Diskussionsschriften

A Canonical Form for Unit Root Processes in the State Space Framework

Dietmar Bauer
Martin Wagner

03-12

July 2003

Universität Bern
Volkswirtschaftliches Institut
Gesellschaftstrasse 49
3012 Bern, Switzerland
Tel: 41 (0)31 631 45 06
Web: www.vwi.unibe.ch

A Canonical Form for Unit Root Processes in the State Space Framework

Dietmar Bauer *
Inst. for Econometrics,
Operations Research and System Theory
TU Wien
Argentinierstr. 8, A-1040 Wien

Martin Wagner†
Department of Economics
University of Bern
Gesellschaftsstrasse 49
CH-3012 Bern

July 1, 2003

*Correspondence to: tel.: ++43 +1 58801 11944, fax: ++43 +1 58801 11999, e-mail: Dietmar.Bauer@tuwien.ac.at

†Currently visiting the Department of Economics of Princeton University whose hospitality is gratefully acknowledged.

Abstract

In this paper we develop a canonical state space representation for rational stochastic processes containing unit roots with integer integration orders at arbitrary points on the unit circle. It is shown that the state space framework, which is – in a certain sense made precise in the paper – equivalent to the ARMA framework, is very suitable for the analysis of unit roots and cointegration issues. The advantages become especially prominent for systems with higher integration orders at the various roots on the unit circle. A unique state space representation is constructed that clearly reveals the integration and cointegration properties. The canonical form given in the paper can be used to construct a parameterization of the class of all rational processes with a given state space unit root structure, which is defined in the paper.

JEL Classification: C13, C32
Keywords: Canonical form, state space representation, unit roots, cointegration

1 Introduction

In modelling economic time series over the last several decades linear dynamic models incorporating unit roots and cointegration have become prominent tools. The literature in this area is primarily focused on (vector) AR or (vector) ARMA models, see e.g. Banerjee et al. (1993) or Johansen (1995). The state space framework, which is - in a specific sense discussed below - equivalent to the ARMA framework, has surprisingly not received a lot of attention for modelling unit root processes. Only a few earlier exceptions exist like e.g. Aoki (1990), Aoki and Havenner (1989), Aoki and Havenner (1997) or some of the contributions in the special issue on state space modelling in Econometric Reviews (1991). These contributions considering unit root analysis in the state space framework can to a certain extent be regarded as preliminary as they are focused only on the unit root $z = 1$ and on integration order 1. This shows an imbalance between the extent to which theory is developed for ARMA models as compared to state space models: For ARMA models representation theory exists for processes with unit roots possibly located at other points than $z = 1$ and of higher order of integration, see e.g. Gregoir (1999a). An eminent feature that emerges from this ARMA cointegration literature is the fact that for higher integration orders even the determination of the integration order (for an exact definition see below) tends to be inconvenient, see e.g. Johansen (1992) or again Gregoir (1999a). It is the purpose of this paper to show that the state space framework is a very convenient alternative to the ARMA framework for discussing unit root processes leading directly to simple and easily interpretable characterizations of the (co)integration properties.

In the discussion we allow for processes with a finite number of unit roots with integer integration orders at arbitrary points of the unit circle. The analysis is restricted to rational processes, in particular to ARMA processes, see the discussion below. Hence, issues like fractional (co)integration are not dealt with. Usually a process is defined to be a unit root process, if it constitutes a solution to a vector difference equation (VDE) with certain properties, see e.g. Johansen (1995, Chapter 3) for the vector AR case. The first point of the paper is to show that any unit root process can also be represented as a solution to suitably defined state space system equations, see below (6) in Section 3. Based upon this, the first main result of the paper is to show that for any given unit root process a unique

state space representation that very clearly reveals the integration and cointegration properties can be given. This unique representation can be used to classify state space systems according to their state space unit root structure. In the state space unit root structure i.a. the locations and orders of the unit roots are collected, see Definition 6 in Section 6 below.

The developed unique state space representation leads directly to Granger type representations and reveals the cointegrating relationships, either static, dynamic or polynomial, via orthogonality relationships. A detailed discussion of polynomial cointegration in the state space framework is outsourced to a companion paper, Bauer and Wagner (2003a). Thus, the representation in the state space framework is indeed more convenient than representation results derived in the AR or ARMA framework, see e.g. Johansen (1992), Johansen and Schaumburg (1999), Lee (1992), Stock and Watson (1993) or Gregoir (1999a). Also the classification of processes according to their state space unit root structure is simpler than the corresponding classification in the ARMA framework.

The second advantage is that the developed canonical form can be used to construct a parameterization. Based on the results presented in this paper Bauer and Wagner (2003b) present one specific parameterization that possesses relevant topological properties important for subsequent statistical analysis. The parameterization can be used e.g. to perform maximum likelihood analysis for classes of processes with fixed state space integration structure. First results are derived in Bauer and Wagner (2002b), where pseudo maximum likelihood estimators and their asymptotic distribution are derived for processes where the integration orders corresponding to all unit roots are equal to 1 (cf. Remark 3). Up to now, estimation for processes with a *general* unit root structure seems to have been limited to autoregressive processes cf. e.g. Gregoir (1999b).

The paper is organized as follows: Section 2 presents the class of processes dealt with in this paper and gives the relevant definitions. Section 3 discusses state space representations. Section 4 links and compares the state space representation to the VAR framework for I(1) processes. Section 5 demonstrates the key ideas using some illustrative examples, while the presentation of the canonical form is given in Section 6. Finally Section 7 concludes the paper. The proof of the main result is contained in the Appendix.

Throughout the paper the following notation is used: I_d denotes the $d \times d$ dimensional

identity matrix, $0^{a\times b}$ the $a \times b$ dimensional null matrix. x' denotes the complex conjugate transpose of a complex vector and $\overline{x}$ the complex conjugate, unless stated otherwise. For a matrix $C \in \mathbb{C}^{s\times c}, c \leq s$, a matrix whose columns span the orthogonal complement of the space spanned by the columns of the matrix C is denoted by $C_\perp \in \mathbb{C}^{s\times(s-c)}$. The ambient space is complex-valued if not stated otherwise explicitly. With abuse of notation we will call a matrix $C \in \mathbb{C}^{s\times c}$ orthonormal, if $C'C = I_c$.

2 The Class of Processes Considered

Unit root processes are usually discussed in a vector ARMA framework. In the following we review the basic facts about ARMA processes needed in the paper. For a detailed discussion in the stationary case refer to Hannan and Deistler (1988) in particular chapters 1 and 2. An s-dimensional process $(y_t)_{t\in\mathbb{N}}$ is an ARMA process, if there exist matrices $A_j \in \mathbb{R}^{s\times s}, j = 1,\ldots,p$ and $B_j \in \mathbb{R}^{s\times s}, j = 1,\ldots,q$, $A_p \neq 0, B_q \neq 0$, a white noise process $(\varepsilon_t)_{t\in\mathbb{N}}$ and (possibly random) initial conditions $y_t, t = 1-p,\ldots,0$ and $\varepsilon_t, t = 1-q,\ldots,0$, such that

$$y_t + A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} = \varepsilon_t + B_1\varepsilon_{t-1} + \ldots + B_q\varepsilon_{t-q}, t \in \mathbb{N}. \tag{1}$$

Considering the polynomials $a(z) = I_s + A_1 z + \ldots + A_p z^p$ and $b(z) = I_s + B_1 z + \ldots + B_q z^q$, the pair $(a(z), b(z))$ is called an ARMA system corresponding to the process $(y_t)_{t\in\mathbb{N}}$. The representation of $(y_t)_{t\in\mathbb{N}}$ as the solution to the vector difference equation (1) is called ARMA representation of $(y_t)_{t\in\mathbb{N}}$. Here z denotes a complex variable. Letting z also denote the backward shift operator [1] the ARMA system equations can compactly (but somewhat loosely) be written as $a(z)y_t = b(z)\varepsilon_t$. It is well known that the solution set to the system of equations (1) is given by the sum of one particular solution and the solutions to the homogenous equation $a(z)y_t = 0, t \in \mathbb{N}$. We call a process $(y_t)_{t\in\mathbb{N}}$ *linearly deterministic*, if $\sup_{t\geq t_0} \|y_t - y_{t|t_0}\| = 0$ for some $t_0 \in \mathbb{N}$, where $y_{t|t_0}$ denotes the best linear prediction of y_t given $y_{t_0}, y_{t_0-1}, \ldots, y_1$. Hence, linearly deterministic processes are perfectly predictable from some time instant on. Therefore typical deterministic terms such as the constant, a polynomial trend or seasonal dummies are linearly deterministic and so are all solutions

[1] Note, that the definition of z requires the setting of initial values $y_0, \varepsilon_0, y_{-1}, \ldots$. Hence the backward-shift operator depends on initial values. However, this is neglected in the notation.

to the homogenous equation $a(z)y_t = 0$.

The ARMA representation of $(y_t)_{t\in\mathbb{N}}$ is not unique: If $(y_t)_{t\in\mathbb{N}}$ solves the ARMA equation corresponding to $(a(z), b(z))$ and $p(z) = I_s + P_1 z + \ldots + P_k z^k$ denotes a polynomial matrix, then $(y_t)_{t\in\mathbb{N}}$ also solves the ARMA equations corresponding to $(p(z)a(z), p(z)b(z))$. If $a(z) = p(z)\tilde{a}(z), b(z) = p(z)\tilde{b}(z)$, then any solution to the equation $\tilde{a}(z)y_t = \tilde{b}(z)\varepsilon_t, t \in \mathbb{N}$ differs from any solution to $a(z)y_t = b(z)\varepsilon_t$ only by a linearly deterministic process z_t, such that $a(z)z_t = 0$. Left coprime pairs $(a(z), b(z))$ play a special role: In this case the roots of $\det a(z)$ determine the stationarity respectively nonstationarity properties of the process $(y_t)_{t\in\mathbb{N}}$. If all roots of $\det a(z)$ in a left coprime pair $(a(z), b(z))$ are outside the unit circle, then there exist initial conditions such that the corresponding solution is stationary. For stationary processes it is well known that it is no restriction of generality to consider only left coprime representations, where moreover the roots of $\det b(z)$ are outside the open unit disc, i.e. $\det b(z) \neq 0, |z| < 1$. The pair of polynomials $(a(z), b(z))$ in this case corresponds to the rational transfer function $k(z) = a(z)^{-1}b(z) = \sum_{j=0}^{\infty} K_j z^j$, which converges on the closed unit disc. In this case the stationary solution $(y_t)_{t\in\mathbb{N}}$ possesses a representation for white noise $(\varepsilon_t)_{t\in\mathbb{Z}}$ as

$$y_t = \sum_{j=0}^{\infty} K_j \varepsilon_{t-j}, t \in \mathbb{N} \tag{2}$$

and corresponds to a special choice of initial conditions. If in a left coprime pair $(a(z), b(z))$ some of the roots of $\det a(z)$ are on the unit circle, the solutions to the corresponding vector difference equations are called unit root processes, which are formally defined below. Let us first define the difference operator at frequency ω:

$$\Delta_\omega(z) = \begin{cases} 1 - e^{i\omega}z, & \omega \in \{0, \pi\} \\ (1 - e^{i\omega}z)(1 - e^{-i\omega}z), & \omega \in (0, \pi). \end{cases} \tag{3}$$

Note that for real valued processes complex roots occur in pairs of complex conjugate roots. Therefore the way we define the differencing operator $\Delta_\omega(z) = (1 - e^{i\omega}z)(1 - e^{-i\omega}z) = 1 - 2(\cos\omega)z + z^2$ already incorporates the assumption of real valued y_t by filtering pairs of complex conjugate roots.

In order to simplify the notation we will use $\Delta := \Delta_0(z)$. Note that the application of a filter of degree q (polynomial degree in the complex variable z) to a process $(y_t)_{t\in\mathbb{N}}$ necessitates q initial conditions to be specified. We are now ready to define unit root processes:

Definition 1 *The s-dimensional real process* $(y_t)_{t\in\mathbb{N}}$ *has unit root structure* $((\omega_1, h_1), \ldots, (\omega_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}))$ *with* $0 \leq \omega_1 < \omega_2 < \ldots < \omega_{l^{\mathbb{R}}} \leq \pi, h_k \in \mathbb{N}, k = 1, \ldots, l^{\mathbb{R}}$, *if there exist random initial values* $y_{1-H}, \ldots, y_0, H = \sum_{k=0}^{l^{\mathbb{R}}}[h_k + h_k \mathbb{I}(\omega_k \notin \{0, \pi\})]$ *with finite second moments and a linearly deterministic process* $(T_t)_{t\in\mathbb{N}}$ *such that*

$$\Delta_{\omega_1}^{h_1}(z) \cdots \Delta_{\omega_{l^{\mathbb{R}}}}^{h_{l^{\mathbb{R}}}}(z) y_t = v_t + T_t, \quad t \in \mathbb{N} \tag{4}$$

for $v_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}$ *corresponding to the Wold representation of the stationary process* $(v_t)_{t\in\mathbb{Z}}$, *where for* $c(z) = \sum_{j=0}^{\infty} c_j z^j$ *with* $\sum_{j=0}^{\infty} \|c_j\| < \infty$ *it holds that* $c(e^{i\omega_k}) \neq 0$ *for* $k = 1, \ldots, l^{\mathbb{R}}$. *Here* $\mathbb{I}(.)$ *denotes the indicator function.*
If $c(z)$ *is a rational function of* z, *then* $(y_t)_{t\in\mathbb{N}}$ *is called a* rational process.

Remark 1 *The assumptions on* $c(z)$ *rule out a number of cases, such as e.g. long memory processes or processes with discontinuous spectral densities. For a detailed discussion of these issues see e.g. Leeb and Pötscher (2001). In this paper we study only rational processes, i.e. processes for which a finite dimensional state space representation exists. Therefore the difficulties encountered in Leeb and Pötscher (2001) do not occur.*

Remark 2 *The requirement that* $c(e^{i\omega_k}) \neq 0$ *only at the points at which filtering has taken place and not for all points on the unit circle allows to classify processes of the form* $y_t + y_{t-2} = \varepsilon_t - \varepsilon_{t-1}$, *with* ε_t *as in the above definition. In the light of the above definition this process* y_t *has integration structure* $\left((\frac{\pi}{2}, 1)\right)$, *but would not be covered by the definition if* $c(z) \neq 0$ *were required for all* $|z| = 1$.

