ near $z = 1$. We achieve this by applying the analytic Fredholm theorem, a complete statement of which is provided in Appendix A.1. To apply this result, we need $\Phi(z)$ to be analytic in z (in fact, it is polynomial in z, hence analytic), and Fredholm operator-valued (meaning that it has finite dimensional kernel and cokernel for all z, which is guaranteed by our compactness condition on the Φ_j's). The analytic Fredholm theorem implies that, for all z in a punctured neighborhood of one, we have

$$\Phi(z)^{-1} = \sum_{k=-d}^{\infty} (z-1)^k \Upsilon_k, \tag{3.3}$$

where $d \in \mathbb{N}$ and $\Upsilon_{-d}, \Upsilon_{-d+1}, \ldots$ is a sequence in $\mathcal{L}_{\mathcal{H}}$. The series in (3.3) is called the Laurent series of $\Phi(z)^{-1}$ around $z = 1$, and converges in $\mathcal{L}_{\mathcal{H}}$. If we assume without loss of generality that $\Upsilon_{-d} \neq 0$, then $\Phi(z)^{-1}$ is said to have a pole of order d at $z = 1$. The operator Υ_{-1} is called the residue of $\Phi(z)^{-1}$ at $z = 1$. A pole of order one is said to be simple. We call the sum of the leading terms indexed by $k = -d, \ldots, -1$ the principal part of the Laurent series, and we call the truncated series excluding these leading terms the analytic part of the Laurent series.

An important implication of the analytic Fredholm theorem is that the leading Laurent coefficients $\Upsilon_{-d}, \ldots, \Upsilon_{-1}$ in (3.3) are all of finite rank. In the representation theory to be developed, this has the effect of ensuring that we always obtain a cointegrating space with finite codimension (i.e., with finite dimensional orthogonal complement).

It will be convenient to introduce some additional notation. We will write Π_0 and Π_1 for the values taken by $\Phi(z)$ and its first derivative at $z = 1$:

$$\Pi_0 = \Phi(1), \quad \Pi_1 = \Phi^{(1)}(1).$$

We also define the linear spaces

$$\alpha_1 = \operatorname{ran} \Pi_0, \quad \beta_1 = (\ker \Pi_0)^{\perp}.$$

Our compactness condition on the Φ_j's ensures that α_1 and β_1 are closed linear spaces with equal and finite codimension.

We have yet to give a formal definition of the I(d) property. For our purposes, it is sufficient to define the I(0) property for standard linear processes. We will need to consider standard linear processes in $\mathcal{H}$ and in $\mathbb{C}$, with innovations in $\mathcal{H}$, so give the following definition of a standard linear process in an arbitrary separable complex Hilbert space $\mathcal{H}'$.

DEFINITION 3.1. *A sequence $(W_t, t \geq t_0)$ in $L^2_{\mathcal{H}'}$ is called a standard linear process in $\mathcal{H}'$ if there is another separable complex Hilbert space $\mathcal{H}$ such that we may write*

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

$$W_t = \sum_{k=0}^{\infty} A_k(\varepsilon_{t-k}), \quad t \geq t_0, \tag{3.4}$$

where $(A_k, k \geq 0)$ *is a norm-summable sequence in* $\mathcal{L}_{\mathcal{H},\mathcal{H}'}$, *and* $(\varepsilon_t, t \in \mathbb{Z})$ *is an iid sequence in* $L^2_{\mathcal{H}}$ *with nonzero covariance operator* $\Sigma \in \mathcal{L}_{\mathcal{H}}$.

Remark 3.3. If we were to require that the two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ are the same, and that $A_0 = \mathrm{I}$, then our definition of a standard linear process in $\mathcal{H}$ would be the same as that of Bosq (2000, p. 183). Our more general definition is needed because if $(W_t, t \geq t_0)$ is a standard linear process in $\mathcal{H}$ with innovations in $\mathcal{H}$, then for any $x \in \mathcal{H}$ we may write

$$\langle x, W_t \rangle = \sum_{k=0}^{\infty} \langle x, A_k(\varepsilon_{t-k}) \rangle = \sum_{k=0}^{\infty} \hat{A}_k^x(\varepsilon_{t-k}), \quad t \geq t_0,$$

where the $\hat{A}_k^x$'s are given by

$$\hat{A}_k^x(y) = \langle x, A_k(y) \rangle, \quad y \in \mathcal{H},$$

and form a norm-summable sequence in $\mathcal{L}_{\mathcal{H},\mathbb{C}}$. The sequence of inner products $(\langle x, W_t \rangle, t \geq t_0)$ is thus a standard linear process in $\mathbb{C}$ with innovations in $\mathcal{H}$.

We may now define the I(d) property for sequences in $L^2_{\mathcal{H}'}$, with $\mathcal{H}'$ an arbitrary separable complex Hilbert space.

DEFINITION 3.2. *We say that a sequence* $(W_t, t \geq t_0)$ *in* $L^2_{\mathcal{H}'}$ *is I(0) if it is a standard linear process in* $\mathcal{H}'$ *admitting a representation* (3.4) *in which* Σ *is positive definite and the* A_k*'s satisfy* $\sum_{k=0}^{\infty} A_k \neq 0$ *and* $\sum_{k=0}^{\infty} k \|A_k\|_{\mathcal{L}_{\mathcal{H},\mathcal{H}'}} < \infty$.

DEFINITION 3.3. *We say that a sequence* $(W_t, t \geq t_0)$ *in* $L^2_{\mathcal{H}'}$ *is I(d) for* $d \in \mathbb{N}$ *if its* d*th difference is an I(0) standard linear process in* $\mathcal{H}'$.

Remark 3.4. The summability condition on the norms of the coefficients A_k in Definition 3.2 is called 1-summability. It was used by Phillips and Solo (1992) to facilitate a version of the Beveridge-Nelson decomposition for a time series whose difference is I(0). In the results to be developed, all processes claimed to be I(0) in fact have coefficients decaying exponentially in norm, so 1-summability is easily satisfied.

Our first result provides an I(1) representation for autoregressive Hilbertian processes for which $\Phi(z)^{-1}$ has a simple pole at $z = 1$. We will discuss the simple pole condition in more detail later in this section.

THEOREM 3.1. *Suppose that Assumption 3.1 is satisfied, and that* $\Phi(z)^{-1}$ *has a simple pole at* $z = 1$. *Let* Υ_{-1} *denote the residue of* $\Phi(z)^{-1}$ *at* $z = 1$, *let* $\tilde{\Psi}(z)$ *denote the analytic part of the Laurent series of* $\Phi(z)^{-1}$ *around* $z = 1$, *and set*

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

$\tilde{\Psi}_k = \tilde{\Psi}^{(k)}(0)/k!$. *A sequence* $(X_t, t \geq -p+1)$ *in* $L^2_{\mathcal{H}}$ *satisfying the law of motion* (3.1) *allows the following representation: for some* $Z_0 \in L^2_{\mathcal{H}}$ *and all* $t \geq 1$ *we have*

$$X_t = Z_0 - \Upsilon_{-1}\left(\sum_{s=1}^{t} \varepsilon_s\right) + \nu_t. \tag{3.5}$$

Here, $(\nu_t, t \geq 1)$ *is a stationary sequence of random elements of* $\mathcal{H}$ *defined by the* $L^2_{\mathcal{H}}$*-convergent series* $\nu_t = \sum_{k=0}^{\infty} \tilde{\Psi}_k(\varepsilon_{t-k})$. *Moreover,*

(1) *The range of* Υ_{-1} *is equal to* $\beta_1^{\perp}$ *and has positive and finite dimension;*
(2) *If* Z_0 *belongs to* $\beta_1^{\perp}$*, then for nonzero* $x \in \mathcal{H}$ *the sequence of inner products* $(\langle x, X_t \rangle, t \geq 1)$ *is* I(0) *if* $x \in \beta_1$*, and is* I(1) *otherwise.*

Remark 3.5. Theorem 3.1 is similar to Theorem 4.1 of Beare, Seo, and Seo (2017), but makes the connection to the analytic behavior of $\Phi(z)^{-1}$ explicit. The latter result is more generally applicable in one respect: compactness of the autoregressive operator is not assumed when $p = 1$. The approach taken here relies on the analytic Fredholm theorem and therefore requires $\Phi(z)$ to be Fredholm, which may not be the case if the autoregressive operators are not compact.

Remark 3.6. The residue Υ_{-1} appearing in Theorem 3.1 has finite rank by the analytic Fredholm theorem. The attractor space, which is the subspace of $\mathcal{H}$ in which the I(1) stochastic trend in the Beveridge-Nelson representation (3.5) takes values, thus has finite dimension. We are therefore outside the framework considered by Chang, Hu, and Park (2016), in which the cointegrating space has finite dimension and the attractor space has finite codimension.

When can we expect the simple pole condition in Theorem 3.1 to be satisfied? Our next result provides equivalent reformulations of this condition that may be easier to check in practice, and a formula for the residue Υ_{-1} in terms of α_1, β_1 and Π_1. (Recall we defined $\Pi_1 = \Phi^{(1)}(1)$.)

THEOREM 3.2. *Suppose that Assumption 3.1(ii) holds. The following four conditions are equivalent.*

(1) $\Phi(z)^{-1}$ *has a simple pole at* $z = 1$.
(2) *The operator* $\Lambda_1 : \beta_1^{\perp} \to \alpha_1^{\perp}$ *obtained by restricting* $\mathrm{P}_{\alpha_1^{\perp}} \Pi_1$ *to* $\beta_1^{\perp}$ *is bijective.*
(3) $\mathcal{H} = \alpha_1 \oplus \Pi_1 \beta_1^{\perp}$.
(4) $\mathcal{H} = \alpha_1 + \Pi_1 \beta_1^{\perp}$.

If $\Phi(z)^{-1}$ *has a simple pole at* $z = 1$*, then its residue at* $z = 1$ *is* $\Upsilon_{-1} = \Lambda_1^{-1} \mathrm{P}_{\alpha_1^{\perp}}$.

Remark 3.7. The closest results we have found to Theorem 3.2 in prior literature are those of Steinberg (1968) and Howland (1971), who worked in a more general Banach space setting. Steinberg (1968) established sufficient conditions

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

for a simple pole, and Howland (1971) established the equivalence of Conditions (1) and (3).

Remark 3.8. Beare, Seo, and Seo (2017, Prop. 4.1) showed that condition (3) of Theorem 3.2 is equivalent to the I(1) condition given by Johansen (1991, Thm. 4.1) in the finite dimensional case $\mathcal{H} = \mathbb{C}^n$. In this setting, we may let $r < n$ be the rank of the $n \times n$ complex matrix Π_0, let α and β be full-rank $n \times r$ complex matrices such that $\Pi_0 = \alpha\beta'$, and let $\alpha_\perp$ and $\beta_\perp$ be full-rank $n \times (n-r)$ complex matrices such that $\alpha'\alpha_\perp = 0$ and $\beta'\beta_\perp = 0$. The Johansen I(1) condition is satisfied when the $(n-r) \times (n-r)$ complex matrix $\alpha_\perp' \Pi_1 \beta_\perp$ is invertible. Examples 4.1–4.3 of Beare, Seo, and Seo (2017) illustrate the reformulation of the Johansen I(1) condition as a direct sum decomposition.

Remark 3.9. The direct sum appearing in condition (3) of Theorem 3.2 is not in general an orthogonal direct sum. Franchi and Paruolo (2019a) showed that, when $\Phi(z)$ is noninvertible at $z = 1$, condition (3) is equivalent to the following orthogonal direct sum decomposition of $\mathcal{H}$:

$$\mathcal{H} = \beta_1 \oplus \left(\ker \Lambda_1 \mathrm{P}_{\beta_1^\perp}\right)^\perp.$$

In their notation, this is $\mathcal{H} = \tau_0 \oplus \tau_1$.