The definition just given is formulated to discuss differencing simultaneously at pairs of complex conjugate unit roots. If one wants to disentangle the effects on y_t of each of a member of a pair of complex conjugate unit roots, the definition can be reformulated counting each complex unit root separately. The corresponding definition of a complex unit root structure is also needed in order to keep the presentation of the results to follow algebraically as simple as possible. Disentangling the contributions to each unit root allows to simplify the algebra of the arguments by using complex quantities.

Definition 2 *The s-dimensional random process* $(y_t)_{t\in\mathbb{N}}$ *has complex unit root structure* $((\omega_1, h_1), \ldots, (\omega_l, h_l))$ *with* $z_k = e^{i\omega_k}$, $0 \leq \omega_1 < \omega_2 < \ldots < \omega_l < 2\pi$ *and* $h_k \in \mathbb{N}$ *for*

$k = 1, \ldots, l$, if there exist random initial conditions $y_{1-H}, \ldots, y_0, H = h_1 + \cdots + h_l$ with finite second moments and a linearly deterministic process $(T_t)_{t\in\mathbb{N}}$ such that

$$\prod_{k=1}^{l}(1 - z_k z)^{h_k} y_t = v_t + T_t, \quad t \in \mathbb{N} \tag{5}$$

with $v_t = c(z)\varepsilon_t$ as in Definition 1.

Here and throughout the rest of the paper the unit roots $\omega_{l^{\mathbb{R}}+1}, \ldots, \omega_l$ denote the unit roots with negative imaginary part, i.e. with frequency in the interval $(\pi, 2\pi)$. For real valued processes for each k such that $\omega_k \in (0, \pi)$ there exists a k^* such that $\omega_{k^*} \in (\pi, 2\pi)$ and with $\omega_{k^*} = 2\pi - \omega_k$. In later sections, to distinguish notationally, we will use the term *complex integrated of order h_k at z_k* if the pair (ω_k, h_k) is contained in the unit root structure, where $z_k = e^{i\omega_k}$.

Remark 3 *Note that in the above definitions the integration orders are defined for the s-dimensional process, and not for the individual components. It follows, however, directly from the definition that for all unit roots ω_k there is at least one component of y_t that is integrated of order h_k at this unit root frequency. This stems from the requirement $c(e^{i\omega_k}) \neq 0 \ \forall \ k = 1, \ldots, l$. Alternatively unit root processes could be defined for scalar processes in total analogy. For a multivariate process the unit root structure could then be defined as the maximum integration order of all components, leading to exactly the same classification as the one given above.*

Remark 4 *Note that due to our definition of the unit root structure deterministic terms are subtracted before defining the integration orders. This implies e.g. that so called trend stationary processes are not integrated according to this definition. It does also not hold that the first differences of processes with unit root structure $((0, 1))$ are stationary. In general they may contain a linearly deterministic component T_t.*

Processes where the only unit root is located at $z = 1$ are as usual called integrated of order m, say, see the following definition:

Definition 3 *A real valued process with unit root structure $((0, m))$ is called I(m) process. A real valued process with unit root structure $((\omega_1, 1), \ldots, (\omega_l, 1))$ is called multiple frequency I(1) process, or short MFI(1) process.*

Let us next define cointegration. For processes with higher integration orders and with unit roots at different points on the unit circle a multitude of possibilities for cointegration and polynomial cointegration of different orders arises. In the following definitions it is understood that pairs of the form $(\omega_k, 0)$ are removed from the unit root structure of the transformed processes. For the vector polynomial $\beta(z) = \sum_{j=0}^{q} \beta_j z^j, \beta_j \in \mathbb{R}^s$ let $\beta(z)'y_t = \sum_{j=0}^{q} \beta_j' y_{t-j}$, where $y_t = 0$ is used for $t < 1$.

Definition 4 *The real valued s-dimensional random process $(y_t)_{t\in\mathbb{N}}$ with unit root structure $((\omega_1, h_1), \ldots, (\omega_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}))$ is called* cointegrated of order $\big((\omega_1, h_1, h_1^p), \ldots, (\omega_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}^p)\big)$, $0 \leq h_k^p \leq h_k, k = 1, \ldots, l^{\mathbb{R}}$, *where* $\max_{k=1,\ldots,l^{\mathbb{R}}}(h_k - h_k^p) > 0$, *if there exists a vector* $\beta \in \mathbb{R}^s, \beta \neq 0$ *such that $(\beta' y_t)_{t\in\mathbb{N}}$ has unit root structure $\big((\omega_1, h_1^p), \ldots, (\omega_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}^p)\big)$. The vector β is in this case called* cointegrating vector of order $\big((\omega_1, h_1, h_1^p), \ldots, (\omega_l, h_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}^p)\big)$. *The real random process $(y_t)_{t\in\mathbb{N}}$ with integration structure $((\omega_1, h_1), \ldots, (\omega_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}))$ is called* polynomially cointegrated of order $\big((\omega_1, h_1, h_1^p), \ldots, (\omega_l, h_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}^p)\big)$, $0 \leq h_k^p \leq h_k, k = 1, \ldots, l^{\mathbb{R}}$, *with* $\max_{k=1,\ldots,l^{\mathbb{R}}}(h_k - h_k^p) > 0$, *if there exists a vector polynomial* $\beta(z) = \sum_{j=0}^{q} \beta_j z^j$, $\beta_j \in \mathbb{R}^s$ *with* $\max_{k=1,\ldots,l^{\mathbb{R}}} \|\beta(z_k)\|(h_k - h_k^p) > 0$, *such that $(\beta(z)'y_t)_{t\in\mathbb{N}}$ has unit root structure $\big((\omega_1, h_1^p), \ldots, (\omega_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}^p)\big)$. The vector polynomial $\beta(z)$ is in this case called* polynomial cointegrating vector of order $\big((\omega_1, h_1, h_1^p), \ldots, (\omega_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}^p)\big)$.

Remark 5 *Analogously to the definition of complex integration, also the definition of cointegration can be extended to complex static and complex dynamic cointegration, by considering the complex unit root structure as the basis in Definition 4 and allowing for complex coefficients $\beta_j \in \mathbb{C}^s$.*

Remark 6 *The condition that a polynomial cointegrating vector has to fulfill the constraint* $\max_{k=1,\ldots,l^{\mathbb{R}}} \|\beta(z_k)\|(h_k - h_k^p) > 0$ *excludes trivial polynomial cointegrating relationships which achieve the reductions in the integration orders simply due to differencing. For a detailed investigation of polynomial cointegration see Bauer and Wagner (2003a). In that paper it is shown that the canonical representation developed in this paper also forms the basis for an elegant and simple analysis of polynomial cointegration in the state space framework.*

In the paper we only deal with rational processes, where hence $\tilde{a}(z)v_t = b(z)\varepsilon_t$ for some $(\tilde{a}(z), b(z))$ such that $\tilde{a}(z)^{-1}b(z) = c(z)$ holds, i.e. where $(v_t)_{t\in\mathbb{N}}$ is a stationary ARMA

process. If furthermore $T_t = 0$ holds, then $(y_t)_{t\in\mathbb{N}}$ is an ARMA process with corresponding ARMA system $(\prod_{k=1}^{l}(1 - z_k z)^{h_k}\tilde{a}(z), b(z))$. Therefore every rational unit root process $(y_t)_{t\in\mathbb{N}}$ is the sum of an ARMA process and a solution to the vector difference equation $\prod_{k=1}^{l}(1 - z_k z)^{h_k} z_t = T_t$ for identical initial values.

3 State Space Realizations

It is well known (see e.g. Hannan and Deistler 1988, Chapter 1) that for every stationary ARMA process $(y_t)_{t\in\mathbb{N}}$ corresponding to the ARMA system $(a(z), b(z))$ there exists an equivalent representation, the so called *state space* representation: For $t \in \mathbb{N}$

$$\begin{aligned} y_t &= Cx_t + \varepsilon_t \\ x_{t+1} &= Ax_t + B\varepsilon_t. \end{aligned} \tag{6}$$

Here $x_t \in \mathbb{C}^n$ denotes the n-dimensional unobserved state vector. The initial conditions x_1 are set in order to render the processes $(x_t)_{t\in\mathbb{N}}$ and $(y_t)_{t\in\mathbb{N}}$ jointly stationary. $A \in \mathbb{C}^{n\times n}, B \in \mathbb{C}^{n\times s}, C \in \mathbb{C}^{s\times n}$ are complex matrices. Usually it is assumed that x_t, A, B and C are real rather than complex, however for some of the following results the use of complex quantities simplifies the algebra and the interpretation considerably. To ensure that y_t is real valued, the matrices A, B and C have to fulfill certain restrictions, which will be commented upon below (cf. Theorem 2). For a given ARMA representation of a rational process one way to obtain a (specific) state space representation corresponding to this process is e.g. given on page 15 of Hannan and Deistler (1988).

The first equation in (6) is termed *observation equation* and the second *state equation*. The dynamics of the system are contained in the state equation. Note that the state is by construction an AR(1) process. It follows that

$$y_t = Cx_t + \varepsilon_t = CAx_{t-1} + CB\varepsilon_{t-1} + \varepsilon_t = \cdots = CA^{t-1}x_1 + \varepsilon_t + \sum_{j=0}^{t-2} CA^j B\varepsilon_{t-j-1}. \tag{7}$$

For stable A, i.e. A is such that $|\lambda_{max}(A)| < 1$, where $\lambda_{max}(A)$ denotes an eigenvalue of maximum modulus, the choice $x_1 = \sum_{j=0}^{\infty} A^j B\varepsilon_{-j}$ (again using white noise $(\varepsilon_t)_{t\in\mathbb{Z}}$) leads to a stationary process $(y_t)_{t\in\mathbb{N}}$ with the representation (cf. (2))

$$y_t = \varepsilon_t + \sum_{j=0}^{\infty} CA^j B\varepsilon_{t-j-1}.$$

Consequently $K_0 = I_s$ and $K_j = CA^{j-1}B, j > 0$ holds. This implies that the transfer function $k(z) = a^{-1}(z)b(z)$ can alternatively be written as $k(z) = I_s + zC(I_n - zA)^{-1}B$. In this case the system (A, B, C) is called *a state space realization of the transfer function* $k(z)$. Letting S_n denote the set of all triples of complex matrices (A, B, C) of appropriate dimension we can define the mapping $\Pi : \bigcup_{n\geq 0} S_n \to U, (A, B, C) \mapsto k(z) = I_s + zC(I_n - zA)^{-1}B$. Here U denotes the set of all $s \times s$ dimensional matrix valued functions, where each entry is a rational function of the complex variable z.

Like ARMA representations also state space representations are not unique. There are two sources of non-uniqueness. For each transfer function there exist state space realizations of different state dimension. A state space system (A, B, C) with state dimension n is called *minimal*, if there exists no realization $(\tilde{A}, \tilde{B}, \tilde{C})$ with state dimension $\tilde{n}$, such that $\Pi(A, B, C) = \Pi(\tilde{A}, \tilde{B}, \tilde{C})$ and $n > \tilde{n}$. The second source of non-uniqueness relates to the choice of the basis of the state. Since the state is not observed this basis can be chosen arbitrarily. A change of coordinates using a nonsingular transformation $T \in \mathbb{C}^{n\times n}$ leads to a different realization (TAT^{-1}, TB, CT^{-1}) of the same transfer function. The two different realizations are called *observationally equivalent*. Given a particular system (A, B, C) the set $\Pi^{-1}(k(z)) \subset \bigcup_{n\geq 0} S_n$ with $k(z) = I_s + zC(I_n - zA)^{-1}B$ is called *observational equivalence class*.

Both sources of non-uniqueness can be investigated using the so called *Hankel* matrix $\mathcal{H} = [K_{i+j-1}]_{i,j\in\mathbb{N}}$. As seen above for state space systems $K_{i+j-1} = CA^{i+j-2}B$ and hence the Hankel matrix can be factored into $\mathcal{H} = \mathcal{O}\mathcal{C}, \mathcal{O} = [C', A'C', (A^2)'C', \ldots]', \mathcal{C} = [B, AB, A^2B, \ldots]$. It is well known (cf. e.g. Theorem 2.4.1. in Hannan and Deistler 1988) that the transfer function $k(z)$ is rational, if and only if the rank of $\mathcal{H}$ is finite. In this case, the rank of $\mathcal{H}$ is equal to the dimension of any minimal state space realization. Therefore the rank of the Hankel matrix is called the *order* of the state space system and the corresponding transfer function. It can be shown that the order coincides with the McMillan degree of the transfer function (Hannan and Deistler 1988, Theorem 2.4.1.). This implies that a state space system is minimal, if and only if $\mathcal{O}$ and $\mathcal{C}$ are of full rank. A change of coordinates using a nonsingular transformation matrix T is equivalent to a different factorization of the Hankel matrix into $\mathcal{O}$ and $\mathcal{C}$ as $\mathcal{H} = \mathcal{O}\mathcal{C} = [\mathcal{O}T^{-1}][T\mathcal{C}]$.

The set of solutions to the system equations (6) has the same structure as the set of

solutions of an ARMA system, since it is in effect constituted by an AR(1) equation for the state and a linear mapping, the observation equation. Hence any solution to $y_t - Cx_t = \varepsilon_t, x_{t+1} - Ax_t = B\varepsilon_t, t \in \mathbb{N}$ is given by the sum of one particular solution and a solution to the equations $y_t - Cx_t = 0, x_{t+1} - Ax_t = 0, t \in \mathbb{N}$, which is simply given by $y_t = CA^{t-1}x_1$ for some initial conditions x_1. This decomposition has already been given in (7). Note that $(\tilde{T}_t)_{t\in\mathbb{N}} = (CA^{t-1}x_1)_{t\in\mathbb{N}}$ is a linearly deterministic process.