Remark 3.10. Our assumption that the operators $\Phi_1, \ldots, \Phi_p$ are compact implies that Π_0 is Fredholm of index zero. If Π_0 were Fredholm but not of index zero, then it would be impossible to satisfy condition (2). This is because bijectivity of Λ_1 requires its domain and codomain to have the same dimension. However, our proof that condition (4) implies condition (1) does not use the index-zero property, and remains valid if the compactness condition on $\Phi_1, \ldots, \Phi_p$ is weakened to require only that $\Phi(z)$ is Fredholm operator-valued.

In the special case where $p = 1$, conditions (3) and (4) of Theorem 3.2 take on a particularly simple form, and another related equivalent condition becomes available. Moreover, the direct sum decomposition asserted by condition (3) serves to define an oblique projection that is the negative of the residue of our simple pole. The following corollary to Theorem 3.2 provides details.

COROLLARY 3.1. *Suppose that Assumption 3.1(ii) holds and that* $p = 1$. *The following four conditions are equivalent.*

(1) $\Phi(z)^{-1}$ *has a simple pole at* $z = 1$.
(2) $\mathcal{H} = \alpha_1 \oplus \beta_1^\perp$.
(3) $\mathcal{H} = \alpha_1 + \beta_1^\perp$.
(4) $\{0\} = \alpha_1 \cap \beta_1^\perp$.

If $\Phi(z)^{-1}$ *has a simple pole at* $z = 1$, *then its residue at* $z = 1$ *is the negative of the projection on* $\beta_1^\perp$ *along* α_1.

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

Remark 3.11. The oblique projection appearing in Corollary 3.1 is in fact the Riesz projection for the unit eigenvalue of Φ_1. Said Riesz projection is defined (Gohberg, Goldberg, and Kaashoek, 1990, p. 9; Markus, 2012, pp. 11–12) by the contour integral

$$P = \frac{1}{2\pi \mathrm{i}} \oint_{\Gamma} (z\mathrm{I} - \Phi_1)^{-1} \mathrm{d}z, \tag{3.6}$$

where Γ is a positively oriented smooth Jordan curve around one separating it from zero and from any other eigenvalues of Φ_1, and where the integral of an $\mathcal{L}_{\mathcal{H}}$-valued function should be understood in the sense of Bochner. Let $\gamma : [0,1] \to \mathbb{C}$ be a smooth parametrization of Γ, and rewrite (3.6) as

$$P = \frac{1}{2\pi \mathrm{i}} \int_0^1 (\gamma(t)\mathrm{I} - \Phi_1)^{-1} \gamma'(t) \mathrm{d}t. \tag{3.7}$$

The image of Γ under the reciprocal transform $z \mapsto z^{-1}$, which we denote Γ', is a positively oriented smooth Jordan curve around one separating it from any other poles of $\Phi(z)^{-1}$ and from zero. It admits the parametrization $t \mapsto 1/\gamma(t) =: \delta(t)$. A little calculus shows that $\gamma'(t) = -\delta'(t)/\delta(t)^2$, and so from (3.7) we have

$$P = \frac{-1}{2\pi \mathrm{i}} \int_0^1 \delta(t)^{-1} (\mathrm{I} - \delta(t)\Phi_1)^{-1} \delta'(t) \mathrm{d}t = \frac{-1}{2\pi \mathrm{i}} \oint_{\Gamma'} z^{-1} \Phi(z)^{-1} \mathrm{d}z. \tag{3.8}$$

The residue theorem therefore tells us that P is the negative of the residue of $z^{-1}\Phi(z)^{-1}$ at $z = 1$, implying that the residue of $\Phi(z)^{-1}$ at $z = 1$ is $-P$. It now follows from Corollary 3.1 that when the direct sum decomposition $\mathcal{H} = \alpha_1 \oplus \beta_1^{\perp}$ is satisfied, the Riesz projection for the unit eigenvalue of Φ_1 is the projection on $\beta_1^{\perp}$ along α_1. This is a nonorthogonal projection except when $\alpha_1 = \beta_1$, which would occur if, for instance, Φ_1 is a normal operator.

Remark 3.12. It is apparent from our discussion in Remark 3.11 that, under any of the equivalent I(1) conditions in Corollary 3.1, the restriction of Π_0 to the range of the Riesz projection P is zero. This is precisely the condition given by Hu and Park (2016) for an AR(1) process in $\mathcal{H}$ to be I(1).

We close this section with several examples of the use of Corollary 3.1 for verifying that $\Phi(z)^{-1}$ has a simple pole at $z = 1$.

Example 3.1
Suppose that $p = 1$ and that Φ_1 is self-adjoint. Then, Π_0 is also self-adjoint, implying that $\alpha_1 = \beta_1$. The direct sum decomposition $\mathcal{H} = \alpha_1 \oplus \beta_1^{\perp}$ appearing in Corollary 3.1 is therefore satisfied, and is in this case an orthogonal direct sum decomposition.

Example 3.2
Let $(e_j, j \in \mathbb{N})$ be an orthonormal basis of $\mathcal{H}$. Suppose that $p = 1$, and that Φ_1 is given by

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

$$\Phi_1(x) = \langle x, e_1\rangle(e_1 + e_2) + \sum_{j=2}^{\infty} \lambda_j \langle x, e_j\rangle e_j, \quad x \in \mathcal{H},$$

with $\lambda_j \in (0,1)$ for $j \geq 2$ and $\lambda_j \to 0$ as $j \to \infty$. For any $x \in \mathcal{H}$ with representation $x = \sum_{j=1}^{\infty} c_j e_j$, $c_j = \langle x, e_j\rangle$, we have

$$(\mathrm{I} - \Phi_1)(x) = (c_2(1-\lambda_2) - c_1)e_2 + \sum_{j=3}^{\infty} c_j(1-\lambda_j)e_j. \tag{3.9}$$

Since $\lambda_j \neq 1$ for all $j \geq 3$, it is clear that $e_j \notin \ker(\mathrm{I} - \Phi_1)$ for all $j \geq 3$. Moreover,

$$(\mathrm{I} - \Phi_1)(c_1 e_1 + c_2 e_2) = (c_2(1-\lambda_2) - c_1)e_2.$$

It follows that

$$\beta_1^{\perp} = \ker(\mathrm{I} - \Phi_1) = \{c_1 e_1 + c_2 e_2 : c_1 = c_2(1-\lambda_2)\}. \tag{3.10}$$

Moreover, it may be deduced that $\alpha_1 = \mathrm{clsp}\{e_j : j \geq 2\}$, the closed linear span of $\{e_j : j \geq 2\}$, as follows. Any $x \in \mathrm{clsp}\{e_j : j \geq 2\}$ may be written as $x = \sum_{j=2}^{\infty} d_j e_j$ for some square-summable sequence $(d_j, j \geq 2)$. We can always find another square-summable sequence $(c_j, j \in \mathbb{N})$ such that

$$d_2 = c_2(1-\lambda_2) - c_1 \quad \text{and} \quad d_j = c_j(1-\lambda_j), \quad j \geq 3. \tag{3.11}$$

Then

$$(\mathrm{I} - \Phi_1)\left(\sum_{j=1}^{\infty} c_j e_j\right) = (c_2(1-\lambda_2) - c_1)e_2 + \sum_{j=3}^{\infty} c_j(1-\lambda_j)e_j = \sum_{j=2}^{\infty} d_j e_j = x,$$

which shows that $x \in \mathrm{ran}(\mathrm{I} - \Phi_1) = \alpha_1$. Thus, $\mathrm{clsp}\{e_j : j \geq 2\} \subseteq \alpha_1$. In addition, it is easily deduced that $\alpha_1 \subseteq \mathrm{clsp}\{e_j : j \geq 2\}$ using (3.9). Therefore, $\alpha_1 = \mathrm{clsp}\{e_j : j \geq 2\}$. From (3.10), we see that the only element of $\beta_1^{\perp}$ belonging to $\mathrm{clsp}\{e_j : j \geq 2\}$ is zero. Thus, the condition $\{0\} = \alpha_1 \cap \beta_1^{\perp}$ appearing in Corollary 3.1 is satisfied.

Example 3.3

Suppose that in Example 3.2 we instead defined $\Phi_1 \in \mathcal{L}_{\mathcal{H}}$ by

$$\Phi_1(x) = \langle x, e_1\rangle(e_1 + e_2 + e_3) + \langle x, e_2\rangle e_2 + \langle x, e_3\rangle e_3 + \sum_{j=4}^{\infty} \lambda_j \langle x, e_j\rangle e_j, \quad x \in \mathcal{H},$$

with $\lambda_j \in (0,1)$ for $j \geq 4$ and $\lambda_j \to 0$ as $j \to \infty$. For any $x \in \mathcal{H}$ with representation $x = \sum_{j=1}^{\infty} c_j e_j$, $c_j = \langle x, e_j\rangle$, we now have

$$(\mathrm{I} - \Phi_1)(x) = -c_1 e_2 - c_1 e_3 + \sum_{j=4}^{\infty} c_j(1-\lambda_j)e_j.$$

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

Since $\lambda_j \neq 1$ for all $j \geq 4$, it is clear that $e_j \notin \ker(\mathrm{I} - \Phi_1)$ for all $j \geq 4$. Moreover,

$$(\mathrm{I} - \Phi_1)(c_1 e_1 + c_2 e_2 + c_3 e_3) = -c_1 e_2 - c_1 e_3. \tag{3.12}$$

It follows that $\beta_1^\perp = \ker(\mathrm{I} - \Phi_1) = \mathrm{sp}\{e_2, e_3\}$. Furthermore, arguments similar to those in Example 3.2 can be used to show that

$$\alpha_1 = \mathrm{clsp}\{e_2 + e_3, e_4, e_5, \ldots\}.$$

It follows that $\alpha_1 \cap \beta_1^\perp = \mathrm{sp}\{e_2 + e_3\}$. Thus, the condition $\{0\} = \alpha_1 \cap \beta_1^\perp$ appearing in Corollary 3.1 is violated.

4. I(2) AUTOREGRESSIVE HILBERTIAN PROCESSES

In this section, we state our results for I(2) autoregressive processes. We continue to assume that the autoregressive coefficients and innovations of such a process satisfy Assumption 3.1. It will be convenient to introduce some additional notation. We define the linear spaces

$$\alpha_2 = \alpha_1 + \Pi_1 \beta_1^\perp, \quad \beta_2 = \beta_1 + \Pi_1^* \alpha_1^\perp. \tag{4.1}$$

The role played by α_2 and β_2 in our I(2) results will be analogous to the role played by α_1 and β_1 in our I(1) results. The notation Π_1^* refers to the adjoint operator to Π_1.

Remark 4.1. The linear spaces $\alpha_2^\perp$ and $\beta_2^\perp$ have equal and finite dimension and are, respectively, the cokernel and kernel of the operator Λ_1 appearing in Theorem 3.2. To see why, observe that α_2 satisfies

$$\alpha_2^\perp = (\alpha_1 + \mathrm{P}_{\alpha_1^\perp} \Pi_1 \beta_1^\perp)^\perp = (\alpha_1 + \mathrm{ran}\, \Lambda_1)^\perp = \ker \Lambda_1^*,$$

and similarly β_2 satisfies

$$\beta_2^\perp = (\beta_1 + \mathrm{P}_{\beta_1^\perp} \Pi_1^* \alpha_1^\perp)^\perp = (\beta_1 + \mathrm{ran}\, \Lambda_1^*)^\perp = \ker \Lambda_1.$$

The operator Λ_1 is invertible under any of the equivalent I(1) conditions given in Theorem 3.2, in which case we must have $\alpha_2 = \beta_2 = \mathcal{H}$. In this section, we are interested in the case where our I(1) conditions fail. This occurs when $\alpha_2^\perp$ and $\beta_2^\perp$ have positive dimension.

Remark 4.2. Simpler expressions for α_2 and β_2 become available when $p = 1$. In this case, we have $\Pi_1 = \Pi_0 - \mathrm{I}$, from which it follows easily that $\Pi_1 \beta_1^\perp = \beta_1^\perp$ and $\Pi_1^* \alpha_1^\perp = \alpha_1^\perp$. We may therefore write

$$\alpha_2 = \alpha_1 + \beta_1^\perp, \quad \beta_2 = \beta_1 + \alpha_1^\perp.$$

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

Our first result in this section provides an I(2) analogue to Theorem 3.1 in Section 3. It establishes an I(2) representation for autoregressive Hilbertian processes for which $\Phi(z)^{-1}$ has a pole of second order at $z = 1$. We will discuss the pole condition in more detail later in this section.

THEOREM 4.1. *Suppose that Assumption 3.1 is satisfied, and that $\Phi(z)^{-1}$ has a pole of second order at $z = 1$. Let Υ_{-2} and Υ_{-1} denote the coefficients in the principal part of the Laurent series of $\Phi(z)^{-1}$ around $z = 1$, let $\tilde{\Psi}(z)$ denote the analytic part of the Laurent series of $\Phi(z)^{-1}$ around $z = 1$, and set $\tilde{\Psi}_k = \tilde{\Psi}^{(k)}(0)/k!$. A sequence $(X_t, t \geq -p+1)$ in $L^2_{\mathcal{H}}$ satisfying the law of motion* (3.1) *allows the following representation: for some $Z_0, Z_1 \in L^2_{\mathcal{H}}$ and all $t \geq 1$, we have*

$$X_t = Z_0 + tZ_1 + \Upsilon_{-2}\left(\sum_{s=1}^{t}\sum_{r=1}^{s}\varepsilon_r\right) - \Upsilon_{-1}\left(\sum_{s=1}^{t}\varepsilon_s\right) + \nu_t. \tag{4.2}$$

Here, $(\nu_t, t \geq 1)$ is a stationary sequence of random elements of $\mathcal{H}$ defined by the $L^2_{\mathcal{H}}$-convergent series $\nu_t = \sum_{k=0}^{\infty}\tilde{\Psi}_k(\varepsilon_{t-k})$. Moreover,

(1) *The range of Υ_{-2} is equal to $\beta_2^{\perp}$ and has positive and finite dimension;*
(2) *If Z_0 and Z_1 belong to $\beta_2^{\perp}$, then for nonzero $x \in \mathcal{H}$ the sequence of inner products $(\langle x, X_t\rangle, t \geq 1)$ is* I(0) *or* I(1) *if $x \in \beta_2$, and is* I(2) *otherwise.*

In view of Claim (2) in Theorem 4.1, we may refer to the linear subspace β_2 as the cointegrating space. For $x \in \beta_2$, the sequence of inner products $(\langle x, \Delta X_t\rangle, t \geq 1)$ may be I(0) or I(1) (ignoring the effect of the deterministic components Z_0 and Z_1). Polynomial cointegration may also occur. This will be discussed more fully in Remark 4.5 below.

Our next result provides an I(2) analogue to Theorem 3.2 in Section 3. It establishes necessary and sufficient conditions for $\Phi(z)^{-1}$ to have a pole of second order at $z = 1$, as assumed in Theorem 4.1. It also establishes formulas for the Laurent coefficients Υ_{-2} and Υ_{-1} appearing in Theorem 4.1. To state our result, it will be convenient to introduce some additional notation. Similar to Π_0 and Π_1, we define

$$\Pi_2 = \frac{1}{2}\Phi^{(2)}(1) - \Pi_1\Pi_0^{+}\Pi_1, \quad \Pi_3 = \frac{1}{6}\Phi^{(3)}(1) - \Pi_1\Pi_0^{+}\Pi_1\Pi_0^{+}\Pi_1.$$

Here and elsewhere, a plus superscript is used to denote the Moore–Penrose inverse of a continuous linear operator between Hilbert spaces with closed range (see, e.g., Ben-Israel and Greville, 2003, chap. 9).

THEOREM 4.2. *Suppose that Assumption 3.1(ii) holds and that $\Phi(z)^{-1}$ does not have a simple pole at $z = 1$. Then, the following four conditions are equivalent.*

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

(1) $\Phi(z)^{-1}$ *has a pole of second order at* $z = 1$.
(2) *The operator* $\Lambda_2 : \beta_2^\perp \to \alpha_2^\perp$ *obtained by restricting* $\mathrm{P}_{\alpha_2^\perp}\Pi_2$ *to* $\beta_2^\perp$ *is bijective.*
(3) $\mathcal{H} = \alpha_2 \oplus \Pi_2\beta_2^\perp$.
(4) $\mathcal{H} = \alpha_2 + \Pi_2\beta_2^\perp$.

If $\Phi(z)^{-1}$ *has a pole of second order at* $z = 1$*, then the coefficients of* $(z-1)^{-2}$ *and* $(z-1)^{-1}$ *in the Laurent series of* $\Phi(z)^{-1}$ *around* $z = 1$ *are given by*

$$\Upsilon_{-2} = \Lambda_2^{-1}\mathrm{P}_{\alpha_2^\perp} \tag{4.3}$$

and

$$\begin{aligned}\Upsilon_{-1} = {} & \Lambda_1^+\mathrm{P}_{\alpha_1^\perp} - \left[\Lambda_1^+\mathrm{P}_{\alpha_1^\perp}\Pi_2 + \Pi_0^+\Pi_1\right]\Upsilon_{-2} - \Upsilon_{-2}\left[\Pi_2\Lambda_1^{+\prime}\mathrm{P}_{\alpha_1^\perp} + \Pi_1\Pi_0^{+\prime}\right] \\ & + \Upsilon_{-2}\left[\Pi_1\Pi_0^{+\prime}\Pi_2 + \Pi_2\Pi_0^{+\prime}\Pi_1 + \Pi_2\Lambda_1^{+\prime}\mathrm{P}_{\alpha_1^\perp}\Pi_2 - \Pi_3\right]\Upsilon_{-2},\end{aligned} \tag{4.4}$$

respectively.

Remark 4.3. The four equivalent conditions in Theorem 4.2 are equivalent to the I(2) condition given by Johansen (1992, Thm. 3) in the finite dimensional case $\mathcal{H} = \mathbb{C}^n$. Continuing with the notation of Remark 3.8, suppose that the Johansen I(1) condition fails, and let φ and η be full-rank $(n-r) \times s$ $(s < n-r)$ complex matrices such that $(\alpha_\perp'\alpha_\perp)^{-1}\alpha_\perp'\Pi_1\beta_\perp(\beta_\perp'\beta_\perp)^{-1} = \varphi\eta'$. Let $\tilde{\alpha}_\perp$ and $\tilde{\beta}_\perp$ be full-rank $n \times (n-r-s)$ complex matrices whose columns are orthogonal to those of $\tilde{\alpha} := (\alpha, \alpha_\perp\varphi)$ and $\tilde{\beta} := (\beta, \beta_\perp\eta)$, respectively. Note that the column spaces of $\tilde{\alpha}$ and $\tilde{\beta}$ are, respectively, the linear spaces α_2 and β_2 defined in (4.1) above. The Johansen I(2) condition is satisfied when the $(n-r-s) \times (n-r-s)$ complex matrix $\tilde{\alpha}_\perp'\Pi_2\tilde{\beta}_\perp$ is invertible. Lemma 4.1 of Beare, Seo, and Seo (2017) implies that $\tilde{\alpha}_\perp'\Pi_2\tilde{\beta}_\perp$ is invertible if and only if $\mathcal{H}$ is the direct sum of the null space of $\tilde{\alpha}_\perp'$ and the column space of $\Pi_2\tilde{\beta}_\perp$. The former space is α_2 and the latter space is $\Pi_2\beta_2^\perp$, so the Johansen I(2) condition is equivalent to condition (3) in Theorem 4.2.

Remark 4.4. Formulas (4.3) and (4.4) in Theorem 4.2 correspond to formulas (12) and (13) of Johansen (2009, Thm. 5), given for the finite dimensional case $\mathcal{H} = \mathbb{C}^n$. The objects C_1, C_2, $\bar{\beta}_1\bar{\alpha}_1'$, $\bar{\beta}\bar{\alpha}'$, $\dot{\Pi}$, θ and $(1/6)\dddot{\Pi} - \dot{\Pi}\bar{\beta}\bar{\alpha}'\ddot{\Pi}\bar{\beta}\bar{\alpha}'\dot{\Pi}$ in Johansen's notation correspond, respectively, to $-\Upsilon_{-1}$, Υ_{-2}, $\Lambda_1^+\mathrm{P}_{\alpha_1^\perp}$, $-\Pi_0^+$, Π_1, Π_2 and Π_3 in our notation.

Remark 4.5. The direct sum appearing in condition (3) of Theorem 4.2 is not in general an orthogonal direct sum. Extending results of Johansen (1992) from the finite dimensional case $\mathcal{H} = \mathbb{C}^n$ to a more general Hilbert space setting, Franchi and Paruolo (2019a) showed that, when $\Phi(z)$ is noninvertible at $z = 1$ and the I(1) condition fails, an equivalent necessary and sufficient condition for a pole

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

of second order is the following tripartite orthogonal direct sum decomposition of $\mathcal{H}$:

$$\mathcal{H} = \beta_1 \oplus \left(\ker \Lambda_1 \mathrm{P}_{\beta_1^\perp}\right)^\perp \oplus \left(\ker \Lambda_2 \mathrm{P}_{\beta_2^\perp}\right)^\perp.$$

In their notation, this is $\mathcal{H} = \tau_0 \oplus \tau_1 \oplus \tau_2$. Moreover, we have the following bipartite orthogonal direct sum decomposition of the cointegrating space β_2:

$$\beta_2 = \beta_1 \oplus \left(\ker \Lambda_1 \mathrm{P}_{\beta_1^\perp}\right)^\perp,$$

or $\beta_2 = \tau_0 \oplus \tau_1$. This decomposition is informative because the two subspaces decomposing the cointegrating space correspond to different kinds of cointegrating behavior. Suppose that Z_0 and Z_1 belong to $\beta_2^\perp$, so that we may ignore the effect of deterministic components. For $x \in \beta_2$ not belonging to β_1, the sequence $(\langle x, X_t\rangle, t \geq 1)$ is I(1). For nonzero $x \in \beta_1$, the sequence $(\langle x, X_t\rangle, t \geq 1)$ may be I(0) or I(1), but the sequence $(\langle x, X_t\rangle - \langle x, \Pi_0^+ \Pi_1 \Delta X_t\rangle, t \geq 1)$ is always I(0). The case where the inner product sequences $(\langle x, X_t\rangle, t \geq 1)$ and $(\langle x, \Pi_0^+ \Pi_1 \Delta X_t\rangle, t \geq 1)$ are I(1) and cointegrated is called polynomial cointegration or multicointegration (Yoo, 1987; Granger and Lee, 1989, 1990; Engsted and Johansen, 1999; Phillips and Kheifets, 2019).

Remark 4.6. From our discussion of the I(2) case in Remark 4.5, it is apparent that (ignoring deterministic components) the sequence $(\langle x, X_t\rangle, t \geq 1)$ is I(0) for precisely those nonzero $x \in \beta_1$ for which $(\langle x, \Pi_0^+ \Pi_1 \Delta X_t\rangle, t \geq 1)$ is stationary. This is the collection of all nonzero $x \in \beta_1$ such that $(\Pi_0^+ \Pi_1)^*(x)$ belongs to β_2, the cointegrating space. Therefore, $(\langle x, X_t\rangle, t \geq 1)$ is I(0) if and only if x is a nonzero element of the linear space

$$\zeta = \beta_1 \cap (\Pi_0^+ \Pi_1 \beta_2^\perp)^\perp.$$

Remark 4.7. When $p = 1$, the linear space $\Pi_2 \beta_2^\perp$ appearing in conditions (3) and (4) of Theorem 4.2 is equal to $(\mathrm{I} - \Pi_0^+)(\alpha_1 \cap \beta_1^\perp)$. To see why, observe that when $p = 1$, we have

$$\Pi_2 = -\Pi_1 \Pi_0^+ \Pi_1 = -(\mathrm{I} - \Pi_0)\Pi_0^+(\mathrm{I} - \Pi_0) = -\Pi_0^+ + \Pi_0^+ \Pi_0 + \Pi_0 \Pi_0^+ - \Pi_0 \Pi_0^+ \Pi_0.$$

From Remark 4.2, we know that $\beta_2^\perp = \alpha_1 \cap \beta_1^\perp$ when $p = 1$. Therefore, since Π_0 is zero on $\beta_1^\perp$, we have

$$\Pi_2 \beta_2^\perp = (\Pi_0 \Pi_0^+ - \Pi_0^+)(\alpha_1 \cap \beta_1^\perp).$$

From the basic properties of Moore–Penrose inverses, we know that $\Pi_0 \Pi_0^+$ is orthogonal projection on α_1, the range of Π_0. The restriction of this projection to $\alpha_1 \cap \beta_1^\perp$ coincides with the identity, so our claim is established.