For ARMA processes with left coprime pairs $(a(z), b(z))$ the locations of the roots of $\det a(z)$ are known to determine the stochastic properties of the corresponding ARMA processes. For minimal state space processes the eigenvalues of A have the same function. Theorem 1.2.2. of Hannan and Deistler (1988) states that the nonzero eigenvalues of A are the inverses of the roots of $\det a(z)$. Hence the condition $|\lambda_{max}(A)| < 1$ is equivalent to the stability assumption $\det a(z) \neq 0, |z| \leq 1$. Here $\lambda_{max}(A)$ denotes an eigenvalue of maximum modulus. Eigenvalues of A on the unit circle hence imply unit roots in the corresponding solution processes. Analogously, the condition $\det b(z) \neq 0, |z| < 1$ is equivalent to the condition $|\lambda_{max}(A - BC)| \leq 1$ (minimum-phase assumption).

The first lemma shows that all unit root processes as given in Definition 1 have a state space representation.

Lemma 1 *For every rational process $(y_t)_{t\in\mathbb{N}}$ with unit root structure $((\omega_1, h_1), \ldots, (\omega_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}))$ there exists a linearly deterministic process $(d_t)_{t\in\mathbb{N}}$, such that the process $(y_t - d_t)_{t\in\mathbb{N}}$ has a state space representation (6). By choosing the linearly deterministic process $(d_t)_{t\in\mathbb{N}}$ appropriately the state space realization (A, B, C) can be assumed minimal.*

Conversely for minimal (A, B, C), every solution to the state space equations (6), where $|\lambda_{max}(A)| = 1$ is a rational process with unit root structure $((\omega_1, h_1), \ldots, (\omega_l, h_l))$, where $z_k = e^{i\omega_k}$ are the eigenvalues of A of unit modulus and $h_k > 0$ are suitable integers.

Proof: Rational processes with some unit root structure are defined as solutions to a difference equation. In the difference equation two terms on the right hand side have been included: A stationary process $v_t = \tilde{a}(z)^{-1}b(z)\varepsilon_t$ and a linearly deterministic process T_t. Due to the linearity of the vector difference equation the solution y_t can be represented as $y_t = y_t^1 + y_t^2$ where y_t^1 is a solution to the equation $\Delta_{\omega_1}^{h_1}(z) \ldots \Delta_{\omega_{l^{\mathbb{R}}}}^{h_l^{\mathbb{R}}}(z) y_t^1 = v_t$

and y_t^2 is a solution to the equation $\Delta_{\omega_1}^{h_1}(z)\dots\Delta_{\omega_{l^{\mathbb{R}}}}^{h_l^{\mathbb{R}}}(z)y_t^2 = T_t$. Denoting the latter solution as d_t (which clearly is linearly deterministic) it suffices to show, that every solution to the difference equation $\Delta_{\omega_1}^{h_1}(z)\dots\Delta_{\omega_{l^{\mathbb{R}}}}^{h_l^{\mathbb{R}}}(z)y_t^1 = v_t$ has a state space representation. By the definition of a rational process, v_t is a stationary ARMA process and hence denotes the stationary solution to the equation $\tilde{a}(z)v_t = b(z)\varepsilon_t, t \in \mathbb{N}$. Therefore $v_t = k(z)\varepsilon_t$ for $k(z) = \tilde{a}^{-1}(z)b(z) = \Pi(\tilde{A}, \tilde{B}, \tilde{C})$ for some state space realization $(\tilde{A}, \tilde{B}, \tilde{C})$. Hence $v_t = \tilde{C}\tilde{x}_t + \varepsilon_t, \tilde{x}_{t+1} = \tilde{A}\tilde{x}_t + \tilde{B}\varepsilon_t, t \in \mathbb{Z}$. Choosing $x_t = [(y_{t-1}^1)', (y_{t-2}^1)', \dots, (y_{t-H}^1)', \tilde{x}_t']', t \in \mathbb{N}$ it follows that $y_t^1 = Cx_t + \varepsilon_t, x_{t+1} = Ax_t + B\varepsilon_t, t \in \mathbb{N}$ for suitably defined (A, B, C), see the discussion in Section 4 below and Hannan and Deistler (1988, p. 15) for details. This shows the first part of the lemma.

Concerning minimality of the state space representation, it follows from the representation $y_t - d_t = CA^{t-1}x_1 + \varepsilon_t + \sum_{j=0}^{t-2} CA^jB\varepsilon_{t-j-1}$ that incorporating $CA^{t-1}x_1$ into d_t and noting that for linearly deterministic d_t also $d_t + CA^{t-1}x_1$ is linearly deterministic, the contribution of the state space system (A, B, C) to y_t is determined only by the impulse response sequence CA^jB and the noise ε_t. This directly implies that minimality is no restriction as this sequence is identical for all (minimal and non-minimal) realizations in the observational equivalence class $\Pi^{-1}(k(z))$.

Conversely, let y_t denote the solution to the state space equations (6) for some minimal system (A, B, C) and initial state x_1. The matrix A is similar to its Jordan normal form J, i.e. there exists a nonsingular transformation $T \in \mathbb{C}^{n\times n}$, such that $TAT^{-1} = J$. Let the eigenvalues be ordered according to their modulus in decreasing order. This ordering implies that $J = \text{diag}(A_1, A_{st})$, where in $A_1 \in \mathbb{C}^{c\times c}$ all eigenvalues of unit modulus are collected[2] and A_{st} contains the remaining stable ones. Denote the corresponding transformed system as $(J, \hat{B}, \hat{C})$, i.e. $\hat{B} = TB, \hat{C} = CT^{-1}$. Let $x_t = [x_{t,1}', x_{t,st}']', x_{t,1} \in \mathbb{C}^c$ and $\hat{B} = [B_1', B_{st}']', \hat{C} = [C_1, C_{st}]$ be partitioned accordingly. Let $a_1(z)$ denote the monic polynomial of minimum degree such that $a_1(A_1^{-1}) = 0$. Since for $\det(I_c - zA_1)$ it holds that $\det(I_c - A_1^{-1}A_1) = 0$ such a polynomial exists and it divides $\det(I_c - zA_1)$. Hence the degree, q say, of $a_1(z)$ is at most equal to c. Therefore the roots of $a_1(z)$ are a subset of the inverses of the eigenvalues of A_1, which are by definition of unit magnitude. Hence $a_1(z) = (-1)^d\Delta_{\omega_1}^{h_1}(z)\dots\Delta_{\omega_{l^{\mathbb{R}}}}^{h_l^{\mathbb{R}}}(z)$ for some unit root structure $((\omega_1, h_1), \dots, (\omega_{l^{\mathbb{R}}}, h_{l^{\mathbb{R}}}))$, where

[2]I.e. c is the sum of the algebraic multiplicities of all eigenvalues of unit modulus.

$d = 0$ or $d = 1$. From the block-diagonal structure $CA^jB = C_1A_1^jB_1 + C_{st}A_{st}^jB_{st}$ follows immediately. Consider $a_1(z)y_t$ for $t > q$:

$$a_1(z)y_t = a_1(z)(C_1x_{t,1} + C_{st}x_{t,st} + \varepsilon_t).$$

Since $|\lambda_{max}(A_{st})| < 1$, an initial state $x_{1,st}$ can be chosen such that $x_{t,st}$ is stationary. Every other solution to the state space equation differs only by a linearly deterministic term. Hence the stationarity of $x_{t,st}$ can be assumed without restriction of generality. Consequently also $a_1(z)x_{t,st}$ is stationary and by assumption ε_t is stationary. Since every matrix fulfills its characteristic equation, we obtain from the definition of $x_{t,1}$

$$\begin{aligned}
&a_1(z)x_{t,1} = x_{t,1} + \alpha_1 x_{t-1,1} + \ldots + \alpha_q x_{t-q,1} = \\
&A_1^q x_{t-q,1} + \textstyle\sum_{j=0}^{q-1} A_1^j B_1 \varepsilon_{t-j-1} + \alpha_1(A_1^{q-1}x_{t-q,1} + \sum_{j=0}^{q-2} A_1^j B_1 \varepsilon_{t-j-2}) + \ldots + \alpha_q x_{t-q,1} \\
&= (A_1^q + \alpha_1 A_1^{q-1} + \ldots + \alpha_q I_q)x_{t-q,1} + \textstyle\sum_{j=0}^{q-1} G_j \varepsilon_{t-j-1} = \sum_{j=0}^{q-1} G_j \varepsilon_{t-j-1}
\end{aligned}$$

since $A_1^q + \alpha_1 A_1^{q-1} + \ldots + \alpha_q I_n = A_1^q(I_n + \alpha_1 A_1^{-1} + \alpha_q A_1^{-q}) = A_1^q a_1(A_1^{-1}) = 0$ and thus $a_1(z)x_{t,1}, t > q$ is seen to be an MA(q-1) process with the coefficients G_j defined by the above equations and hence in particular stationary. For $1 \leq t \leq q$ equality can be obtained by a suitable choice of T_t. Over-differencing of $x_{t,1}$ by the filtering is ruled out by minimality of the degree of $a_1(z)$. From the definition of $a_1(z)$ it can also be observed that the unit root structure is given by the factorization of the minimal A_1-annihilating polynomial. □

The above lemma shows that also in the unit root case the state space representation covers the same class of processes as the ARMA representation. The lemma in addition states that it is no restriction of generality to consider only minimal representations. Therefore in the remainder of the paper we are only concerned with minimal state space systems. Note here again that the linearly deterministic component d_t is included for two purposes: It absorbs by appropriate choice all effects of the initial values ($CA^{t-1}x_1$) and it allows to incorporate deterministic components like constants and deterministic trends and cycles in the analysis.

The major aim of this paper is to provide a unique state space representation for all unit root processes that clearly reveals the integration and cointegration properties. The representation result is based on so called canonical forms:

Definition 5 *A canonical form of the set $M \subset U$ is a mapping $\phi : M \to \bigcup_{j\geq 0} S_j : k(z) \mapsto (A, B, C)$.*

In words thus a canonical form selects for each transfer function $k(z) \in M$ one representative (A, B, C) of the class of observationally equivalent systems $\Pi^{-1}(k(z))$ corresponding to $k(z)$. This implies that for every state space realization there exists one unique observationally equivalent realization in the image of the canonical form. Usually canonical forms are defined on subsets of U. The partitioning of U is determined by characteristics of the contained transfer functions. In this paper the discussion is limited to the sets M_n of transfer functions $k(z)$ of order n, where $\det k(z) \neq 0, |z| < 1$ and $k(z)$ has no pole for $|z| < 1$. Thus, the canonical form will be defined on $\bigcup_{n\in\mathbb{N}} M_n$.
There is a variety of possibilities of constructing such a mapping (see e.g. Hannan and Deistler, 1988, Chapter 2, for a discussion of the so called *echelon* canonical form). In this paper we develop a specific canonical form for state space systems corresponding to unit root processes, which has the special feature that the orders of integration corresponding to the solutions to a given state space system can be directly read off from the canonical form. The possibility to construct such a representation is seen to be an advantage of the state space framework over the ARMA framework.

4 The I(1) VAR Models in the State Space Framework

The overwhelming majority of the literature on unit root processes is concerned with $I(1)$ processes, often studied in a VAR framework (cf. e.g. Johansen 1995, and the references contained therein). In this section we exemplify the links between the autoregressive and the state space framework using the notation of Johansen (1995) for VAR processes. Relating to the discussion in Section 2, $(y_t)_{t\in\mathbb{N}}$ is an s-dimensional (here $s > 1$) VAR process, if there exist matrices $\Pi_j \in \mathbb{R}^{s\times s}, j = 1, \ldots, p$ such that the pair of polynomial matrices $a(z) = I_s - \Pi_1 z - \ldots - \Pi_p z^p, b(z) = I_s$ corresponds to an ARMA representation of the process $(y_t)_{t\in\mathbb{N}}$. In this case

$$y_t = \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \ldots + \Pi_p y_{t-p} + \varepsilon_t, \quad t \in \mathbb{N}$$

for initial conditions $y_{1-p}, \ldots, y_0$. One state space representation for $(y_t)_{t\in\mathbb{N}}$ is derived by defining the state $x_t = [y'_{t-1}, \ldots, y'_{t-p}]' \in \mathbb{R}^{ps}$. This leads to the following minimal

representation:

$$
\begin{aligned}
y_t &= \underbrace{\begin{bmatrix} \Pi_1 & \Pi_2 & \dots & \Pi_p \end{bmatrix}}_{C} x_t + \varepsilon_t, \\
x_{t+1} &= \underbrace{\begin{bmatrix} \Pi_1 & \Pi_2 & \dots & \dots & \Pi_p \\ I_s & 0 & \cdots & \cdots & 0 \\ 0 & I_s & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_s & 0 \end{bmatrix}}_{A} x_t + \underbrace{\begin{bmatrix} I_s \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{B} \varepsilon_t.
\end{aligned}
$$

Note that the above state equation resembles the so called companion form of the VAR (see e.g. page 15 of Johansen 1995). In the sequel we describe a set of basis transformations that parallels the transformation of the above VAR to its *error correction representation* (see e.g. page 45 of Johansen 1995). Starting with the transformation matrix

$$
T = \begin{bmatrix} I_s & 0 & \dots & 0 \\ I_s & -I_s & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ 0 & \dots & I_s & -I_s \end{bmatrix}
$$

the state is transformed to $\tilde{x}_t = Tx_t = [y'_{t-1}, \Delta y'_{t-1}, \dots, \Delta y'_{t-p+1}]'$. This results by construction in the new state being the stacked vector of all lagged regressors from the VAR error correction representation. It follows from straightforward computations that this basis change transforms the system matrices to:

$$
\tilde{A} = \left[\begin{array}{c|cccccc} I_s + \Pi & \Gamma_1 & \dots & \dots & \dots & \Gamma_{p-1} \\ \hline \Pi & \Gamma_1 & \dots & \dots & \dots & \Gamma_{p-1} \\ 0 & I_s & 0 & \dots & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & I_s & 0 \end{array}\right], \tilde{B} = \left[\begin{array}{c} I_s \\ \hline I_s \\ 0 \\ \vdots \\ 0 \end{array}\right], \tilde{C} = \left[\begin{array}{c|ccc} I_s + \Pi & \Gamma_1 & \dots & \Gamma_{p-1} \end{array}\right]
$$

where $\Pi = -I_s + \sum_{j=1}^{p} \Pi_j, \Gamma_i = -\sum_{j=i+1}^{p} \Pi_j$ for $i = 1, \dots, p-1$. It is well known that the process is integrated, if Π is not of full rank. As in Johansen (1995) let $\Pi = \alpha\beta', \alpha, \beta \in \mathbb{R}^{s \times r}$, where α and β are assumed to be of full column rank and $0 < r < s$ is assumed here to exclude the boundary cases of an integrated process with no cointegration ($r = 0$) and a stationary process ($r = s$). To ensure that the process is integrated only of order 1 we

assume additionally that $\alpha'_\perp(I_s - \sum_{i=1}^{p-1}\Gamma_i)\beta_\perp$ has full rank (cf. e.g. Johansen 1995, p. 49). Then there exist r linearly independent cointegrating relationships and consequently the number of common trends is equal to $c = s - r$. Note that after this transformation the observation equation reads already as $y_t = \tilde{C}\tilde{x}_t + \varepsilon_t = (I_s + \Pi)y_{t-1} + \sum_{i=1}^{p-1}\Gamma_i\Delta y_{t-i} + \varepsilon_t$. The first block-row of the state equation is of the same form and the second block-row of the state equations reproduces the VAR error correction form.