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

Remark 4.8. When $p = 1$, the linear space ζ appearing in Remark 4.6 is equal to $\beta_1 \cap (\Pi_0^+(\alpha_1 \cap \beta_1^\perp))^\perp$. To see why, observe that when $p = 1$, we have $\beta_2^\perp = \alpha_1 \cap \beta_1^\perp$ (shown in Remark 4.2) and $\Pi_0^+ \Pi_1 = \Pi_0^+ \Pi_0 - \Pi_0^+$. From the basic properties of Moore–Penrose inverses, we know that $\Pi_0^+ \Pi_0$ is orthogonal projection on β_1, the corange of Π_0. Thus, $\Pi_0^+ \Pi_0$ is zero on $\alpha_1 \cap \beta_1^\perp$, and we have $\Pi_0^+ \Pi_1 \beta_2^\perp = \Pi_0^+(\alpha_1 \cap \beta_1^\perp)$, establishing our claim.

Remark 4.9. As discussed in Remark 3.11, when $p = 1$, the negative of the residue Υ_{-1} is a projection, and is called the Riesz projection for the unit eigenvalue of Φ_1. The space on which $-\Upsilon_{-1}$ projects is called the generalized eigenspace for the unit eigenvalue, and its dimension is called the algebraic multiplicity of the unit eigenvalue. Contained within the generalized eigenspace is the usual eigenspace $\beta_1^\perp$, whose dimension is called the geometric multiplicity of the unit eigenvalue (Gohberg, Goldberg, and Kaashoek, 1990, p. 26). The generalized eigenspace is in fact the orthogonal complement to ζ, the subspace of the cointegrating space β_2 yielding I(0) inner products, as defined in Remark 4.6. To see why, observe that $\zeta^\perp$ is the subspace occupied by the I(1) and I(2) trends in the representation (4.2), which is the sum of the ranges of Υ_{-2} and Υ_{-1}. But from formula (4.3) in Theorem 4.2, we know that the range of Υ_{-2} is $\beta_2^\perp$, which is contained in the usual eigenspace $\beta_1^\perp$, and therefore contained in the generalized eigenspace, which is the range of Υ_{-1}. Therefore, $\zeta^\perp$ is the generalized eigenspace. In view of Remark 4.8, we have

$$\zeta^\perp = \beta_1^\perp + \Pi_0^+(\alpha_1 \cap \beta_1^\perp). \tag{4.5}$$

From Corollary 3.1, we know that for $p = 1$ the I(1) condition fails precisely when $\alpha_1 \cap \beta_1^\perp \neq \{0\}$. Since the Moore–Penrose inverse Π_0^+ defines a bijection from α_1 to β_1, and $\beta_1^\perp$ is finite dimensional, we deduce that when $p = 1$ and the I(1) condition fails we must have $\dim \beta_1^\perp < \dim \zeta^\perp$. Thus, we see that when the I(1) condition fails, the algebraic multiplicity of the unit eigenvalue exceeds its geometric multiplicity. This contrasts with the situation when the I(1) condition is satisfied, where, as is apparent from our discussion in Remark 3.11, the algebraic and geometric multiplicities of the unit eigenvalue are equal. The fact that the equality of geometric and algebraic multiplicities of the unit eigenvalue implies an I(1) representation was observed by Johansen (1996, Cor. 4.3) in the finite dimensional case $\mathcal{H} = \mathbb{C}^n$.

Remark 4.10. Hu and Park (2016) provide the following condition for an AR(1) process in $\mathcal{H}$ to be I(2): the restriction of Π_0 to the generalized eigenspace $\zeta^\perp$ must be nilpotent of degree two. We can verify this condition using the expression we obtained for $\zeta^\perp$ in (4.5). Based on this expression, any element of $\zeta^\perp$ may be written as $x + \Pi_0^+(y)$, where $x \in \beta_1^\perp$ and $y \in \alpha_1 \cap \beta_1^\perp$. Observe that

$$\Pi_0(x + \Pi_0^+(y)) = \Pi_0 \Pi_0^+(y) = y \quad \text{and} \quad \Pi_0^2(x + \Pi_0^+(y)) = \Pi_0(y) = 0,$$

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

since $\Pi_0\Pi_0^+$ is orthogonal projection on α_1. Thus, the restriction of Π_0 to $\zeta^\perp$ is zero if and only if $\alpha_1 \cap \beta_1^\perp = \{0\}$, which is one of our necessary and sufficient conditions for a simple pole from Corollary 3.1. Assuming that this condition fails, and that our pole is of second order, we find that the restriction of Π_0 to $\zeta^\perp$ is nilpotent of degree two.

To illustrate the preceding results on I(2) representations, we discuss three examples. The first two examples correspond to Examples 4.2–4.3 of Johansen (1996) and Beare, Seo, and Seo (2017), with $\gamma = 2$ in the latter example so that the solution is I(2). These examples are finite (two) dimensional.

Example 4.1
Consider the first-order autoregressive law of motion $X_t = \Phi_1(X_{t-1}) + \varepsilon_t$ in $\mathcal{H} = \mathbb{C}^2$ with autoregressive coefficient matrix

$$\Phi_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

The only eigenvalue of Φ_1 is one. With a little algebra we find that

$$\alpha_1 = \beta_1^\perp = \alpha_2 = \beta_2^\perp = \operatorname{sp}\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \Pi_2\beta_2^\perp = \operatorname{sp}\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

We depict these subspaces of $\mathbb{C}^2$ in Figure 1(a), displaying them as subspaces of $\mathbb{R}^2$ since Φ_1 has real elements. (We do the same for Example 4.2 in Figure 1(b).) It is apparent that the (version for $p = 1$ of the) I(1) condition $\mathbb{C}^2 = \alpha_1 \oplus \beta_1^\perp$

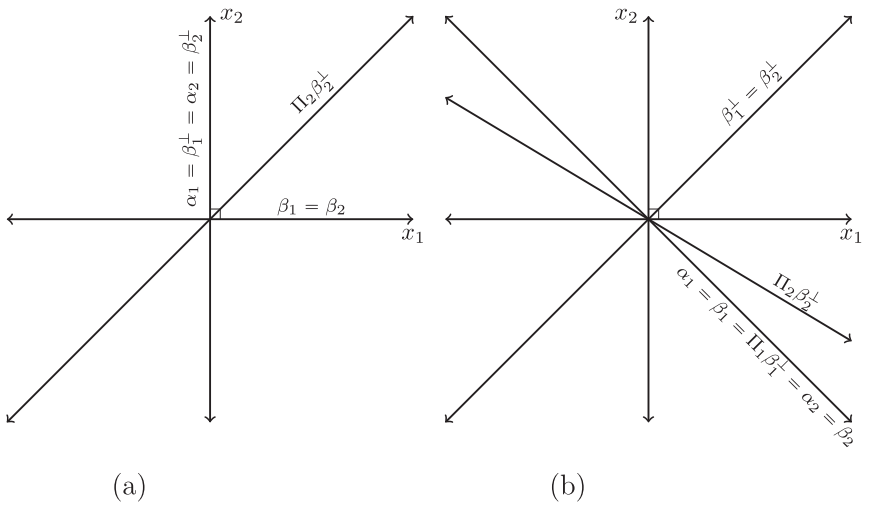


FIGURE 1. Visual aid to Examples 4.1–4.2.

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

is not satisfied, and that the I(2) condition $\mathbb{C}^2 = \alpha_2 \oplus \Pi_2\beta_2^\perp$ is satisfied, so our autoregressive law of motion generates an I(2) process. It is easily verified that $\zeta = \{0\}$, so in this case it follows from Theorem 4.1 that (ignoring deterministic components) the inner product sequence $(\langle x, X_t\rangle, t \geq 1)$ is I(1) for nonzero $x \in \beta_2$ and I(2) for nonzero $x \notin \beta_2$.

Example 4.2
Consider the second-order autoregressive law of motion $X_t = \Phi_1(X_{t-1}) + \Phi_2(X_{t-2}) + \varepsilon_t$ in $\mathcal{H} = \mathbb{C}^2$ with autoregressive coefficient matrices

$$\Phi_1 = \begin{bmatrix} \frac{5}{4} & \frac{7}{4} \\ -\frac{1}{4} & \frac{5}{4} \end{bmatrix}, \quad \Phi_2 = \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix}.$$

The points of noninvertibility of $\Phi(z)$ are $z = 1$ and $z = 2$. It is straightforward to establish that

$$\alpha_1 = \beta_1 = \Pi_1\beta_1^\perp = \alpha_2 = \beta_2 = \mathrm{sp}\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \Pi_2\beta_2^\perp = \mathrm{sp}\begin{bmatrix} 5 \\ -3 \end{bmatrix}.$$

We depict these subspaces in Figure 1(b). It is apparent that the I(1) condition $\mathbb{C}^2 = \alpha_1 \oplus \Pi_1\beta_1^\perp$ is not satisfied, and that the I(2) condition $\mathbb{C}^2 = \alpha_2 \oplus \Pi_2\beta_2^\perp$ is satisfied, so our autoregressive law of motion generates an I(2) process. As in Example 4.1, we have $\zeta = \{0\}$, so that (ignoring deterministic components) $(\langle x, X_t\rangle, t \geq 1)$ is I(1) for nonzero $x \in \beta_2$ and I(2) for nonzero $x \notin \beta_2$.

Our final example is infinite dimensional, and builds on Example 3.3 above.

Example 4.3
Consider the setting of Example 3.3. We can use Theorem 4.2 to determine whether we have a pole of second order. The Moore–Penrose inverse Π_0^+ satisfies $\Pi_0^+\Pi_0 = \mathrm{P}_{\beta_1}$. Applying Π_0^+ to both sides of the equality $(\mathrm{I} - \Phi_1)(-e_1) = e_2 + e_3$ reveals that $\mathrm{P}_{\beta_1}(-e_1) = \Pi_0^+(e_2 + e_3)$, which simplifies to $\Pi_0^+(e_2 + e_3) = -e_1$ since $\beta_1 = \mathrm{sp}\{e_2, e_3\}^\perp$. In view of Remark 4.7, it follows that

$$\Pi_2\beta_2^\perp = (\mathrm{I} - \Pi_0^+)(\alpha_1 \cap \beta_1^\perp) = (\mathrm{I} - \Pi_0^+)\mathrm{sp}\{e_2 + e_3\} = \mathrm{sp}\{e_1 + e_2 + e_3\}.$$

In view of Remark 4.2, we have

$$\alpha_2 = \alpha_1 + \beta_1^\perp = \mathrm{clsp}\{e_2 + e_3, e_4, e_5, \ldots\} + \mathrm{sp}\{e_2, e_3\} = \mathrm{clsp}\{e_2, e_3, e_4, \ldots\}.$$

Thus, we see that $\mathcal{H}$ is the sum of the linear subspaces α_2 and $\Pi_2\beta_2^\perp$, and deduce from Theorem 4.2 that $\Phi(z)^{-1}$ has a pole of second order at $z = 1$. The associated cointegrating space is

$$\beta_2 = \beta_1 + \alpha_1^\perp = \mathrm{sp}\{e_2, e_3\}^\perp + \mathrm{clsp}\{e_2 + e_3, e_4, e_5, \ldots\}^\perp = \mathrm{clsp}\{e_1, e_2 - e_3, e_4, \ldots\}.$$

The I(2) stochastic trend takes values in the orthogonal complement to this space, which is

$$\beta_2^\perp = \mathrm{sp}\{e_2 + e_3\}.$$

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

In view of Remark 4.9, the I(1) stochastic trend takes values in the larger space

$$\zeta^{\perp} = \beta_1^{\perp} + \Pi_0^{+}(\alpha_1 \cap \beta_1^{\perp}) = \mathrm{sp}\{e_2, e_3\} + \mathrm{sp}\{e_1\} = \mathrm{sp}\{e_1, e_2, e_3\}.$$

It follows that (ignoring deterministic components) for nonzero $x \in \mathcal{H}$ the inner product sequence $(\langle x, X_t \rangle, t \geq 1)$ is

I(0) for $x \in \mathrm{cl\,sp}\{e_4, e_5, \ldots\}$,

I(1) for $x \in \mathrm{cl\,sp}\{e_1, e_2 - e_3, e_4, \ldots\} \setminus \mathrm{cl\,sp}\{e_4, e_5, \ldots\}$,

I(2) for $x \notin \mathrm{cl\,sp}\{e_1, e_2 - e_3, e_4, \ldots\}$.

5. CONCLUDING REMARKS

The results established in this article extend long-known results on I(1) and I(2) representations for autoregressive processes in $\mathbb{C}^n$ to a more general Hilbert space setting. It may be desirable to extend our results further to a Banach space setting. In a Banach space setting, we no longer have the luxury of using inner products to formulate a suitable notion of cointegration. It is instead natural to define cointegration in terms of the order of integration of continuous linear functionals of an integrated process, and the cointegrating space as a subspace of the topological dual. An unpublished manuscript of the second author (Seo, 2018) provides some results in this direction, and also investigates relaxing the compactness condition we have imposed here on autoregressive operators. It is interesting to drop compactness because in this case the attractor space may be infinite dimensional; the difficulty is that the analytic Fredholm theorem becomes unavailable.

Beyond representation theory, research on the development of statistical procedures for analyzing cointegrated functional time series is a priority. A new manuscript by Nielsen, Seo, and Seong (2019) develops a procedure for estimating the dimension of an attractor space based on sequential variance ratio tests, with apparently good small sample behavior. If the size of the literature on cointegration in the finite dimensional setting provides any indication, there remains enormous scope for further research on estimation, testing, and forecasting with cointegrated functional time series.

REFERENCES

Al Sadoon, M. (2018) The linear systems approach to rational expectations models. *Econometric Theory* 34, 628–658.

Beare, B.K. (2017) The Chang-Kim-Park model of cointegrated density-valued time series cannot accomodate a stochastic trend. *Econ Journal Watch* 14, 133–137.

Beare, B.K. & W.-K. Seo (2017) Representation of I(1) autoregressive Hilbertian processes. ArXiv e-prints 1701.08149v1 [math.ST].

Beare, B.K., J. Seo, & W.-K. Seo (2017) Cointegrated linear processes in Hilbert space. *Journal of Time Series Analysis* 38, 1010–1027.

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

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

Bosq, D. (2000) *Linear Processes in Function Spaces*. Springer.
Cerovecki, C. & S. Hörmann (2017) On the CLT for discrete Fourier transforms of functional time series. *Journal of Multivariate Analysis* 154, 282–295.
Chan, N.H. & C.Z. Wei (1987) Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* 15, 1050–1063.
Chan, N.H. & C.Z. Wei (1988) Limiting distributions of least squares estimates of unstable autoregressive processes. *Annals of Statistics* 16, 367–401.
Chang, Y., B. Hu, & J.-Y. Park (2016) On the error correction model for functional time series with unit roots. Mimeo, Indiana University.
Chang, Y., C.-S. Kim, & J.-Y. Park (2016) Nonstationarity in time series of state densities. *Journal of Econometrics* 192, 152–167.
Cheng, X. & P.C.B. Phillips (2009) Semiparametric cointegrating rank selection. *Econometrics Journal* 12, S83–S104.
Cheng, X. & P.C.B. Phillips (2012) Cointegrating rank selection in models with time-varying variance. *Journal of Econometrics* 169, 155–165.
Davidson, J.E.H., D.F. Hendry, F. Srba, & S. Yeo (1978) Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom. *Economic Journal* 88, 661–692.
Engle, R.F. & C.W.J. Granger (1987) Co-integration and error correction: Representation, estimation and testing. *Econometrica* 55, 251–276.
Engsted, T. & S. Johansen (1999) Granger's representation theorem and multicointegration. In R.F. Engle & H. White (eds.), *Cointegration, Causality and Forecasting: Festschrift in Honour of Clive Granger*, pp. 200-211. Oxford University Press.
Faliva, M. & M.G. Zoia (2002) Matrix polynomials and their inversion: The algebraic framework of unit-root econometrics representation theorems. *Statistica* 62, 187–202.
Faliva, M. & M.G. Zoia (2009) *Dynamic Model Analysis*: *Advanced Matrix Methods and Unit-root Econometrics Representation Theorems*, 2nd ed. Springer.
Faliva, M. & M.G. Zoia (2011) An inversion formula for a matrix polynomial about a (unit) root. *Linear and Multilinear Algebra* 59, 541–556.
Franchi, M. (2007) The integration order of vector autoregressive processes. *Econometric Theory* 23, 546–553.
Franchi, M. & P. Paruolo (2011) Inversion of regular analytic matrix functions: Local Smith form and subspace duality. *Linear Algebra and its Applications* 435, 2896–2912.
Franchi, M. & P. Paruolo (2016) Inverting a matrix function around a singularity via local rank factorization. *SIAM Journal of Matrix Analysis and Applications* 37, 774–797.
Franchi, M. & P. Paruolo (2019a) Cointegration in functional autoregressive processes. *Econometric Theory*, doi: https://doi.org/10.1017/S0266466619000306.
Franchi, M. & P. Paruolo (2019b) A general inversion theorem for cointegration. *Econometric Reviews*, in press.
Gohberg, I., S. Goldberg, & M.A. Kaashoek (1990) *Classes of Linear Operators, vol. 1*. Birkhäuser.
Granger, C.W.J. (1981) Some properties of times series data and their use in econometric model specification. *Journal of Econometrics* 16, 121–130.
Granger, C.W.J. (1983) Cointegrated variables and error-correcting models. Mimeo, University of California, San Diego.
Granger, C.W.J. (1986) Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics* 48, 213–228.
Granger, C.W.J. & T.H. Lee (1989) Investigation of production, sales and inventory relationships using multicointegration and non-symmetric error correction models. *Journal of Applied Econometrics* 4, S145–S159.
Granger, C.W.J. & T.H. Lee (1990) Multicointegration. In G.F. Rhodes & T.B. Fomby (eds.), *Advances in Econometrics, Vol. 8*: *Co-integration, Spurious Regressions, and Unit Roots*, pp. 71–84. JAI Press.

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

Hörmann, S. & P. Kokoszka (2012) Functional time series. In T.S. Rao, S.S. Rao, & C.R. Rao (eds.), *Handbook of Statistics, Volume 30*: *Time Series Analysis—Methods and Applications*, chapter 7. pp. 157–186. North-Holland.

Howland, J.S. (1971) Simple poles of operator-valued functions. *Journal of Mathematical Analysis and Applications* 36, 12–21.

Hu, B. & J.-Y. Park (2016) Econometric analysis of functional dynamics in the presence of persistence. Mimeo, Indiana University.

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

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

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

Johansen, S. (1996) *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

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

Kurtz, T.G. & P. Protter (1991) Weak limit theorems for stochastic integrals and stochastic differential equations. *Annals of Probability* 19, 1035–1070.

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

Markus, A.S. (2012) *Introduction to the Spectral Theory of Polynomial Operator Pencils*. American Mathematical Society.

Nielsen, M. Ø., W.-K. Seo, & D. Seong (2019) Variance ratio test for the number of stochastic trends in functional time series. Queen's Economics Department Working paper No. 1420.

Park, J.-Y. & P.C.B. Phillips (1988) Statistical inference in regressions with integrated processes: Part I. *Econometric Theory* 4, 468–497.

Park, J.-Y. & P.C.B. Phillips (1989) Statistical inference in regressions with integrated processes: Part II. *Econometric Theory* 5, 95–131.

Phillips, P.C.B. (1986) Understanding spurious regressions in econometrics. *Journal of Econometrics* 33, 311–340.

Phillips, P.C.B. (1988) Regression theory for near-integrated time series. *Econometrica* 56, 1021–1043.

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

Phillips, P.C.B. & S.N. Durlauf (1986) Multiple time series regression with integrated processes. *Review of Economic Studies* 53, 473–495.

Phillips, P.C.B. & B.E. Hansen (1990) Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57, 99–125.

Phillips, P.C.B. & I. Kheifets (2019) On multicointegration. Mimeo, Yale University.

Phillips, P.C.B. & J.-Y. Park (1988) Asymptotic equivalence of ordinary least squares and generalized least squares in regressions with integrated regressors. *Journal of the American Statistical Association* 83, 111–115.

Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.

Sargan, J.D. (1964) Wages and prices in the United Kingdom: A study in econometric methodology. In P.E. Hart, G. Mills, & J.K. Whitacker (eds.), *Econometric Analysis for National Economic Planning*, pp. 23–63. Butterworths.

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

Seo, W.-K. (2018) Cointegration and representation of integrated autoregressive processes in function space. ArXiv e-prints 1712.08748v3 [math.FA].

Seo, W.-K. & B.K. Beare (2019) Cointegrated linear processes in Bayes Hilbert space. *Statistics and Probability Letters* 147, 90–95.

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

Steinberg, S. (1968) Meromorphic families of compact operators. *Archive for Rational Mechanics and Analysis* 31, 372–379.

Yoo, B.-S. (1987) Co-integrated time series: Structure, forecasting and testing. Doctoral thesis, University of California, San Diego.

APPENDIX A: Mathematical Appendix

A.1. Spectral Properties of Operator-Valued Functions

Consider a map $A : U \to \mathcal{L}_{\mathcal{H}}$, where U is some open connected subset of $\mathbb{C}$. We say that A is holomorphic on an open connected set $D \subseteq U$ if, for each $z_0 \in D$, the limit

$$A^{(1)}(z_0) := \lim_{z \to z_0} \frac{A(z) - A(z_0)}{z - z_0} \tag{A.1}$$

exists in the norm of $\mathcal{L}_{\mathcal{H}}$. It can be shown (Gohberg, Goldberg, and Kaashoek, 1990, pp. 7–8) that holomorphicity on D in fact implies analyticity on D, meaning that, for every $z_0 \in D$, we may represent A on D in terms of a power series

$$A(z) = \sum_{k=0}^{\infty} (z - z_0)^k A_k, \quad z \in D, \tag{A.2}$$

where $A_0, A_1, \ldots$ is a sequence in $\mathcal{L}_{\mathcal{H}}$ not depending on z. When the power series terminates, so that $A(z)$ is polynomial in z, then we say that A is an operator pencil.

The set of points $z \in U$ at which the operator $A(z)$ is noninvertible is called the spectrum of A, and denoted $\sigma(A)$. The spectrum is always a closed set, and if A is analytic on U, then $A(z)^{-1}$ depends analytically on $z \in U \setminus \sigma(A)$ (Markus, 2012, p. 56). A lot more can be said about $\sigma(A)$ and the behavior of $A(z)^{-1}$ if we assume that $A(z)$ is a Fredholm operator for every $z \in U$. In this case, we have the following result, a proof of which may be found in Gohberg, Goldberg, and Kaashoek (1990, pp. 203–204). It is a crucial input to our main results.