Next use a specific choice for $\beta_\perp$ such that $\beta'\beta_\perp = 0$ and $\beta'_\perp\beta_\perp = I_c$ and define $\bar{\beta} = \beta(\beta'\beta)^{-1}$. Denote with $S = [\beta_\perp, \beta]'$, with inverse $S^{-1} = [\beta_\perp, \bar{\beta}]$. Use the matrix S to perform another basis change with $\breve{T} = \operatorname{diag}\left[S, I_{(p-1)s}\right]$. This results in the transformed state and system matrices:

$$\breve{x}_t = [(\beta'_\perp y_{t-1})', (\beta' y_{t-1})', \Delta y'_{t-1}, \ldots, \Delta y'_{t-p+1}]',$$

$$\breve{A} = \left[\begin{array}{c||c|cccc} I_c & \beta'_\perp\alpha & \beta'_\perp\Gamma_1 & \ldots & \ldots & \ldots & \beta'_\perp\Gamma_{p-1} \\ \hline\hline 0 & I_r + \beta'\alpha & \beta'\Gamma_1 & \ldots & \ldots & \ldots & \beta'\Gamma_{p-1} \\ \hline 0 & \alpha & \Gamma_1 & \ldots & \ldots & \ldots & \Gamma_{p-1} \\ \vdots & 0 & I_s & 0 & \ldots & \ldots & 0 \\ \vdots & \vdots & 0 & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & I_s & 0 \end{array}\right], \breve{B} = \left[\begin{array}{c} \beta'_\perp \\ \hline\hline \beta' \\ \hline I_s \\ 0 \\ \vdots \\ 0 \end{array}\right],$$

$$\breve{C} = \left[\begin{array}{c||c|ccc} \beta_\perp & \alpha + \bar{\beta} & \Gamma_1 & \ldots & \Gamma_{p-1} \end{array}\right].$$

The matrix A is in the above representation already seen to have the eigenvalue 1 with multiplicity greater or equal to c. According to the assumptions on $a(z)$ it follows that $\lambda_{max}(A) = 1$. Thus, analyze the eigenvalues of the south-east block of the matrix $\breve{A}$, separated by the double lines, $\breve{A}_{st}$ say. We show by an indirect argument that under the I(1) assumption $|\lambda_{max}(\breve{A}_{st})| < 1$ holds. Therefore, let $z = [z'_1, z'_2, \ldots, z'_p]', z_1 \in \mathbb{R}^r, z_j \in \mathbb{R}^s, j \geq 2$ denote a vector such that $\breve{A}_{st}z = z$. It follows from the structure of $\breve{A}_{st}$ that $z_2 = \ldots = z_p$. Further one obtains from the first two block-rows that $z_2 = \alpha z_1 + \sum_{j=1}^{p-1}\Gamma_j z_2$ and $z_1 = z_1 + \beta'\alpha z_1 + \beta'\sum_{j=1}^{p-1}\Gamma_j z_2 = z_1 + \beta' z_2$. This shows that $(I_s - \sum_{j=1}^{p-1}\Gamma_j)z_2 = \alpha z_1$ and $\beta' z_2 = 0$. This system of equations has a nonzero solution if and only if $\alpha'_\perp(I_s - \sum_{i=1}^{p-1}\Gamma_i)\beta_\perp$ is rank deficient. Hence under the condition of no integration of order higher than 1 it follows that $|\lambda_{max}(\breve{A}_{st})| < 1$.

The state space representation $(\breve{A}, \breve{B}, \breve{C})$ has a direct interpretation in terms of the quantities used in the VAR cointegration literature. Let us start with the state equation. The

first block-row, above the double lines, of the state equation gives the dynamic behavior of the c common trends $\breve{x}_{t+1,1} = \beta'_\perp y_t$. It can easily be shown, and is well known from the VAR cointegration literature, that this part of the state is an I(1) process with no cointegration. The second block-row gives the dynamics of the cointegrating relationships, $\beta' y_t$ and thus describes for suitable initialization a stationary process. The first block-row of the remaining equations (below the single line) replicates the error correction representation and the latter rows are shifting the lagged differences in order to capture the short-run dynamics of $(y_t)_{t\in\mathbb{N}}$ in the AR(1) state equation.
Inserting in the observation equation leads to

$$y_t = \breve{C}\breve{x}_t + \varepsilon_t = \beta_\perp \beta'_\perp y_{t-1} + (\alpha + \bar{\beta})\beta' y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t$$

which decomposes in the specific normalization chosen the impact from the stochastic trends, the stationary directions and the lagged differences on y_t.
The above transformations separate in the state equation the nonstationary and the stationary components. In the representation developed in this paper we go further. By one further transformation of the basis of the state we can obtain nonstationary state components that follow a vector random walk, i.e. whose first difference is white noise. In the above example this means that a further basis change has to be performed that gives the A-matrix a block-diagonal structure with the $c \times c$ identity matrix in the upper corner. This then completely separates the unit root dynamics from the stationary dynamics, due to the block-diagonal structure. Concerning the part of the system matrices corresponding to the stationary components of the state it is then with analogous basis changes possible to transform it in the corresponding Jordan normal form as well. However, such a transformation does not seem to be of much value for the discussion in this section and hence we refrain from presenting the details in this respect.

5 Some Illustrative Examples

In the previous sections already a few important observations concerning the behavior of solutions to the system equations (6) have been made. In particular the relationship between the eigenvalues of A and the integration properties of the process y_t being the

solution of equations (6) is important: For minimal systems the presence of eigenvalues of A on the unit circle is equivalent to integration of the process y_t at the various corresponding frequencies, see Lemma 1. One main goal of this paper is to provide a unique state space representation for rational unit root processes. The main strategy in this respect is to successively impose restrictions on the system (A, B, C), such that the set of observationally equivalent systems, which furthermore fulfill all imposed restrictions is increasingly restricted and ultimately is a singleton. It has been shown in Lemma 1 that it is no restriction of generality to consider only minimal state space systems.

The eigenvalues are directly seen when the matrix A is given in its Jordan normal form, see e.g. Meyer (2000). For this reason, as we want to construct a canonical form that clearly reveals the integration and cointegration properties, we restrict ourselves to system representations, where the A-matrix is in (reordered) Jordan normal form.[3] Since any square matrix A is related to its Jordan normal form via $TAT^{-1} = J$, for suitable nonsingular T, there always exist observationally equivalent realizations, where the matrix A is in (reordered) Jordan normal form. However, as we will see below, restricting A to be in (reordered) Jordan normal form is in general not a sufficient restriction to achieve identification. In other words, usually there exist more than one minimal realization of a given transfer function $k(z)$ of the form (J, B, C), with J denoting a matrix in (reordered) Jordan normal form. Hence state space representations with the matrix A in (reordered) Jordan normal form are not unique. Therefore further restrictions have to be imposed in order to characterize a unique realization.

In this section the issue of imposing appropriate restrictions is discussed for several examples that are intended to reveal the type of restrictions that can be put in place in such a way that the cointegration properties of the system are clearly revealed. A secondary aim of this section is to present the canonical state space representation of the most commonly used unit root processes.

[3]We will see in Section 6 that for processes where identical Jordan blocks appear in the Jordan normal form a specific reordering of the Jordan normal form to obtain the so called *reordered* Jordan normal form is very convenient.

5.1 Example 1: I(1) Processes

Let us start with the simplest but most important example, with cointegration in I(1) processes. In this case all unit roots are located at $z = 1$, i.e. in the Jordan normal form only one Jordan segment (using the notation of Meyer, 2000) is present in the nonstationary part. All Jordan blocks composing the Jordan segment have to be of size one, i.e. there are no generalized eigenvectors corresponding to the eigenvalue 1, to ensure that the process is indeed $I(1)$, cf. Archontakis (1998), Theorem 1 below or the discussion for I(1) VAR processes in the previous section. This is also documented by the fact that in this case the minimal annihilating polynomial must be equal to $\Delta = (1 - z)$. As a counterexample consider the following state equation that leads to higher integration orders:

$$x_{t+1} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} x_t + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \varepsilon_t.$$

Denoting the coordinates of x_t by $x_{t,1}, x_{t,2}$ and $x_{t,3}$ respectively, we obtain: $x_{t+1,3} = x_{t,3} + \varepsilon_t, x_{t+1,2} = x_{t,2} + x_{t,3}, x_{t+1,1} = x_{t,1} + x_{t,2}$. Using the difference operator we directly obtain $\Delta x_{t,3} = \varepsilon_t, \Delta x_{t,2} = x_{t,3}, \Delta x_{t,1} = x_{t,2}$. Both representations directly show that $x_{t,3}$ is integrated of order one, $x_{t,2}$ is integrated of order two and $x_{t,1}$ is integrated of order three. This example leads to two observations: First, the integration order is equal to the size of the Jordan block, compare again to Archontakis (1998). Second, the components of the state corresponding to a Jordan block are linked to each other in a chain of subsequently increasing integration orders. It is easy to see that the same relationships hold also for higher dimensions and also for unit roots not equal to $z = 1$.

For a minimal state space representation it can be shown (cf. the proof of Theorem 1) that the output has the same unit root structure as the state. This shows that for minimal representations the integration orders can be deduced from the state equation. However, for non-minimal system representations this equivalence does not prevail.

Thus, for I(1) processes the A-matrix in Jordan normal form is given by $J = \mathrm{diag}(I_c, J_{st})$, i.e. the multiplicity of the unit root is denoted by c. J_{st} corresponds to the eigenvalues smaller than one in absolute value.[4] Partition the matrices B and C accordingly,

[4]The following discussion will make clear that the stationary part can be dealt with using any canonical form for stationary state space models. Thus, we transform it to Jordan normal form only for simplifying the discussion at this point.

i.e. $B = [B_1', B_{st}']'$ and $C = [C_1, C_{st}]$ with $C_1 \in \mathbb{C}^{s\times c}$ and $B_1 \in \mathbb{C}^{c\times s}$. The Granger representation of y_t (see e.g. Engle and Granger 1987) is in this particular case, for $x_{1,1} = 0, x_{1,st} = \sum_{j=0}^{\infty} A_{st}^j B_{st}\varepsilon_{-j}$, directly given by solving the system equations (6) for $y_t = C_1 B_1 \sum_{j=1}^{t-1} \varepsilon_j + k_{st}(z)\varepsilon_t$, where $k_{st}(z) = I_s + zC_{st}(I - A_{st}z)^{-1}B_{st}$ denotes the transfer function of the stationary part of the system. Hence we see that the use of the Jordan normal form separates the nonstationary from the stationary part of the system. It will be seen below that the dynamics corresponding to different unit roots are decoupled from each other for the same reason, when the A-matrix is in Jordan normal form respectively reordered Jordan normal form. Now, due to the fact that in minimal representations the matrices C_1 and B_1 both have full rank (see again the proof of Theorem 1 below), the number of common trends in x_t is restricted to be smaller or equal to the dimension of the output, i.e. $c \leq s$. Furthermore, from the above Granger type representation we directly see that the cointegrating space for y_t is the orthogonal complement to the column space of C_1, henceforth denoted by $C_1^{\perp}$. This directly implies also that the dimension of the cointegrating space, r say, is equal to $s - c$. Thus, in minimal state space representations of I(1) processes the well known duality, for I(1) processes, between the number of common trends and the dimension of the cointegrating space is directly seen.

The above system representation is however not unique. Look only at the nonstationary part again. Any nonsingular transformation matrix $T \in \mathbb{R}^{c\times c}$ to transform the nonstationary part, delivers an observationally equivalent realization (I_c, TB_1, C_1T^{-1}) of the nonstationary part of the system, which also has its A-matrix in Jordan normal form. This implies that further restrictions have to be imposed in order to achieve uniqueness of the system representation. These additional restrictions can be put on the matrices B_1 and C_1. A specific way of imposing a sufficient set of restrictions on the matrices C_1 and B_1 in order to identify them from the product C_1B_1 is described in the following lemma. Note at this point as a remark that this is only one possible way to achieve identifiability.

Lemma 2 *Let $X \in \mathbb{C}^{s\times s}$ denote a matrix of rank c. Then there exists a unique decomposition $X = CB, C \in \mathbb{C}^{s\times c}, B \in \mathbb{C}^{c\times s}$, such that $C'C = I_c$ and B is positive upper triangular. Here a matrix $B = [b_{i,j}]_{i=1,\ldots,c,j=1,\ldots,s}$ is called positive upper triangular (p.u.t.), if there exist indices $1 \leq j_1 < j_2 < \ldots < j_c \leq s$, such that $b_{i,j} = 0, j < j_i, b_{i,j_i} > 0$. I.e. B is of the*

form

$$\begin{bmatrix} 0 & \cdots & 0 & b_{1,j_1} & x & & \dots & & x \\ 0 & & \dots & & 0 & b_{2,j_2} & x & & \\ & & & & & & & & \\ 0 & & & \dots & & & 0 & b_{c,j_c} & x \end{bmatrix} \tag{8}$$

where x denotes arbitrary entries.