Analytic Fredholm Theorem. *Let $A : U \to \mathcal{L}_{\mathcal{H}}$ be an analytic Fredholm operator-valued function, and assume that $A(z)$ is invertible for some $z \in U$. Then, $\sigma(A)$ is at most countable and has no accumulation point in U. Furthermore, for $z_0 \in \sigma(A)$ and $z \in U \setminus \sigma(A)$ sufficiently close to z_0, we have*

$$A(z)^{-1} = \sum_{k=-m}^{\infty} (z - z_0)^k N_k, \tag{A.3}$$

where $m \in \mathbb{N}$ and $N_{-m}, N_{-m+1}, \ldots$ is a sequence in $\mathcal{L}_{\mathcal{H}}$ not depending on z. The operator N_0 is Fredholm of index zero and the operators $N_{-m}, \ldots, N_{-1}$ are of finite rank.

The analytic Fredholm theorem tells us that $A(z)^{-1}$ is analytic except at a discrete set of points, which are poles. The technical term for this property of $A(z)^{-1}$ is meromorphicity. For further reading on the spectral properties of operator-valued functions we suggest Gohberg, Goldberg, and Kaashoek (1990) and Markus (2012).

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

A.2. Proofs

Here, we provide proofs of all numbered results. The ordering of the proofs differs from the order in which the results were stated, so as to respect logical antecedence.

Proof of Theorem 3.2. It is obvious that (3) ⇒ (4), because when we write $\mathcal{H} = \alpha_1 \oplus \Pi_1 \beta_1^{\perp}$, we are asserting that $\mathcal{H} = \alpha_1 + \Pi_1 \beta_1^{\perp}$ and that $\alpha_1 \cap \Pi_1 \beta_1^{\perp} = \{0\}$. To establish the equivalence of the four conditions, it remains to show that (4) ⇒ (1) ⇒ (2) ⇒ (3). Under Assumption 3.1(ii), $\Phi(z)$ is analytic, Fredholm operator-valued, noninvertible at $z = 1$, and invertible at $z = 0$. The analytic Fredholm theorem therefore implies that $\Phi(z)^{-1}$ is analytic on a punctured neighborhood D of $z = 1$ with a pole of order d at $z = 1$, and for $z \in D$ admits a Laurent series around $z = 1$ as in (3.3). Furthermore, $\Phi(z)$ is analytic on $D \cup \{1\}$ and thus for $z \in D$ admits the Taylor series

$$\Phi(z) = \sum_{k=0}^{\infty} \frac{1}{k!} \Phi^{(k)}(1)(z-1)^k. \tag{A.4}$$

Combining (3.3) and (A.4) we obtain, for $z \in D$,

$$\mathrm{I} = \left(\sum_{k=-d}^{\infty} \Upsilon_k (z-1)^k \right) \left(\sum_{k=0}^{\infty} \frac{1}{k!} \Phi^{(k)}(1)(z-1)^k \right) \tag{A.5}$$

$$= \sum_{k=-d}^{\infty} \left(\sum_{j=0}^{d+k} \frac{1}{j!} \Upsilon_{k-j} \Phi^{(j)}(1) \right) (z-1)^k. \tag{A.6}$$

Suppose that condition (1) is false, meaning that $d > 1$. Then, the coefficients of $(z-1)^{-d}$ and $(z-1)^{-d+1}$ in the expansion of the identity in (A.6) must be zero. That is,

$$\Upsilon_{-d} \Pi_0 = 0, \tag{A.7}$$

$$\Upsilon_{-d+1} \Pi_0 + \Upsilon_{-d} \Pi_1 = 0. \tag{A.8}$$

Equation (A.7) implies that $\Upsilon_{-d} \alpha_1 = \{0\}$, whereas equation (A.8) implies that $\Upsilon_{-d} \Pi_1 \beta_1^{\perp} = \{0\}$. If the Condition (4) were valid, we could conclude that $\Upsilon_{-d} = 0$; however, this is impossible since Υ_{-d} is the leading coefficient in the Laurent series (3.3), which is nonzero by construction. Thus, if Condition (4) is true then Condition (1) must also be true: (4) ⇒ (1).

We next show that (1) ⇒ (2). Suppose that (1) is true, meaning that $d = 1$. The coefficients of $(z-1)^{-1}$ and $(z-1)^0$ in the expansion of the identity in (A.6) must be equal to 0 and I, respectively. Since $d = 1$, this means that

$$\Upsilon_{-1} \Pi_0 = 0, \tag{A.9}$$

$$\Upsilon_0 \Pi_0 + \Upsilon_{-1} \Pi_1 = \mathrm{I}. \tag{A.10}$$

It is apparent from (A.10) that $\Upsilon_{-1} \Pi_1(x) = x$ for all $x \in \beta_1^{\perp}$, and from (A.9) that $\Upsilon_{-1} \mathrm{P}_{\alpha_1} = 0$. We deduce that

$$\Upsilon_{-1} \mathrm{P}_{\alpha_1^{\perp}} \Pi_1(x) = x \quad \text{for all } x \in \beta_1^{\perp}. \tag{A.11}$$

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

This shows that Λ_1 is left-invertible, hence injective. The spaces $\alpha_1^{\perp}$ and $\beta_1^{\perp}$ are of equal and finite dimension due to the fact that Π_0 is Fredholm of index zero under Assumption 3.1(ii), so injectivity of Λ_1 implies bijectivity. Thus, we have shown that (1) $\Rightarrow$ (2).

We next show that (2) $\Rightarrow$ (3). This amounts to showing that (2) $\Rightarrow$ (4) and that (2) implies

$$\alpha_1 \cap \Pi_1 \beta_1^{\perp} = \{0\}. \tag{A.12}$$

Condition (2) implies that $\mathrm{P}_{\alpha_1^{\perp}} \Pi_1 \beta_1^{\perp} = \alpha_1^{\perp}$. We therefore have

$$\mathcal{H} = \alpha_1 + \mathrm{P}_{\alpha_1^{\perp}} \Pi_1 \beta_1^{\perp} = \alpha_1 + (\mathrm{I} - \mathrm{P}_{\alpha_1}) \Pi_1 \beta_1^{\perp}. \tag{A.13}$$

Since every element of $(\mathrm{I} - \mathrm{P}_{\alpha_1}) \Pi_1 \beta_1^{\perp}$ is the sum of an element of α_1 and an element of $\Pi_1 \beta_1^{\perp}$, this shows that every element of $\mathcal{H}$ can be written as the sum of an element of α_1 and an element of $\Pi_1 \beta_1^{\perp}$, and so it is proved that (2) $\Rightarrow$ (4). To establish that Condition (2) also implies (A.12), we observe that any element $y \in \alpha_1 \cap \Pi_1 \beta_1^{\perp}$ may be written as $y = \Pi_1(x)$ for some $x \in \beta_1^{\perp}$. Projecting both sides of this equality on $\alpha_1^{\perp}$ gives $0 = \mathrm{P}_{\alpha_1^{\perp}} \Pi_1(x)$. The bijectivity of Λ_1 asserted by Condition (2) thus requires us to have $x = 0$, implying that $y = 0$. Thus, (A.12) is proved under Condition (2), and we have shown that (2) $\Rightarrow$ (3).

It remains to verify that the residue Υ_{-1} is as claimed when Condition (1) is satisfied. We showed earlier in the proof that Condition (1) implies (A.11). We deduce from this that

$$\Upsilon_{-1} \mathrm{P}_{\alpha_1^{\perp}} \Pi_1 \Lambda_1^{-1} \mathrm{P}_{\alpha_1^{\perp}} = \Lambda_1^{-1} \mathrm{P}_{\alpha_1^{\perp}}, \tag{A.14}$$

or, more simply, $\Upsilon_{-1} \mathrm{P}_{\alpha_1^{\perp}} = \Lambda_1^{-1} \mathrm{P}_{\alpha_1^{\perp}}$. We also showed earlier in the proof that Condition (1) implies that $\Upsilon_{-1} \mathrm{P}_{\alpha_1} = 0$, or equivalently $\Upsilon_{-1} \mathrm{P}_{\alpha_1^{\perp}} = \Upsilon_{-1}$. Thus, $\Upsilon_{-1} = \Lambda_1^{-1} \mathrm{P}_{\alpha_1^{\perp}}$, as claimed. ■

Proof of Corollary 3.1. Since $\Pi_1 = -\Phi_1$ and $\Phi_1(x) = x$ for all $x \in \beta_1^{\perp}$, we must have $\Pi_1 \beta_1^{\perp} = \beta_1^{\perp}$. The equivalence of Conditions (1), (2), and (3) therefore follows from Theorem 3.2.

Obviously (2) $\Rightarrow$ (4). We will show that (4) $\Rightarrow$ (1) by showing that (4) implies bijectivity of Λ_1, which was established in Theorem 3.2 to be necessary and sufficient for a simple pole. The operator Λ_1 reduces in the case $p = 1$ to the restriction of $-\mathrm{P}_{\alpha_1^{\perp}}$ to $\beta_1^{\perp}$, so its kernel is $\alpha_1 \cap \beta_1^{\perp}$. Since the domain and codomain of Λ_1 are spaces of equal and finite dimension, Condition (4) therefore implies that Λ_1 is bijective. Thus, (4) $\Rightarrow$ (1).

It remains to show that, when $p = 1$ and we have a simple pole, $-\Lambda_1^{-1} \mathrm{P}_{\alpha_1^{\perp}}$ corresponds to projection on $\beta_1^{\perp}$ along α_1. Clearly $-\Lambda_1^{-1} \mathrm{P}_{\alpha_1^{\perp}}$ has kernel α_1 and range $\beta_1^{\perp}$, so it remains only to show idempotency. Since $-\Lambda_1$ is equal to the restriction of $\mathrm{P}_{\alpha_1^{\perp}}$ to $\beta_1^{\perp}$ when $p = 1$, we have

$$(-\Lambda_1^{-1} \mathrm{P}_{\alpha_1^{\perp}})(-\Lambda_1^{-1} \mathrm{P}_{\alpha_1^{\perp}}) = \Lambda_1^{-1}(-\Lambda_1)\Lambda_1^{-1} \mathrm{P}_{\alpha_1^{\perp}} = -\Lambda_1^{-1} \mathrm{P}_{\alpha_1^{\perp}},$$

as claimed. ■

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

Proof of Theorem 3.1. Under Assumption 3.1(ii), $\Phi(z)$ is analytic, Fredholm operator-valued, noninvertible at $z = 1$, and invertible elsewhere in the closed unit disk. The analytic Fredholm theorem therefore implies that $\Phi(z)^{-1}$ is analytic on an open disk centered at zero with radius exceeding one, except at the point $z = 1$, where it has a pole, which we have assumed to be simple. Let $\Psi(z) = (1-z)\Phi(z)^{-1}$, defined at $z = 1$ by analytic continuation.

The fact that $\Psi(z)$ and $\tilde{\Psi}(z)$ are analytic on an open disk centered at zero with radius exceeding one implies that the coefficients of their Taylor series around zero, $\Psi(z) = \sum_{k=0}^{\infty} \Psi_k z^k$ and $\tilde{\Psi}(z) = \sum_{k=0}^{\infty} \tilde{\Psi}_k z^k$, decay exponentially in norm. Under Assumption 3.1(i), the two series $\sum_{k=0}^{\infty} \Psi_k(\varepsilon_{t-k})$ and $\sum_{k=0}^{\infty} \tilde{\Psi}_k(\varepsilon_{t-k})$ thus converge in $L^2_{\mathcal{H}}$, the latter validly defining $\nu_t \in L^2_{\mathcal{H}}$. Applying the equivalent linear filters induced by $(1-z)\Phi^{-1}(z)$ and $\Psi(z)$ to either side of the equality $X_t - \sum_{j=1}^{p} \Phi_j(X_{t-j}) = \varepsilon_t$, we find that

$$\Delta X_t = \sum_{k=0}^{\infty} \Psi_k(\varepsilon_{t-k}), \quad t \geq 1, \tag{A.15}$$

a moving average representation for ΔX_t. Moreover, since $\Psi(z) = -\Upsilon_{-1} + (1-z)\tilde{\Psi}(z)$, we may rewrite (A.15) as

$$\Delta X_t = -\Upsilon_{-1}(\varepsilon_t) + \Delta \nu_t, \quad t \geq 1. \tag{A.16}$$

Clearly, the process given by

$$X_0^* = \nu_0, \quad X_t^* = -\Upsilon_{-1}\left(\sum_{s=1}^{t} \varepsilon_s\right) + \nu_t, \quad t \geq 1, \tag{A.17}$$

is a solution to the difference equation (A.16). It is completed by adding the solution to the homogeneous equation $\Delta X_t = 0$, which is any time invariant $Z_0 \in L^2_{\mathcal{H}}$. Therefore, we obtain the representation (3.5).

Since Υ_{-1} is the residue of $\Phi(z)^{-1}$ at $z = 1$, it is apparent from the formula $\Upsilon_{-1} = \Lambda_1^{-1} \mathrm{P}_{\alpha_1^{\perp}}$ given in Theorem 3.2 that the range of Υ_{-1} is $\beta_1^{\perp}$. Moreover, the range of Υ_{-1} is of positive and finite dimension by the analytic Fredholm theorem. This establishes Claim (1).

It remains to establish Claim (2). Supposing that Z_0 belongs to $\beta_1^{\perp}$, if we take the inner product of a nonzero $x \in \beta_1$ with both sides of (3.5) we obtain

$$\langle x, X_t \rangle = \langle x, \nu_t \rangle = \sum_{k=0}^{\infty} \hat{\Psi}_k^x(\varepsilon_{t-k}), \quad t \geq 1,$$

where the $\hat{\Psi}_k^x$'s are given by

$$\hat{\Psi}_k^x(y) = \langle x, \tilde{\Psi}_k(y) \rangle, \quad y \in \mathcal{H},$$

and decay exponentially in norm in $\mathcal{L}_{\mathcal{H},\mathbb{C}}$. (To see why, it may be helpful to recall Remark 3.3.) The sequence of inner products $(\langle x, X_t \rangle, t \geq 1)$ is thus a standard linear process in

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

$\mathbb{C}$ with innovations in $\mathcal{H}$. The innovation covariance operator Σ is positive definite under Assumption 3.1(i), and so $(\langle x, X_t\rangle, t \geq 1)$ is I(0) if $\sum_{k=0}^{\infty} \hat{\Psi}_k^x \neq 0$. Noting that

$$\sum_{k=0}^{\infty} \hat{\Psi}_k^x(y) = \left\langle x, \sum_{k=0}^{\infty} \tilde{\Psi}_k(y) \right\rangle, \quad y \in \mathcal{H},$$

we see that $\sum_{k=0}^{\infty} \hat{\Psi}_k^x = 0$ if and only if x belongs to the orthogonal complement to the range of $\sum_{k=0}^{\infty} \tilde{\Psi}_k$. But $\sum_{k=0}^{\infty} \tilde{\Psi}_k$ is the analytic part of the Laurent series of $\Phi(z)^{-1}$ around $z = 1$ evaluated at $z = 1$, which is Υ_0. We showed in the proof of Theorem 3.2 that (A.10) holds in the presence of a simple pole. From this equality, we deduce that the sum of the ranges of Υ_0 and Υ_{-1} is $\mathcal{H}$. It is therefore impossible for a nonzero $x \in \beta_1$ to belong to the orthogonal complement to the range of Υ_0, and we conclude that $\sum_{k=0}^{\infty} \hat{\Psi}_k^x \neq 0$. Thus, $(\langle x, X_t\rangle, t \geq 1)$ is I(0).

If we instead take the inner product of some $x \notin \beta_1$ with both sides of (3.5) we obtain

$$\langle x, X_t\rangle = \langle x, Z_0\rangle - \left\langle x, \Upsilon_{-1}\left(\sum_{s=1}^{t} \varepsilon_s\right)\right\rangle + \langle x, v_t\rangle, \quad t \geq 1.$$

Differencing yields

$$\begin{aligned}\langle x, \Delta X_t\rangle &= -\langle x, \Upsilon_{-1}(\varepsilon_t)\rangle + \langle x, \Delta v_t\rangle \\ &= -\hat{\Upsilon}_{-1}^x(\varepsilon_t) + \sum_{k=0}^{\infty} \hat{\Psi}_k^x(\varepsilon_{t-k}) - \sum_{k=0}^{\infty} \hat{\Psi}_{k+1}^x(\varepsilon_{t-k-1}), \quad t \geq 2,\end{aligned}$$

where the operator $\hat{\Upsilon}_{-1}^x \in \mathcal{L}_{\mathcal{H},\mathbb{C}}$ is defined by

$$\hat{\Upsilon}_{-1}^x(y) = \langle x, \Upsilon_{-1}(y)\rangle, \quad y \in \mathcal{H}.$$

Thus, we see that $(\langle x, \Delta X_t\rangle, t \geq 2)$ is a standard linear process in $\mathbb{C}$ with innovations in $\mathcal{H}$, with coefficients summing to $-\hat{\Upsilon}_{-1}^x$. Since x does not belong to β_1, the orthogonal complement to the range of Υ_{-1}, we know that $\hat{\Upsilon}_{-1}^x \neq 0$. Thus, $(\langle x, \Delta X_t\rangle, t \geq 2)$ is I(0) and $(\langle x, X_t\rangle, t \geq 1)$ is I(1). ■

Proof of Theorem 4.2. It is obvious that (3) ⇒ (4), so to establish the equivalence of the four conditions, we will show that (4) ⇒ (1) ⇒ (2) ⇒ (3).

To show (4) ⇒ (1), suppose that (1) is false; we will deduce that then (4) must also be false. Applying the analytic Fredholm theorem in the same way as in the proof of Theorem 3.2, we obtain identity (A.6) for z in a punctured neighborhood $D \subset U$ of $z = 1$. Here, d is the order of the pole of $\Phi(z)^{-1}$ at $z = 1$, and $\Upsilon_{-d} \neq 0$. A simple pole is ruled out by assumption, while a pole of second order is ruled out since we are maintaining that (1) is not satisfied, so we must have $d > 2$. The coefficients of $(z-1)^{-d}$, $(z-1)^{-d+1}$ and $(z-1)^{-d+2}$ in the expansion of the identity in (A.6) must therefore be zero. That is,

$$\Upsilon_{-d}\Pi_0 = 0, \tag{A.18}$$

$$\Upsilon_{-d+1}\Pi_0 + \Upsilon_{-d}\Pi_1 = 0, \tag{A.19}$$

$$\Upsilon_{-d+2}\Pi_0 + \Upsilon_{-d+1}\Pi_1 + \Upsilon_{-d}\Pi_2 + \Upsilon_{-d}\Pi_1\Pi_0^{+}\Pi_1 = 0. \tag{A.20}$$

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

From (A.18), we have $\Upsilon_{-d}\alpha_1 = 0$, implying that $\Upsilon_{-d}\alpha_2 = \Upsilon_{-d}\Pi_1\beta_1^{\perp}$. And from (A.19) we have $\Upsilon_{-d}\Pi_1\beta_1^{\perp} = 0$. Therefore,

$$\Upsilon_{-d}\alpha_2 = 0. \tag{A.21}$$

From (A.20) we have

$$\Upsilon_{-d}\Pi_2\beta_2^{\perp} = [\Upsilon_{-d+1} + \Upsilon_{-d}\Pi_1\Pi_0^{+}]\Pi_1\beta_2^{\perp}.$$

By composing both sides of (A.19) with Π_0^{+}, and noting that $\Pi_0\Pi_0^{+}$ coincides with the identity on α_1, we see that $\Upsilon_{-d+1} + \Upsilon_{-d}\Pi_1\Pi_0^{+}$ is zero on α_1. Since $\Pi_1\beta_2^{\perp}$ is contained in α_1, we deduce that

$$\Upsilon_{-d}\Pi_2\beta_2^{\perp} = 0. \tag{A.22}$$

If (4) were true then (A.21) and (A.22) would together imply that $\Upsilon_{-d} = 0$. But this is impossible because d is the order of our pole at $z = 1$ and the associated Laurent coefficient must be nonzero. Thus, (4) $\Rightarrow$ (1).

Next, we show that (1) $\Rightarrow$ (2). The coefficients of $(z-1)^{-2}$, $(z-1)^{-1}$, and $(z-1)^{0}$ in the expansion of the identity in (A.6) must be equal to 0, 0 and I, respectively. Suppose that (1) is true. Then, equations (A.18), (A.19), and (A.21) above hold with $d = 2$, and in place of (A.20) we have

$$\Upsilon_0\Pi_0 + \Upsilon_{-1}\Pi_1 + \Upsilon_{-2}\Pi_2 + \Upsilon_{-2}\Pi_1\Pi_0^{+}\Pi_1 = \mathrm{I}. \tag{A.23}$$

It is apparent from (A.23) that

$$\Upsilon_{-2}\Pi_2(x) = x - [\Upsilon_{-1} + \Upsilon_{-2}\Pi_1\Pi_0^{+}]\Pi_1(x) \quad \text{for all } x \in \beta_2^{\perp}.$$

As observed earlier, it follows from (A.19) that $\Upsilon_{-1} + \Upsilon_{-2}\Pi_1\Pi_0^{+}$ is zero on α_1. Since $\Pi_1\beta_2^{\perp}$ is contained in α_1, we deduce that $\Upsilon_{-2}\Pi_2$ coincides with the identity on $\beta_2^{\perp}$. Moreover, it follows from (A.21) that $\Upsilon_{-2}\Pi_2 = \Upsilon_{-2}\mathrm{P}_{\alpha_2^{\perp}}\Pi_2$. We conclude that $\Upsilon_{-2}\mathrm{P}_{\alpha_2^{\perp}}\Pi_2(x) = x$ for all $x \in \beta_2^{\perp}$. This shows that Λ_2 is left-invertible, hence injective. The spaces $\alpha_2^{\perp}$ and $\beta_2^{\perp}$ are of equal and finite dimension under Assumption 3.1(ii) (recall Remark 4.1), so injectivity of Λ_2 implies bijectivity. Thus, we have shown that (1) $\Rightarrow$ (2).

It can be shown that (2) $\Rightarrow$ (3) by arguing exactly as we did in the proof of the corresponding implication in Theorem 3.2, but replacing all subscripts of 1 with a subscript of 2.