Proof: Consider the SVD of $X = USV'$, where $U, V \in \mathbb{C}^{s\times c}, S \in \mathbb{R}^{c\times c}$. Choosing $\tilde{C} = U$ leads to $\tilde{C}'\tilde{C} = I_c$. It is straightforward to show that for $\tilde{B} = SV'$ there exists exactly one orthonormal matrix $Q \in \mathbb{C}^{c\times c}$, such that $B = Q\tilde{B}$ is in positive upper triangular (p.u.t.) form (see e.g. Ober 1996). Then $C = \tilde{C}Q'$ and B fulfill all assumptions of the lemma. The uniqueness of Q implies uniqueness of C and B. □

Due to the block-diagonal structure for the stationary part any canonical form developed for stationary processes can be employed, e.g. the echelon canonical form (Hannan and Deistler 1988) or balanced canonical forms (Ober 1996). The orthonormality property of the matrix C_1 in the canonical form representation also simplifies the computation of the cointegrating space.

5.2 Example 2: I(2) processes

Let us next consider a trivariate system corresponding to an I(2) process. Look at a system where A has the eigenvalue 1 with algebraic multiplicity 3 and geometric multiplicity 2:

$$A = \operatorname{diag}(A_1, A_{st}) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & a_{st} \end{bmatrix}, B = \begin{bmatrix} B_1 \\ B_{st} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ b_{st1} & b_{st2} & b_{st3} \end{bmatrix},$$

$$C = [C_1, C_{st}] = \begin{bmatrix} 1 & 0 & 1 & c_{1st} \\ 0 & 1 & 1 & c_{2st} \\ 0 & 0 & 1 & c_{3st} \end{bmatrix}.$$

Here (A_{st}, B_{st}, C_{st}) is such that $\Pi(A_{st}, B_{st}, C_{st}) = k_{st}(z)$ is a stable minimum-phase transfer function. In the following we want to find a unique minimal state space representation of the transfer function $k(z) = \Pi(A, B, C)$. From Example 1 we already know that for this example the first component of the state is I(2) and its first difference is up to stationary terms equal to the second component of the state.

In constructing a canonical representation, again the system matrix A is the starting point. Also in this example the A-matrix is already in Jordan normal form, J say again. Thus, analogously to the above example the first question to analyze is which transformation matrices obey $TJT^{-1} = J$. In the case of a non-diagonal Jordan normal form, not every nonsingular matrix T fulfills this equation. Look again only at the nonstationary components, then the invariance of the Jordan normal form can also be written as:

$$J_1 = T_1 J_1 T_1^{-1} = T_1 I_3 T_1^{-1} + T_1 N_1 T_1^{-1} = I_3 + N_1$$

with

$$N_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus, the set of transformation matrices $\mathcal{M} = \{T_1 : T_1 J_1 T_1^{-1} = J_1\}$ is restricted to fulfill the equation $T_1 N_1 T_1^{-1} = N_1$. This holds if and only if T_1 is of the form:

$$T_1 = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{11} & 0 \\ 0 & t_{32} & t_{33} \end{bmatrix}$$

with t_{ij} such that T_1 is non-singular. Thus, we see that in the case of higher integration orders the requirement that the part of the A-matrix corresponding to the nonstationary components is in Jordan normal form leads to restrictions on the set of feasible transformation matrices $\mathcal{M}$. The restrictions depend upon the off-diagonal elements in the Jordan normal form. In general $\mathcal{M}$ is not reduced to a singleton due to the occurrence of Jordan blocks of size larger than one, therefore again further restrictions have to be imposed. The required restrictions are imposed on the columns of C_1 and the rows of B_1. Denote the first three columns of C corresponding to the nonstationary components of the state by $C_1 = \begin{bmatrix} C_1^{1,E} & C_1^{2,G} & C_1^{2,E} \end{bmatrix}$, in the example given by:

$$C_1^{1,E} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad C_1^{2,G} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad C_1^{2,E} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

The superscripts are chosen as follows, already in view of the general case discussed in the following section: Superscript $1, E$ indicates that this column corresponds to a state

component of highest integration order and a column of A corresponding to an eigenvector; superscript $2, G$ indicates that these columns correspond to state components of second highest integration order. The second superscript G indicates the correspondence to a generalized eigenvector, as this column corresponds to the generalized eigenvector of the Jordan block of size 2; the second superscript E indicates the correspondence to an eigenvector of A. For the column(s) corresponding to the highest integration order therefore only a second superscript E occurs.

Now we will describe, in a constructive fashion, how to obtain the canonical form representation of the system, defined in Theorem 2 below. To describe our canonical form it is required to find restrictions on the sub-blocks composed of columns of C_1 that are fulfilled for exactly one transformed matrix $\hat{C}_1 = C_1 T$, with T such that $TJT^{-1} = J$. For the given example we have:

$$\hat{C}_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{11} & 0 \\ 0 & t_{32} & t_{33} \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12}+t_{32} & t_{13}+t_{33} \\ 0 & t_{11}+t_{32} & t_{33} \\ 0 & t_{32} & t_{33} \end{bmatrix}.$$

Start a recursive argument with $\hat{C}_1^{1,E}$: If $\hat{C}_1^{1,E}$ is required to be an orthonormal matrix, this implies $t_{11} = \pm 1$. This restriction is already fulfilled for $C_1^{1,E}$. Next, also for the column corresponding to the other eigenvector (the eigenvector corresponding to the Jordan block of size one), i.e. for $\hat{C}_1^{2,E}$ we require $(\hat{C}_1^{2,E})'\hat{C}_1^{2,E} = I$ and additionally its orthogonality to $\hat{C}_1^{1,E}$. This directly implies that $t_{13} = -t_{33}$ and $t_{33} = \pm 1/\sqrt{2}$. This leaves only t_{12} and t_{32} as free entries and the signs of t_{11} and t_{33} undetermined, the remaining entries of T already having uniquely defined values. As a final restriction on $\hat{C}_1$, the column corresponding to the generalized eigenvector is required to be orthogonal to the eigenvector columns, i.e. $(\hat{C}_1^{2,G})'[\hat{C}_1^{1,E}, \hat{C}_1^{2,E}] = 0$, introducing the restrictions $t_{12} + t_{32} = 0, t_{11} = -2t_{32}$. This leads to

$$T = \begin{bmatrix} t_{11} & t_{11}/2 & -t_{33} \\ 0 & t_{11} & 0 \\ 0 & -t_{11}/2 & t_{33} \end{bmatrix}, T^{-1} = \begin{bmatrix} t_{11} & 0 & t_{11} \\ 0 & t_{11} & 0 \\ 0 & 1/(2t_{33}) & 1/t_{33} \end{bmatrix}, t_{11} = \pm 1, t_{33} = \pm 1/\sqrt{2}.$$

Noting that the sign of t_{11} determines the sign of the first nonzero entry in the second row of $\hat{B}_1 = T^{-1}B_1$ and that the same is true for t_{33} and the third row of $\hat{B}_1$ the restriction of

these rows to start with positive entries determines the matrix T uniquely as

$$T_0 = \begin{bmatrix} 1 & 1/2 & -1/\sqrt{2} \\ 0 & 1 & 0 \\ 0 & -1/2 & 1/\sqrt{2} \end{bmatrix}$$

with which the given realization (A, B, C) has to be transformed in order to arrive at the (therefore unique) canonical form representation given by:

$$\hat{C} = \begin{bmatrix} 1 & 0 & 0 & c_{1st} \\ 0 & 1/2 & 1/\sqrt{2} & c_{2st} \\ 0 & -1/2 & 1/\sqrt{2} & c_{3st} \end{bmatrix}, \hat{B} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ \frac{3}{2}\sqrt{2} & \sqrt{2} & 1/\sqrt{2} \\ b_{st1} & b_{st2} & b_{st3} \end{bmatrix}.$$

The system $(A, \hat{B}, \hat{C})$ is the only minimal realization of the class of observationally equivalent state space systems corresponding to the transfer function $k(z) = I_s + zC(I_n - zA)^{-1}B$ that fulfills the required restrictions. Note again that for the stationary part any canonical form can be employed, e.g. the echelon canonical form. For this reason the stationary part can essentially be neglected in the discussion.

The canonical representation just developed offers the advantage of revealing the cointegrating spaces immediately: As in Example 1, the cointegrating relationships can be found via orthogonality conditions to certain columns of $\hat{C}$ (again assuming $x_{1,st} = \sum_{j=0}^{\infty} A_{st}^j B_{st} \varepsilon_{-j}$):

$$y_t = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} x_{t,1} + \begin{pmatrix} 0 \\ 1/2 \\ -1/2 \end{pmatrix} x_{t,2} + \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} x_{t,3} + k_{st}(z)\varepsilon_t. \tag{9}$$

Note that the components $x_{t,2}$ and $x_{t,3}$ corresponding to different Jordan blocks are both I(1) but not cointegrated, as follows from the representation

$$\begin{bmatrix} x_{t,2} \\ x_{t,3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ \frac{3}{2}\sqrt{2} & \sqrt{2} & 1/\sqrt{2} \end{bmatrix} \sum_{j=1}^{t-1} \varepsilon_{t-j} + \begin{bmatrix} x_{1,2} \\ x_{1,3} \end{bmatrix}$$

and the fact that ε_t has a nonsingular covariance matrix. Taking this into account, the cointegrating relationships are readily found. Any vector of the form $\beta = [0, b_2, b_3]$ reduces the integration order in $\beta' y_t$ to at most integration order 1. It reduces the integration order immediately to 0, if and only if it is also orthogonal to $\hat{C}_1^{2,E}$ and $\hat{C}_1^{2,G}$, i.e. to the second and third column of $\hat{C}$. However, in our example the matrix built of the first three columns of $\hat{C}$ has full rank, and thus no cointegrating relationship that reduces the integration order

from 2 to 0 exists.

In the present I(2) system also polynomial cointegration, see Granger and Lee (1989) or Gregoir and Laroque (1994), occurs. We abstain here from a discussion of polynomial cointegration and refer the interested reader to Bauer and Wagner (2003a) for a detailed discussion of polynomial cointegration in state space systems. In that paper we show that the developed canonical representation is very suitable for the analysis of polynomially cointegrated systems. The canonical system representation highlights the relationships between the state components that give rise to polynomial cointegration and allows to recover all polynomial cointegrating relationships via orthogonality constraints.

5.3 Example 3: Multiple Frequency I(1) Processes

As a third example consider a rational *multiple frequency* I(1), or short MFI(1), process. By this we denote (cf. Definition 3) a process that is integrated at finitely many frequencies with the corresponding integration orders all equal to 1. The primary example of such a process is the case of seasonal integration, where e.g. for quarterly observations unit roots may be present at ± 1 and $\pm i$. From the previous discussion we already know that for the integration order to be 1, the Jordan segments corresponding to the different unit roots must be composed only of Jordan blocks of size one. Let $z_1, \ldots, z_l$ denote the l distinct unit roots, ordered according to increasing frequency $\omega_k \in [0, 2\pi)$, i.e. $\omega_k < \omega_{k+1}, k = 1, \ldots, l-1$ for $z_k = e^{i\omega_k}$. Then the system can be written, with the A-matrix in Jordan normal form, as:

$$\begin{bmatrix} x_{t+1,1} \\ \vdots \\ x_{t+1,l} \\ x_{t+1,st} \end{bmatrix} = \begin{bmatrix} z_1 I_{c_1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & z_l I_{c_l} & 0 \\ 0 & \ldots & 0 & J_{st} \end{bmatrix} \begin{bmatrix} x_{t,1} \\ \vdots \\ x_{t,l} \\ x_{t,st} \end{bmatrix} + \begin{bmatrix} B_1 \\ \vdots \\ B_l \\ B_{st} \end{bmatrix} \varepsilon_t$$

$$y_t = \begin{bmatrix} C_1 & \cdots & C_l & C_{st} \end{bmatrix} \begin{bmatrix} x_{t,1} \\ \vdots \\ x_{t,l} \\ x_{t,st} \end{bmatrix} + \varepsilon_t.$$

where $x_{t,k}, k = 1, \ldots, l$ denotes the state components corresponding to the unit root z_k, $x_{t,st}$ denotes the stationary (for suitable initial values) components of the state and B and C are partitioned accordingly. Starting with an initial state x_1 the above system has a

solution y_t as:

$$y_t = C_1 B_1 \sum_{j=1}^{t-1} z_1^{j-1}\varepsilon_{t-j} + \cdots + C_l B_l \sum_{j=1}^{t-1} z_l^{j-1}\varepsilon_{t-j} + C_1 B_1 z_1^{t-1} x_{1,1} + \cdots + C_l B_l z_l^{t-1} x_{1,l} + k_{st}(z)\varepsilon_t. \tag{10}$$

Here $x_{1,k}, k = 1, \ldots, l$ denotes the components of the initial state corresponding to the different segments in the Jordan normal form and $x_{1,st}$ is assumed to be chosen such that $k_{st}(z)\varepsilon_t = \varepsilon_t + \sum_{j=0}^{t-2} C_{st} J_{st}^j B_{st} \varepsilon_{t-j-1} + C_{st} J_{st}^{t-1} x_{1,st}$ is stationary. The representation (10) shows an advantage of the state space representation: It directly leads to a Granger type representation that reveals the contributions of the nonstationary components corresponding to the different unit roots to the output. Starting from autoregressive or ARMA representations requires more complicated computations to arrive at this type of representation (cf. e.g. Johansen and Schaumburg 1999).

Due to the real valuedness of $(y_t)_{t\in\mathbb{N}}$ a number of restrictions are imposed on the system matrices. For each unit root $z_k = e^{i\omega_k}, \omega_k \notin \{0, \pi\}$ there exists an index k' such that $z_{k'} = \bar{z}_k$, i.e. complex unit roots occur in pairs of conjugate complex roots. For these pairs also $c_{k'} = c_k$, $C_{k'} = \bar{C}_k$ and $B_{k'} = \bar{B}_k$ hold. This follows immediately from the fact that CA^jB has to be real valued for all $j \geq 0$ for real valued y_t. Taking these restrictions into account, the contribution to the output stemming from a pair of conjugate complex unit roots is given by

$$C_k B_k \sum_{j=1}^{t-1} z_k^{j-1}\varepsilon_{t-j} + \bar{C}_k \bar{B}_k \sum_{j=1}^{t-1} \bar{z}_k^{j-1}\varepsilon_{t-j} + C_k B_k z_k^{t-1} x_{1,k} + \bar{C}_k \bar{B}_k \bar{z}_k^{t-1} x_{1,k'}, \tag{11}$$

which is directly seen to be real valued if also the initial states to conjugate complex pairs of roots are conjugate complex. Thus, the contribution to the output stemming from a pair of complex conjugate unit roots is seen to be the sum of conjugate complex stochastic cycles plus a contribution from the initial state. Similar restrictions concerning the stable complex roots of course apply also to the stationary part. The representation (10), (11) also shows directly that the real valued MFI(1) process $(y_t)_{t\in\mathbb{N}}$ has unit root structure $\{(\omega_1, 1), \ldots, (\omega_l^{\mathbb{R}}, 1)\}$, where $\omega_1 < \ldots < \omega_{l^{\mathbb{R}}}$ denote the $l^{\mathbb{R}}$, compare Definition 1, frequencies in the range $[0, \pi]$, as $(1 - z_k z)\left[\sum_{j=1}^{t-1} z_k^{j-1}\varepsilon_{t-j}\right] = \varepsilon_{t-1}, t > 2$ for $k = 1, \ldots, l$. In a minimal state space representation however more information is present, as also the number of state components integrated at a specific frequency and therefore the dimension of the

cointegrating space at this frequency is directly seen in the A-matrix in Jordan normal form. This fact leads to the definition of the state space unit root structure in Definition 6 below.

As for the I(1) example the cointegrating spaces are directly seen from the Granger type representation (10). Since we are concerned only with real valued cointegrating spaces, for complex unit roots and integration of order 1 we have to distinguish between static and *dynamic* cointegrating relationships (see e.g. Johansen and Schaumburg 1999) to recover the full set of possible cointegrating relationships.

Note first at this point if we consider complex cointegration (cf. Remark 5), i.e. if we allow for $\beta \in \mathbb{C}^s$ then the orthogonality constraint $\beta' C_k = 0$, where $C_k \in \mathbb{C}^{s \times c_k}$ say, leads to a $s - c_k$ dimensional complex cointegrating space corresponding to unit root z_k. Thus, the link, discussed above in the I(1) example, between the number of common cycles and the dimension of the cointegrating space prevails also for the case of complex unit roots in the MFI(1) case. Considering only real valued cointegrating vectors breaks this link.

For a real valued output process it immediately follows that the (*real valued*) cointegrating spaces corresponding to complex conjugate unit roots coincide, since for $\beta' C_k = 0$ it follows that $\overline{\beta' C_k} = \beta' \bar{C}_k = 0$. Therefore a cointegrating vector at the unit root z_k is also a cointegrating vector at the unit root $\bar{z}_k$. The orthogonality constraint $\beta' C_k = 0$ (solved over $\mathbb{R}$) can be rewritten in real form as $\beta'[\mathcal{R}(C_k), \mathcal{I}(C_k)] = 0$, with $\mathcal{R}$ denoting the real and $\mathcal{I}$ denoting the imaginary part of a complex number. Full column rank (in $\mathbb{C}^s$) of C_k does not imply full column rank (in $\mathbb{R}^s$) of $[\mathcal{R}(C_k), \mathcal{I}(C_k)]$. The latter matrix can take on any rank $c_k, c_k + 1, \ldots, \min(2c_k, s)$. Thus, in a real valued discussion there is no link between the number of common cycles and the dimension of the *static* cointegrating space corresponding to a complex unit root of order 1.

Note at this point as a remark that the above orthogonality constraint $\beta'[\mathcal{R}(C_k), \mathcal{I}(C_k)] = 0$ is exactly the same orthogonality constraint as the one that arises from the real valued system representation, where the sub-blocks (J_k, B_k, C_k) and $(\bar{J}_k, \bar{B}_k, \bar{C}_k)$ are transformed to a real valued sub-system comprising both blocks, see (15). In the corresponding real valued system representation, the real C-matrix corresponding to the pair of complex conjugate unit roots, $C_{k,\mathbb{R}}$ say, is given by $[\mathcal{R}(C_k), \mathcal{I}(C_k)]$.

The focus on real valued cointegration gives rise to *dynamic* cointegrating relationships. In

the MFI(1) cases these are polynomial cointegrating relationships of degree 1, i.e. of the form $\beta(z) = \beta_0 + \beta_1 z$ with $\beta_0, \beta_1 \in \mathbb{R}^s$. To see the argument, the state space representation is again very revealing. Look only at one term of representation (10) to obtain for $z_t = \sum_{j=1}^{t-1} z_k^{j-1} \varepsilon_{t-j}$ using $z z_t = z_{t-1} = \sum_{j=1}^{t-2} z_k^{j-1} \varepsilon_{t-j-1}$ that for $t > 2$

$$(\beta_0' + \beta_1' z) C_k B_k z_t = \beta_0' C_k B_k \varepsilon_{t-1} + \left[\beta_0' C_k z_k + \beta_1' C_k\right] B_k \sum_{j=1}^{t-2} z_k^{j-1} \varepsilon_{t-j-1}.$$

Thus dynamic cointegration at the unit root z_k occurs for

$$\begin{bmatrix} \beta_0' & \beta_1' \end{bmatrix} \begin{bmatrix} C_k z_k \\ C_k \end{bmatrix} = 0. \tag{12}$$

The dynamic cointegrating relationships are found via orthogonality relationships over a real space of dimension $2s$ by equating the real and the imaginary part of (12) separately. Note that equivalently again the real valued system representation can be taken to recover the dynamic cointegrating spaces as

$$\begin{bmatrix} \beta_0' & \beta_1' \end{bmatrix} \begin{bmatrix} C_{k,\mathbb{R}} J_{k,\mathbb{R}} \\ C_{k,\mathbb{R}} \end{bmatrix} = \beta_0' \begin{bmatrix} \mathcal{R}(C_k) & \mathcal{I}(C_k) \end{bmatrix} \begin{bmatrix} \mathcal{R}(z_k) I_{c_k} & \mathcal{I}(z_k) I_{c_k} \\ -\mathcal{I}(z_k) I_{c_k} & \mathcal{R}(z_k) I_{c_k} \end{bmatrix} + \beta_1' \begin{bmatrix} \mathcal{R}(C_k) & \mathcal{I}(C_k) \end{bmatrix} = 0. \tag{13}$$

The matrix, respectively the matrix product, on the right hand sides of equations (12) and (13) can be shown to have full column rank. Thus, the dynamic cointegrating spaces are seen to be of dimension $2(s - c_k)$ for the complex unit roots in MFI(1) processes. This reestablishes the analogy to the well known relation for the unit roots ± 1. Note that in the above space the static cointegrating relationships are contained as a subset with $\beta_1 = 0$. It is also obvious that if β is a static cointegrating relationship, then $\beta(z) = 0 + \beta z$ is a dynamic cointegrating relationship. However, one that does not add additional insights.

The above discussion also makes clear that in the case of complex unit roots and a focus on real cointegration it suffices to investigate the system blocks in the complex representation corresponding to the unit roots with frequencies in the interval $[0, \pi]$. This, as has been illustrated in the example and holds true in general, is equivalent to consider the real valued blocks corresponding to pairs of complex conjugate unit roots.

Given the above discussion, a *complex valued* canonical representation of MFI(1) processes is easily obtained by replicating the result from the I(1) case for each of the Jordan segments corresponding to the different unit roots. Thus, we require the matrix A to be in

Jordan normal form, the blocks C_k corresponding to the Jordan segments $z_k I_{c_k}$ to fulfil $C_k' C_k = I_{c_k}$ and the blocks B_k to be p.u.t. The block structure of the Jordan normal form allows to analyze each block separately. This follows from the fact that the set of all transformation matrices T that solve $TJT^{-1} = J$ consists only of block-diagonal matrices $T = \text{diag}(T_1, \ldots, T_l, T_{st})$, see the proof of Theorem 2 in the Appendix. A real valued canonical form also follows (cf. Theorem 2).

In Bauer and Wagner (2002b) pseudo maximum likelihood estimation is considered for MFI(1) processes, based on the just discussed canonical state space representation. Both consistency and the asymptotic distribution of the estimates are derived.

6 The Canonical Form

In Lemma 1 in Section 3 it has been shown that for every rational unit root process there exists a minimal state space representation. It has also become clear that these representations are not unique. This stems from the fact that even under the restriction to minimal realizations (A, B, C) of order n, the set of observationally equivalent state space systems can be described by $\mathcal{M}$, the set of all $n \times n$ nonsingular matrices. In the preceding examples section we have already seen possible ways to restrict the set of observationally equivalent state space realizations obeying certain restrictions to singletons. In this section the ideas introduced already in the examples are applied to construct for any given transfer function $k(z) \in M_n$ a unique state space realization, say $(\hat{A}, \hat{B}, \hat{C})$. The construction of this canonical realization is recursive and proceeds by sequentially imposing restrictions until exactly one system in the equivalence class fulfills all restrictions. Let $\mathcal{M}_i \subset \mathcal{M}$ denote the set of all matrices $T \in \mathcal{M}$ for which starting from a realization (A, B, C) fulfilling all restrictions imposed up to the i-th step, also every realization (TAT^{-1}, TB, CT^{-1}) for $T \in \mathcal{M}_i$ fulfills all restrictions. Thus, $\mathcal{M}_i$ represents the class of feasible transformation matrices after step i. The stepwise imposition of further restrictions implies $\mathcal{M}_{i+1} \subset \mathcal{M}_i$. The aim is to reduce $\mathcal{M}_i$ to a singleton, i.e. to arrive at $\mathcal{M}_\iota = \{I_n\}$ for some ι.

All arguments required for the construction of the canonical form have already been presented in the examples. The A-matrix in Jordan normal form clearly displays the contributions due to the different unit roots. It has been indicated in Section 3 that the canonical form is based on a *reordered* version of the Jordan normal form. Before presenting the

canonical form in Theorem 2, the following Theorem 1 shows why the components of the state are arranged in the way discussed next. Let us start from the A-matrix of a minimal state space representation (A, B, C) given in its Jordan normal form, where we assume that the Jordan segments corresponding to the unit roots $z_1, \ldots, z_l$ with $z_k = e^{i\omega_k}, \omega_k \in [0, 2\pi)$ are ordered according to increasing frequency, where again each unit root is taken into account separately. The matrix A is thus given by:

$$A = \begin{bmatrix} J_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & J_l & 0 \\ 0 & \ldots & 0 & J_{st} \end{bmatrix}.$$

The matrices $J_1, \ldots, J_l$ are the Jordan segments corresponding to the unit roots $z_1, \ldots, z_l$. The matrix J_{st} accounts for all eigenvalues of A smaller than one in absolute value. In the A-matrix all information concerning the unit roots is contained. Each unit root corresponds to an eigenvalue of A and therefore to a diagonal block J_k of J. It is now possible to reorder the Jordan blocks within the Jordan segments – this amounts to a transformation of (A, B, C) with some matrix T – to arrive at the reordered Jordan normal form for each of the Jordan segments J_k given by:

$$J_k = \begin{bmatrix} z_k I_{d_1} & [I_{d_1}, 0^{d_1 \times (d_2 - d_1)}] & 0 & 0 & 0 \\ 0^{d_2 \times d_1} & z_k I_{d_2} & [I_{d_2}, 0^{d_2 \times (d_3 - d_2)}] & 0 & \vdots \\ 0 & 0 & z_k I_{d_3} & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & [I_{d_{m_k-1}}, 0^{d_{m_k-1} \times (d_{m_k} - d_{m_k-1})}] \\ 0 & 0 & 0 & 0 & z_k I_{d_{m_k}} \end{bmatrix} \quad (14)$$

Some notation has to be explained: The integer m_k denotes the size of the largest Jordan block corresponding to the unit root z_k. The integers $1 \leq d_1 \leq d_2 \leq \ldots \leq d_{m_k}$ denote the differences of the dimension of the image of $(J_k - z_k I)^{m_k - j}$ and the dimension of the image of $(J_k - z_k I)^{m_k - j + 1}$ for $j = 1, \ldots, m_k$. Note that in order to keep notation as simple as possible the dependence of the integers $d_j = d_j^k$ on k is suppressed when only one typical block is investigated. For later use define $N_k = J_k - z_k I$ and also $d^k = \sum_{j=1}^{m_k} d_j^k$. Notwithstanding the already large amount of notation, some more notation is required to proceed: Let $B_k = [(B_k^1)', \ldots, (B_k^{m_k})']', B_k^j \in \mathbb{C}^{d_j \times s}, C_k = [C_k^1, \ldots, C_k^{m_k}], C_k^j \in \mathbb{C}^{s \times d_j}, x_{t,k} =$

$[(x_{t,k}^1)', \ldots, (x_{t,k}^{m_k})']', x_{t,k}^j \in \mathbb{C}^{d_j}$ be partitioned according to (14). After collecting the notation the following theorem now clarifies the properties of the state for minimal systems (A, B, C) and also makes clear why the reordering as discussed is useful:

Theorem 1 *Let (J, B, C) denote a system, where the matrix J is block-diagonal with diagonal blocks of the form (14). Let the corresponding blocks C_k, B_k and $x_{t,k}$ be partitioned accordingly. Further partition $C_k^j = \begin{bmatrix} C_k^{j,G} & C_k^{j,E} \end{bmatrix}$, with $C_k^{j,E} \in \mathbb{C}^{s \times (d_j - d_{j-1})}$ and $C_k^{j,G} \in \mathbb{C}^{s \times d_{j-1}}, j = 1, \ldots, m_k$, where $d_0 = 0$ is used. Define furthermore $\bar{C}_k^E = [C_k^{1,E}, \ldots, C_k^{m_k,E}] \in \mathbb{C}^{s \times d_{m_k}}$.*
Then the system (J, B, C) is minimal if and only if $B_k^{m_k}$ has full row rank, $\bar{C}_k^E$ has full column rank and the representation (A_{st}, B_{st}, C_{st}) of the stationary subsystem is minimal. In this case $x_{t,k}^j$ is integrated with complex unit root structure $((\omega_k, m_k - j + 1))$ where $z_k = e^{i\omega_k}$.
Full rank of $\bar{C}_k^E$ implies $d_{m_k} \leq s$ and full rank of $C_k^{j,E}, j = 1, \ldots, m_k$.