It remains to verify our formulas for the Laurent coefficients Υ_{-2} and Υ_{-1}. Formula (4.3) for Υ_{-2} may be obtained by arguing exactly as we did in the proof of the residue formula in Theorem 3.2, but replacing all subscripts of 1 with a subscript of 2. Formula (4.4) for Υ_{-1} takes a bit more work. The coefficients of $(z-1)^{-2}$, $(z-1)^{-1}$, $(z-1)^{0}$, and $(z-1)^{1}$ in the expansion of the identity in (A.6) must be equal to 0, 0, I and 0, respectively. When $d = 2$, this yields equations (A.18), (A.19), and (A.23) above, as well as

$$\Upsilon_1\Pi_0 + \Upsilon_0\Pi_1 + \Upsilon_{-1}[\Pi_2 + \Pi_1\Pi_0^{+}\Pi_1] + \Upsilon_{-2}[\Pi_3 + \Pi_1\Pi_0^{+}\Pi_1\Pi_0^{+}\Pi_1] = 0. \tag{A.24}$$

Our Hilbert space $\mathcal{H}$ satisfies $\mathcal{H} = \alpha_1 \oplus (\alpha_2 \cap \alpha_1^{\perp}) \oplus \alpha_2^{\perp}$. Thus, $\Upsilon_{-1} = \Upsilon_{-1}\mathrm{P}_{\alpha_1} + \Upsilon_{-1}\mathrm{P}_{\alpha_2 \cap \alpha_1^{\perp}} + \Upsilon_{-1}\mathrm{P}_{\alpha_2^{\perp}}$. We shall proceed by obtaining expressions for the three terms on

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

the right-hand side of this equality separately. As observed earlier, it follows from (A.19) that $\Upsilon_{-1} + \Upsilon_{-2}\Pi_1\Pi_0^+$ is zero on α_1. The Moore–Penrose inverse Π_0^+ is zero on $\alpha_1^\perp$, so we have

$$\Upsilon_{-1}\mathrm{P}_{\alpha_1} = -\Upsilon_{-2}\Pi_1\Pi_0^+. \tag{A.25}$$

Next, we observe that $\Upsilon_{-1}\Lambda_1$ is equal to the restriction of $\Upsilon_{-1}\Pi_1 - \Upsilon_{-1}\mathrm{P}_{\alpha_1}\Pi_1$ to $\beta_1^\perp$. Consequently, Υ_{-1} coincides with $\Upsilon_{-1}\Pi_1\Lambda_1^+ - \Upsilon_{-1}\mathrm{P}_{\alpha_1}\Pi_1\Lambda_1^+$ on the range of Λ_1, which is $\alpha_2 \cap \alpha_1^\perp$, and so we have

$$\begin{aligned}\Upsilon_{-1}\mathrm{P}_{\alpha_2\cap\alpha_1^\perp} &= \Upsilon_{-1}\Pi_1\Lambda_1^+\mathrm{P}_{\alpha_2\cap\alpha_1^\perp} - \Upsilon_{-1}\mathrm{P}_{\alpha_1}\Pi_1\Lambda_1^+\mathrm{P}_{\alpha_2\cap\alpha_1^\perp} \\ &= \Upsilon_{-1}\Pi_1\Lambda_1^+\mathrm{P}_{\alpha_1^\perp} - \Upsilon_{-1}\mathrm{P}_{\alpha_1}\Pi_1\Lambda_1^+\mathrm{P}_{\alpha_1^\perp},\end{aligned} \tag{A.26}$$

with the second equality following from the fact that $\Lambda_1^+(\alpha_2^\perp \cap \alpha_1^\perp) = \{0\}$. It is apparent from (A.23) that

$$\Upsilon_{-1}\Pi_1\Lambda_1^+ = [\mathrm{I} - \Upsilon_{-2}\Pi_2 - \Upsilon_{-2}\Pi_1\Pi_0^+\Pi_1]\Lambda_1^+. \tag{A.27}$$

If we substitute into (A.26) our expressions for $\Upsilon_{-1}\Pi_1\Lambda_1^+$ and $\Upsilon_{-1}\mathrm{P}_{\alpha_1}$ appearing in (A.27) and (A.25), respectively, we obtain

$$\Upsilon_{-1}\mathrm{P}_{\alpha_2\cap\alpha_1^\perp} = [\mathrm{I} - \Upsilon_{-2}\Pi_2]\Lambda_1^+\mathrm{P}_{\alpha_1^\perp}. \tag{A.28}$$

We still require a suitable expression for $\Upsilon_{-1}\mathrm{P}_{\alpha_2^\perp}$. Our formula (4.3) reveals that the range of Υ_{-2} is contained in $\beta_1^\perp$, implying that $\Pi_0\Upsilon_{-2} = 0$. Therefore, by composing both sides of (A.24) with Υ_{-2}, we obtain

$$\Upsilon_0\Pi_1\Upsilon_{-2} + \Upsilon_{-1}[\Pi_2 + \Pi_1\Pi_0^+\Pi_1]\Upsilon_{-2} + \Upsilon_{-2}[\Pi_3 + \Pi_1\Pi_0^+\Pi_1\Pi_0^+\Pi_1]\Upsilon_{-2} = 0. \tag{A.29}$$

Our formula (4.3) also reveals that the range of $\Pi_1\Upsilon_{-2}$ is contained in α_1, implying that $\Pi_0\Pi_0^+\Pi_1\Upsilon_{-2} = \Pi_1\Upsilon_{-2}$. Therefore, by composing both sides of (A.23) with $\Pi_0^+\Pi_1\Upsilon_{-2}$, we obtain

$$\Upsilon_0\Pi_1\Upsilon_{-2} = [\mathrm{I} - \Upsilon_{-1}\Pi_1 - \Upsilon_{-2}\Pi_2 - \Upsilon_{-2}\Pi_1\Pi_0^+\Pi_1]\Pi_0^+\Pi_1\Upsilon_{-2}. \tag{A.30}$$

We may substitute the right-hand side of (A.30) for $\Upsilon_0\Pi_1\Upsilon_{-2}$ in (A.29). After some simplification this yields

$$\Upsilon_{-1}\Pi_2\Upsilon_{-2} + \Pi_0^+\Pi_1\Upsilon_{-2} + \Upsilon_{-2}[\Pi_3 - \Pi_2\Pi_0^+\Pi_1]\Upsilon_{-2} = 0. \tag{A.31}$$

Formula (4.3) implies that $\mathrm{P}_{\alpha_2^\perp}\Pi_2\Upsilon_{-2} = \mathrm{P}_{\alpha_2^\perp}$. We may therefore rewrite (A.31) as

$$\Upsilon_{-1}\mathrm{P}_{\alpha_2^\perp} = -\Upsilon_{-1}\mathrm{P}_{\alpha_2}\Pi_2\Upsilon_{-2} - \Pi_0^+\Pi_1\Upsilon_{-2} + \Upsilon_{-2}[\Pi_2\Pi_0^+\Pi_1 - \Pi_3]\Upsilon_{-2}. \tag{A.32}$$

We may substitute the sum of the right-hand sides of (A.25) and (A.28) for $\Upsilon_{-1}\mathrm{P}_{\alpha_2}$ in (A.32). After some simplification this yields

$$\begin{aligned}\Upsilon_{-1}\mathrm{P}_{\alpha_2^\perp} = &-[\Lambda_1^+\mathrm{P}_{\alpha_1^\perp}\Pi_2 + \Pi_0^+\Pi_1]\Upsilon_{-2} \\ &+ \Upsilon_{-2}[\Pi_1\Pi_0^+\Pi_2 + \Pi_2\Pi_0^+\Pi_1 + \Pi_2\Lambda_1^+\mathrm{P}_{\alpha_1^\perp}\Pi_2 - \Pi_3]\Upsilon_{-2}.\end{aligned} \tag{A.33}$$

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

Finally, if we sum the right-hand sides of (A.25), (A.28), and (A.33), we obtain the desired formula (4.4) for Υ_{-1}. ■

Proof of Theorem 4.1. Under Assumption 3.1(ii) we may apply the analytic Fredholm theorem to deduce that $\Phi(z)^{-1}$ is analytic on an open disk centered at zero with radius exceeding one, except at the point $z = 1$, where it has a pole, which we assume here to be of second order. Let $\Psi(z) = (1-z)^2\Phi(z)^{-1}$, defined at $z = 1$ by analytic continuation.

As in the proof of Theorem 3.1, the coefficients in the Taylor series $\sum_{k=0}^{\infty}\Psi_k z^k$ and $\sum_{k=0}^{\infty}\tilde{\Psi}_k z^k$ decay exponentially in norm, while $\sum_{k=0}^{\infty}\Psi_k(\varepsilon_{t-k})$ and $\sum_{k=0}^{\infty}\tilde{\Psi}_k(\varepsilon_{t-k})$ converge in $L^2_{\mathcal{H}}$, the latter validly defining $\nu_t \in L^2_{\mathcal{H}}$. Applying the equivalent linear filters induced by $(1-z)^2\Phi^{-1}(z)$ and $\Psi(z)$ to either side of the equality $X_t - \sum_{j=1}^{p}\Phi_j(X_{t-j}) = \varepsilon_t$, we find that

$$\Delta^2 X_t = \sum_{k=0}^{\infty}\Psi_k(\varepsilon_{t-k}), \quad t \geq 2, \tag{A.34}$$

a moving average representation for $\Delta^2 X_t$. Moreover, since $\Psi(z) = \Upsilon_{-2} - (1-z)\Upsilon_{-1} + (1-z)^2\tilde{\Psi}(z)$, we may rewrite (A.34) as

$$\Delta^2 X_t = \Upsilon_{-2}(\varepsilon_t) - \Upsilon_{-1}(\Delta\varepsilon_t) + \Delta^2\nu_t, \quad t \geq 2. \tag{A.35}$$

Clearly, the process given by

$$X_0^* = \nu_0, \quad X_t^* = \Upsilon_{-2}\left(\sum_{s=1}^{t}\sum_{r=1}^{s}\varepsilon_r\right) - \Upsilon_{-1}\left(\sum_{s=1}^{t}\varepsilon_s\right) + \nu_t, \quad t \geq 1, \tag{A.36}$$

is a solution to the difference equation (A.35). It is completed by adding the solution to the homogeneous equation $\Delta^2 X_t = 0$, which is $Z_0 + tZ_1$ for any time invariant $Z_0, Z_1 \in L^2_{\mathcal{H}}$. Therefore, we obtain the representation (4.2).

It was established in Theorem 4.2 that $\Upsilon_{-2} = \Lambda_2^{-1}\mathrm{P}_{\alpha_2^{\perp}}$, so the range of Υ_{-2} is $\beta_2^{\perp}$. Moreover, the range of Υ_{-2} is of positive and finite dimension by the analytic Fredholm theorem. This establishes Claim (1).

It remains to establish Claim (2). Supposing that Z_0 and Z_1 belong to $\beta_2^{\perp}$, if we take the inner product of a nonzero $x \in \beta_2$ with both sides of (4.2) we obtain

$$\langle x, X_t\rangle = -\left\langle x, \Upsilon_{-1}\left(\sum_{s=1}^{t}\varepsilon_s\right)\right\rangle + \langle x, \nu_t\rangle, \quad t \geq 1. \tag{A.37}$$

As in the proof of Theorem 3.1, $(\langle x, \nu_t\rangle, t \geq 1)$ is a standard linear process in $\mathbb{C}$ with innovations in $\mathcal{H}$, and is not I(0) if and only if x belongs to the orthogonal complement to the range of Υ_0. We showed in the proof of Theorem 4.2 that (A.23) holds in the presence of a pole of second order. From this equality we deduce that the sum of the ranges of Υ_0, Υ_{-1} and Υ_{-2} is $\mathcal{H}$. It is therefore impossible for a nonzero $x \in \beta_2$ to be orthogonal to both of the ranges of Υ_0 and Υ_{-1}. Thus, either the first term on the right-hand side of (A.37) is zero or I(1) and the second is I(0) (if x is not orthogonal to the range of Υ_0), or the first term is I(1) and the second is a standard linear process (if x is not orthogonal to the range of Υ_{-1}). In either case, $(\langle x, X_t\rangle, t \geq 1)$ is I(0) or I(1).

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press

If we instead take the inner product of some $x \notin \beta_2$ with both sides of (4.2), and twice difference, we obtain

$$\langle x, \Delta^2 X_t \rangle = \langle x, \Upsilon_{-2}(\varepsilon_t) \rangle - \langle x, \Upsilon_{-1}(\Delta \varepsilon_t) \rangle + \langle x, \Delta^2 \nu_t \rangle, \quad t \geq 3. \tag{A.38}$$

Similar to the proof of Theorem 3.1, we thereby find that $(\langle x, \Delta^2 X_t \rangle), t \geq 3)$ is a standard linear process in $\mathbb{C}$ with innovations in $\mathcal{H}$, with coefficients summing to $-\hat{\Upsilon}^x_{-2}$, where

$$\hat{\Upsilon}^x_{-2}(y) = \langle x, \Upsilon_{-2}(y) \rangle, \quad y \in \mathcal{H}.$$

Since x does not belong to β_2, we know that $\hat{\Upsilon}^x_{-2} \neq 0$. Thus, $(\langle x, \Delta^2 X_t \rangle, t \geq 3)$ is I(0) and $(\langle x, X_t \rangle, t \geq 1)$ is I(2). ■

https://doi.org/10.1017/S0266466619000276 Published online by Cambridge University Press