Proof: It has been noted that minimality of a state space system is equivalent to full rank of the corresponding matrices $\mathcal{O}$ and $\mathcal{C}$. From the block-diagonal structure of J, it follows that the k-th block row of $\mathcal{C}$ can be written as $\mathcal{C}_k = [B_k, J_k B_k, J_k^2 B_k, \ldots]$. Full row rank of this matrix is equivalent to full row rank of $[B_k, N_k B_k, N_k^2 B_k, \ldots]$. Assume, that $B_k^{m_k}$ is of full row rank for $k = 1, \ldots, l$. The full rank of $\mathcal{C}$ in this case is shown via contradiction. Hence let the vector $u \in \mathbb{C}^{d^k}$ be such that $u' N_k^j B_k = 0, j = 0, 1, 2, \ldots$. Decompose $u' = [u_1', \ldots, u_{m_k}'], u_j \in \mathbb{C}^{d_j}$. Then from $u' N_k^{m_k - 1} B_k = 0$ and the full rank assumption on $B_k^{m_k}$ it follows that $u_1 = 0$. This can be seen from a tedious but straightforward computation of $N_k^{m_k - 1}$. Using $u_1 = 0$, the equation $u' N_k^{m_k - 2} B_k = 0$ implies $u_2 = 0$. Iterating this argument leads to $u_j = 0, j = 1, \ldots, m_k - 1$. Finally the equality $u' B_k = 0$ results in $u_{m_k}' B_k^{m_k} = 0$ implying $u_{m_k} = 0$ using again the full rank of $B_k^{m_k}$. This shows, that full row rank of $B_k^{m_k}$ implies full rank of $\mathcal{C}$. Conversely this argument can be used to show rank reduction of $\mathcal{C}$ based on reduced rank of $B_k^{m_k}$.
From full rank of $B_k^{m_k}$ it follows for minimal (J, B, C) that $x_{t+1,k}^{m_k} = z_k x_{t,k}^{m_k} + B_k^{m_k} \varepsilon_t$ has complex unit root structure $((\omega_k, 1))$ for $z_k = e^{i\omega_k}$ and contains m_k linearly independent complex stochastic cycles. Proceeding now to $x_{t+1,k}^{m_k - 1} = z_k x_{t,k}^{m_k - 1} + \left[I_{d_{m_k - 1}}, 0^{d_{m_k - 1} \times (d_{m_k} - d_{m_k - 1})} \right] x_{t,k}^{m_k} + B_k^{m_k - 1} \varepsilon_t$ shows complex unit root structure $((\omega_k, 2))$ for these components of the state $x_{t,k}$. Continuing recursively one obtains that the first d_1 components of $x_{t,k}$ have complex unit

root structure $((\omega_k, m_k))$.
Minimality of the realization also requires full column rank of $\mathcal{O} = [C', J'C', (J^2)'C', \ldots]'$. Analogously to the proof above this is equivalent to full column rank of $\bar{C}_k^E$. $\square$
Restricting the part of the A-matrix containing the eigenvalues on the unit circle to be in reordered Jordan normal form and fixing a canonical form for the stationary part of the transfer function restricts the set of feasible transformation matrices T to $\mathcal{M}_1 = \{T = \text{diag}(T_1, \ldots, T_l, I) : T_k J_k T_k^{-1} = J_k, k = 1, \ldots, l\}$ as is shown in the proof of Theorem 2 in the appendix. The set $\mathcal{M}_1$ is generally not a singleton containing only the identity matrix. This necessitates the imposition of further restrictions, which we impose on sub-blocks of both C_k and B_k. Again, the idea is to formulate restrictions in such a way that the cointegration properties of y_t are highlighted. The result is summarized in the following Theorem 2, whose proof is as mentioned above contained in the appendix.

Theorem 2 *For each rational process $(y_t)_{t\in\mathbb{N}}$ with complex unit root structure $((\omega_1, h_1), \ldots, (\omega_l, h_l))$ there exists a unique state space representation obeying the following restrictions:*

- *The nonstationary part of the A-matrix is block-diagonal, where the diagonal blocks are of reordered Jordan normal form (14) and the unit roots are ordered according to increasing frequency $\omega_k \in [0, 2\pi)$. For all k the equality $m_k = h_k$ holds.*

- *For each of the matrices $C_k, k = 1, \ldots, l$ corresponding to unit root z_k the following restrictions hold, using the notation of Theorem 1: $(\bar{C}_k^E)'\bar{C}_k^E = I$ and $(C_k^{j,G})'C_k^{i,E} = 0, i \leq j$ for $j = 1, \ldots, m_k$ and $k = 1, \ldots, l$.*

- *Using the partitioning $B_k^{m_k} = [(B_k^{m_k,1})', (B_k^{m_k,2})', \ldots, (B_k^{m_k,m_k})']', B_k^{m_k,j} \in \mathbb{C}^{(d_j^k - d_{j-1}^k)\times s}$, each sub-block $B_k^{m_k,j}$ is p.u.t. for $j = 1, \ldots, m_k$ and $k = 1, \ldots, l$.*

- *The stationary part of the transfer function, (A_{st}, B_{st}, C_{st}) is in a canonical form for stationary state space models, e.g. in echelon canonical form.*

For each real valued rational process $(y_t)_{t\in\mathbb{N}}$ this representation additionally has the following properties: For each $z_k \notin \{1, -1\}$ there exists an index k' such that $\bar{z}_k = z_{k'}$ and $\bar{J}_k = J_{k'}, \bar{C}_k = C_{k'}, \bar{B}_k = B_{k'}$. The stationary system matrices (A_{st}, B_{st}, C_{st}) are real valued.

Transforming the two sub-blocks corresponding to k and k' that correspond to complex conjugate unit roots separately for all indices $k = 1, \ldots, l^{\mathbb{R}}$ such that $z_k \notin \{1, -1\}$, the following transformation leads to real valued matrices $(A_{\mathbb{R}}, B_{\mathbb{R}}, C_{\mathbb{R}})$ composed of blocks:

$$\begin{aligned} J_{k,\mathbb{R}} = \begin{bmatrix} I_{d^k} & I_{d^k} \\ iI_{d^k} & -iI_{d^k} \end{bmatrix} \begin{bmatrix} J_k & 0 \\ 0 & \bar{J}_k \end{bmatrix} \begin{bmatrix} I_{d^k} & I_{d^k} \\ iI_{d^k} & -iI_{d^k} \end{bmatrix}^{-1} &= \begin{bmatrix} \mathcal{R}(J_k) & \mathcal{I}(J_k) \\ -\mathcal{I}(J_k) & \mathcal{R}(J_k) \end{bmatrix}, \\ B_{k,\mathbb{R}} = \begin{bmatrix} I_{d^k} & I_{d^k} \\ iI_{d^k} & -iI_{d^k} \end{bmatrix} \begin{bmatrix} B_k \\ \bar{B}_k \end{bmatrix} &= \begin{bmatrix} 2\mathcal{R}(B_k) \\ -2\mathcal{I}(B_k) \end{bmatrix}, \\ C_{k,\mathbb{R}} = \begin{bmatrix} C_k & \bar{C}_k \end{bmatrix} \begin{bmatrix} I_{d^k} & I_{d^k} \\ iI_{d^k} & -iI_{d^k} \end{bmatrix}^{-1} &= \begin{bmatrix} \mathcal{R}(C_k) & \mathcal{I}(C_k) \end{bmatrix}. \end{aligned}$$

Here $\mathcal{R}$ denotes the real part and $\mathcal{I}$ the imaginary part of a complex quantity.
The mapping ϕ attaching this unique real valued state space realization (A, B, C) to the transfer function $k(z) \in M_n$ is a canonical form.

The proof itself consists of a verification of the fact that the formulated restrictions are indeed sufficient to select a unique minimal state space realization. The proof is constructive and recursive and its main ideas have already been displayed in Section 5, most notably in Example 2.
In the chosen canonical representation the cointegrating relationships are clearly revealed. Using the partitioning and notation introduced above, it follows that every rational unit root process has a unique representation as

$$y_t = Cx_t + d_t + \varepsilon_t = \sum_{k=1}^{l} \sum_{j=1}^{m_k} C_k^j x_{t,k}^j + C_{st} x_{t,st} + d_t + \varepsilon_t$$

with $C_k^j x_{t,k}^j$ of complex unit root structure $((\omega_k, m_k - j + 1))$ for $C_k^j \neq 0$ according to Theorem 1. Thus, in $\beta' y_t$ with $\beta \in \mathbb{R}^s$ such that $\beta'[C_k^1, C_k^2, \ldots, C_k^j] = 0$ and $\beta' C_k^{j+1} \neq 0$, the order of integration corresponding to the unit root z_k is reduced to $m_k - j$, whereas y_t itself is integrated of order m_k at z_k. Note again that in case that z_k is a member of a pair of complex conjugate unit roots, the vector β from above also reduces the integration order of y_t at $\bar{z}_k$ to $m_k - j$, compare also the discussion of Example 3. This property of revealing the cointegrating relationships is the main argument for the ordering of the components of the state corresponding to different integration orders. Also polynomial cointegrating relationships can be recovered via orthogonality relationships when the system is represented in the developed canonical form, see Bauer and Wagner (2003a). Thus, the canonical

state space representation reveals more information concerning the integration properties of $(y_t)_{t\in\mathbb{N}}$ than is summarized in the unit root structure defined in Section 2. This leads us to define a state space unit root structure.

Definition 6 *The s-dimensional real random process $(y_t)_{t\in\mathbb{N}}$ with minimal state space representation (6) has, using the notation of the above discussion, state space unit root structure $\Omega = \{(\omega_1, (d_1^1, \ldots, d_{m_1}^1)), \ldots, (\omega_l, (d_1^l, \ldots, d_{m_l}^l))\}, 0 \leq d_1^k \leq d_2^k \leq \ldots \leq d_{m_k}^k \leq s$ for all $k = 1, \ldots, l$, if the matrix A in the unique representation given in Theorem 2 is of the form $A = diag(J_1, \ldots, J_l, J_{st})$, where $|\lambda_{max}(J_{st})| < 1$ and the matrices J_k are of the form (14).*

7 Summary and Conclusions

In this paper we have shown that the state space framework is very useful for the analysis of (co)integrated processes. We have developed a specific canonical representation of unit root processes where we allow for unit roots with integer integration orders at any point on the unit circle. The developed state space representation clearly reveals the integration and also the cointegration properties, also or especially for processes with unit roots spread over the unit circle and with higher integration orders. The information is summarized in the defined state space unit root structure. The equivalence to ARMA models also overcomes the common limitation of the literature to AR models.

The developed representation and the discussed equivalence to ARMA models leads us to conclude that the state space framework offers some potential advantages for understanding the dynamic properties of unit root processes with a complicated unit root structure. To illustrate this, several examples covering the main cases found to be relevant for empirical analysis are discussed in detail. In one of the examples it is seen that the developed canonical state space representation directly leads to Granger type representations for rational MFI(1) processes. Furthermore it is also illustrated in the examples that the static, dynamic and polynomial cointegrating spaces corresponding to the different unit roots can be recovered via orthogonality relationships. For a detailed discussion of polynomial cointegration based on the representation developed in this paper see Bauer and Wagner (2003a).

Two related papers contain first applications of the results developed in this paper. In

Bauer and Wagner (2002b) pseudo maximum likelihood estimates and their asymptotic distributions are derived for MFI(1) processes. For the standard I(1) case not only pseudo maximum likelihood estimates are available, but also so called subspace algorithms are applied for estimation, see Bauer and Wagner (2002a). In that paper based on the subspace algorithm estimates also tests for the cointegrating rank are presented.

A couple of important problems remain unanswered and are left for further research. Firstly the inclusion of exogenous or deterministic variables in the canonical form representation is not discussed in this paper. In the present paper all deterministic variables are summarized in d_t and excluded from further analysis. It has been shown above that there is an inherent non-identifiability of the deterministic components and the initial state. The pseudo maximum likelihood analysis in the MFI(1) framework shows that the asymptotic properties of the estimates differ for deterministic components that can be attributed to an initial state and the remaining ones (cf. Bauer and Wagner 2002b). This needs to be further investigated. Secondly, the derivation of pseudo maximum likelihood estimates also for higher order integrated systems is straightforward using a parameterization based on the canonical form. The properties of these estimates, however, are far from obvious. Finally testing procedures, most importantly for the cointegrating ranks, need to be developed.

Acknowledgements

The first author acknowledges support by the Austrian FWF, project P-14438-INF. We thank Manfred Deistler, Benedikt Pötscher and Søren Johansen for directing our attention to the problem and for useful discussions. We also want to express our gratitude to Thomas Ribarits for meticulous proof reading, noting that any remaining errors or shortcomings are solely our responsibility.

References

Aoki, M. (1990). *State Space Modeling of Time Series*. Springer, New York.

Aoki, M. and A. Havenner, 1989. A Method for Approximate Representation of Vector Valued Time Series and its Relation to Two Alternatives. *Journal of Econometrics* **42**, 181–199.

Aoki, M. and A. Havenner, eds. (1997). *Applications of Computer Aided Time Series Modeling*. Lecture Notes in Statistics. Springer, New York.

Archontakis, F. (1998). An Alternative Proof of Granger's Representation Theorem for I(1) Systems Through Jordan Matrices. *Journal of the Italian Statistical Society* **7**, 111–127.

Banerjee, A., J.J. Dolado, J.W. Galbraith and D.F. Hendry (1993). *Co-integration, Error Correction and the Econometric Analysis of Non-Stationary Data*. Oxford University Press, Oxford.

Bauer, D. and M. Wagner (2002a). Estimating Cointegrated Systems Using Subspace Algorithms. *Journal of Econometrics* **111**, p. 47-84.

Bauer, D. and M. Wagner (2002b). Asymptotic Properties of Pseudo Maximum Likelihood Estimates for Multiple Frequency I(1) Processes. Mimeo.

Bauer, D. and M. Wagner (2003a). On Polynomial Cointegration in the State Space Framework. Submitted to Econometric Theory.

Bauer, D. and M. Wagner (2003b). A Parameterization for Rational Unit Root Processes in the State Space Framework. Mimeo

Econometric Reviews: Special Issue on State Space Modelling (1991). *Econometric Reviews*, **10**, Nr. 1.

Engle, R.F. and C.W.J. Granger (1987). Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica* **55**, 251–276.

Granger, C.W.J. and T.H. Lee (1989). Investigation of Production, Sales and Inventory Relationships using Multicointegration and Non-symmetric Error Correction Models. *Journal of Applied Econometrics* **4**, 145–159.

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

Gregoir, S. (1999a). Multivariate Time Series with Various Hidden Unit Roots, Part I: Integral Operator Algebra and Representation Theorem. *Econometric Theory* **15**, 435–468.

Gregoir, S. (1999b). Multivariate Time Series with Various Hidden Unit Roots, Part II: Estimation and Test. *Econometric Theory* **15**, 469–518.

Hannan, E. and M. Deistler (1988). *The Statistical Theory of Linear Systems*. Wiley, New York.

Johansen, S. (1992). A Representation of Vector Autoregressive Processes Integrated of Order 2. *Econometric Theory* **8**, 188–202.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Auto-Regressive Models*. Oxford University Press, Oxford.

Johansen, S. and E. Schaumburg (1999). Likelihood Analysis of Seasonal Cointegration. *Journal of Econometrics* **88**, 301–339.

Lee, H.S. (1992). Maximum Likelihood Inference on Cointegration and Seasonal Cointegration. *Journal of Econometrics* **54**, 1–47.

Leeb, H. and B.M. Pötscher (2001). The Variance of an Integrated Process need not diverge to Infinity, and Related Results on Partial Sums of Stationary Processes. *Econometric Theory* **17**, 671–685.

Meyer, C.D. (2000). *Matrix Analysis and Applied Linear Algebra*. SIAM, Philadelphia.

Ober, R. (1996). Balanced Canonical Forms. In: *Identification, Adaptation, Learning* (S. Bittanti and G. Picci, eds.), 120–183. Springer, New York.

Stock, J.H. and M.W. Watson (1993). A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica* **61**, 783–820.

A Proof of Theorem 2

It follows from Lemma 1 that for y_t as in the formulation of the theorem a minimal state space representation, say (A, B, C), exists. To show that the restrictions formulated in the theorem select a unique state space realization we proceed as follows: Firstly, we show that starting from any given minimal realization

(A, B, C) there exists a transformation matrix $T \in \mathbb{C}^{n\times n}$, such that $(\hat{A}, \hat{B}, \hat{C}) = (TAT^{-1}, TB, CT^{-1})$ fulfills all restrictions stated. Secondly, we show that this transformation matrix T is unique.
The proof is constructive and recursive: First note that the set of all observationally equivalent minimal systems can be described by $\mathcal{M} = \{T \in \mathbb{C}^{n\times n} : \det T \neq 0\}$. In the course of the proof we sequentially impose restrictions on the system. This results in a corresponding sequence of sets $\mathcal{M}_i$ that describe all observationally equivalent state space realizations that fulfill all restrictions up to step i. As we are sequentially adding restrictions, $\mathcal{M}_{i+1} \subset \mathcal{M}_i$ follows. The proof is completed, if it is shown that the imposition of all restrictions formulated in the theorem implies $\mathcal{M}_\iota = \{I_n\}$, for some suitable index ι. The first restriction imposed is that A has to be in block-diagonal form, with the diagonal blocks corresponding to the unit roots in reordered Jordan normal form as given in (14). This leads to $\mathcal{M}_1 = \{T = \mathrm{diag}(T_u, T_{st}) : T$ nonsingular, $T_u J_u T_u^{-1} = J_u\}$. Due to the block-diagonal structure of J, it follows that in $\mathcal{M}_1$ only block-diagonal matrices are contained, i.e. matrices of the form $T = \mathrm{diag}(T_1, \ldots, T_l, T_{st})$. This fact implies that each block corresponding to one of the unit roots, i.e the blocks J_k, as well as the stationary block, can be analyzed separately. To see the block-diagonal structure note that $TJT^{-1} = J$ implies $TJ^mT^{-1} = J^m$ and thus $Tp(J)T^{-1} = p(J)$ for all polynomials $p(z)$. Using the polynomial $p_k(z) = (J - z_k I)^{m_k}$ one obtains $T(J - z_k I)^{m_k} T^{-1} = (J - z_k I)^{m_k}$. This latter matrix has a null-block as its k-th block-row and block-column respectively, and therefore null-blocks are induced in the off-diagonal blocks of the k-th block-row of T. The same argument repeated for all z_k (and also for the eigenvalues smaller than one in absolute value) shows the block-diagonal structure of T for $TJT^{-1} = J$. It is this fact, already used in the examples, that directly implies that the stable part can be dealt with using any standard canonical form and parameterization developed for the stationary case, e.g. balanced canonical forms (Ober, 1996) or echelon forms (Hannan and Deistler, 1988). Hence, (transforming and) restricting the stationary part to be in a specific canonical form leads to $\mathcal{M}_2 = \{T = \mathrm{diag}(T_1, \ldots, T_l, I) : T_k J_k T_k^{-1} = J_k, k = 1, \ldots, l\}$.
Given the above discussion it is sufficient to look at only one block corresponding to one of the unit roots at a time. For any matrix $T \in \mathcal{M}_2$ for all k the relationship $T_k J_k T_k^{-1} = J_k$ has to hold, which is equivalent to $T_k N_k = N_k T_k$. Now denote with $\tilde{I}_{d_j} = [I_{d_j}, 0^{d_j \times (d_{j+1} - d_j)}]$ and partition the matrix $T_k = [T_{i,j}]$, for $i, j = 1, \ldots, m_k$, according to the partitioning of N_k, i.e. $T_{i,j} \in \mathbb{C}^{d_i \times d_j}$ neglecting again the dependence upon k for notational simplicity both in $T_{i,j}$ and in the indices d_i. The structure of the matrices N_k, see again (14), implies that $T_k N_k = N_k T_k$ leads to $T_{m_k,j} = 0$ for $1 \leq j < m_k$. As a further restriction $T_{i,j}\tilde{I}_{d_j} = \tilde{I}_{d_i} T_{i+1,j+1}$ emerges. Thus, it follows that

$$T_{i+1,j+1} = \begin{bmatrix} T_{i,j} & 0 \\ T^{(2)}_{i+1,j+1} & T^{(1)}_{i+1,j+1} \end{bmatrix}$$

where $T^{(2)}_{i+1,j+1} \in \mathbb{C}^{(d_{i+1} - d_i)\times d_j}$ and $T^{(1)}_{i+1,j+1} \in \mathbb{C}^{(d_{i+1} - d_i)\times (d_{j+1} - d_j)}$. The above relationships and restrictions give a complete characterization of all matrices T_k that fulfill $T_k N_k = N_k T_k$. From $T_{m_k,j} = 0$ for $1 \leq j < m_k$ and the above relation it follows that $T_{i,j} = 0, 1 \leq j < i$, hence all

matrices T_k are upper block-triangular.
Up to now all restrictions on the realization and thus on the set of feasible transformation matrices stem from the invariance requirement of the A-matrix in reordered Jordan normal form. In a final step now again recursively a *unique* transformation matrix T is given, that transforms any minimal realization (A, B, C) of $k(z)$ with the A-matrix in the required format to a system representation where $\hat{C} = CT$ and $\hat{B} = T^{-1}B$ fulfill all restrictions formulated in the theorem. Look again at one block corresponding to one unit root only. Due to the upper block-triangular structure of feasible matrices T_k, the restrictions can be imposed sequentially. Denote with $\hat{C}_k = C_k T_k$ and $\hat{B}_k = T_k^{-1} B_k$ and start the argument in the first block-column of the following equations:

$$\hat{C}_k = C_k T_k = \left[C_k^1, C_k^2, \ldots, C_k^{m_k}\right] \begin{bmatrix} T_{1,1} & T_{1,2} & \cdots & T_{1,m_k} \\ 0 & T_{2,2} & \ldots & T_{2,m_k} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & T_{m_k,m_k} \end{bmatrix}, \tag{15}$$

$$\hat{B}_k = T_k^{-1} B_k = \begin{bmatrix} T_{1,1}^{-1} & x & \cdots & x \\ 0 & T_{2,2}^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & x \\ 0 & \cdots & 0 & T_{m_k,m_k}^{-1} \end{bmatrix} \begin{bmatrix} B_k^1 \\ B_k^2 \\ \vdots \\ B_k^{m_k} \end{bmatrix}. \tag{16}$$

The first block-column of equation (15) simply reads as $\hat{C}_k^{1,E} = C_k^{1,E} T_{1,1}$, the last block-row of equation (16) is given by $\hat{B}_k^{m_k} = T_{m_k,m_k}^{-1} B_k^{m_k}$. From the structure of T_{m_k,m_k} it follows, that

$$T_{m_k,m_k}^{-1} = \begin{bmatrix} T_{1,1}^{-1} & 0 \\ x & x \end{bmatrix}$$

where x denotes entries whose specific value is irrelevant for the argument. Therefore conditions to uniquely determine $T_{1,1}$ from $\hat{C}_k^{1,E} = C_k^{1,E} T_{1,1}, \hat{B}_k^{m_k,1} = T_{1,1}^{-1} B_k^{m_k,1}$ are sought. This is equivalent to finding a unique decomposition of $C_k^{1,E} B_k^{m_k,1}$. One solution to this problem is given in Lemma 2, which leads to the restrictions imposed in the theorem.
Next relabel the state space system, letting (A, B, C) denote the system, where A is in the required Jordan normal form and additionally $(C_k^{1,E})' C_k^{1,E} = I_{d_1}$ and $B_k^{m_k,1}$ is in p.u.t. form. This results in a restricted set $\mathcal{M}_3$, where $T \in \mathcal{M}_3$ fulfills the same restrictions imposed for $\mathcal{M}_2$ and additionally $T_{1,1} = I_{d_1}$ holds.
Now the argument proceeds with (again omitting the index k):

$$T_{1,2} = [T_{1,2}^{(2)} \quad T_{1,2}^{(1)}], \quad T_{2,2} = \begin{bmatrix} I_{d_1} & 0 \\ T_{2,2}^{(2)} & T_{2,2}^{(1)} \end{bmatrix}.$$

Using this information, the second block-column of (15) reads (after the first transformation), itself partitioned into its two sub-blocks as in Lemma 1, as

$$[\hat{C}_k^{2,G}, \hat{C}_k^{2,E}] = [C_k^{1,E} T_{1,2}^{(2)} + C_k^{2,G} + C_k^{2,E} T_{2,2}^{(2)}, C_k^{1,E} T_{1,2}^{(1)} + C_k^{2,E} T_{2,2}^{(1)}] \tag{17}$$

From the equation $\hat{C}_k^{2,G} = C_k^{1,E} T_{1,2}^{(2)} + C_k^{2,G} + C_k^{2,E} T_{2,2}^{(2)}$ it follows that, since $[C_k^{1,E}, C_k^{2,E}]$ has full column rank due to minimality (cf. Lemma 1), it is possible to achieve orthogonality:

$(\hat{C}_k^{2,G})'[C_k^{1,E}, C_k^{2,E}] = 0$. Furthermore the matrices $T_{1,2}^{(2)}$ and $T_{2,2}^{(2)}$ achieving orthogonality are unique.

Next from the second part of the block-equation (17), $\hat{C}_k^{2,E} = C_k^{1,E}T_{1,2}^{(1)} + C_k^{2,E}T_{2,2}^{(1)}$ results. In Theorem 2, $\hat{C}_k^{2,E}$ is required to fulfil $(\hat{C}_k^{2,E})'\hat{C}_k^{2,E} = I$ and to be orthogonal to $C_k^{1,E}$. For each given $T_{2,2}^{(1)}$ there exists a unique $T_{1,2}^{(1)}$, such that $(\hat{C}_k^{2,E})'C_k^{1,E} = 0$. The restriction $(\hat{C}_k^{2,E})'\hat{C}_k^{2,E} = I_{d_2-d_1}$ determines $T_{2,2}^{(1)}$ up to right-multiplication with an orthonormal matrix. The p.u.t. restriction on $\hat{B}_k^{m_k,2}$ leads to a unique matrix $T_{2,2}^{(1)}$ in analogy to the argument given above.

The argument is now repeated also for the other block-columns of equations (15) and (16). The analysis proceeds recursive and the proof is via induction. Thus, assume that up to some $2 \leq j < m_k$ the following restrictions are imposed on C: $(C_k^{j,G})'C_k^{i,E} = 0, i \leq j, (C_k^{j,E})'C_k^{i,E} = 0, i < j, (C_k^{j,E})'C_k^{j,E} = I$ and $B_k^{m_k,i}, i \leq j$ is p.u.t. Concerning the sub-blocks of the matrix T_k these restrictions imply that $T_{i,i} = I_{d_i-d_{i-1}}$ for $i = 1, \ldots, j$, $T_{i,r} = 0$, for $i \leq j, r \leq j, i \neq r$ and

$$T_{i,j+1} = \begin{bmatrix} 0 & 0 \\ T_{i,j+1}^{(2)} & T_{i,j+1}^{(1)} \end{bmatrix}, i \leq j, \quad T_{j+1,j+1} = \begin{bmatrix} I_{d_j} & 0 \\ T_{j+1,j+1}^{(2)} & T_{j+1,j+1}^{(1)} \end{bmatrix}$$

Now exactly the same arguments as discussed above for the first and second block-column lead to the induction from j to $j+1$, which finishes the proof. □