Chapman University
Chapman University Digital Commons

Mathematics, Physics, and Computer Science
Faculty Articles and Research

Science and Technology Faculty Articles and
Research

2010

Linear Stochastic State Space Theory in the White Noise Space Setting

Daniel Alpay
Chapman University, alpay@chapman.edu

David Levanony
Ben-Gurion University of the Negev

Ariel Pinhas
Ben-Gurion University of the Negev

Recommended Citation

Linear Stochastic State Space Theory in the White Noise Space Setting

LINEAR STATE SPACE THEORY IN THE WHITE NOISE SPACE SETTING

DANIEL ALPAY, DAVID LEVANONY, AND ARIEL PINHAS

ABSTRACT. We study state space equations within the white noise space setting. A commutative ring of power series in a countable number of variables plays an important role. Transfer functions are rational functions with coefficients in this commutative ring, and are characterized in a number of ways. A major feature in our approach is the observation that key characteristics of a linear, time invariant, stochastic system are determined by the corresponding characteristics associated with the deterministic part of the system, namely its average behavior.

CONTENTS

1. INTRODUCTION

In a preceding paper, see [1], the first two authors began a study of linear stochastic systems within the framework of the white noise space.

1

There, the emphasis was on stability theorems associated with convolution systems. The present paper is concerned with state space theory. Specifically, we study systems defined by state space equations when randomness is allowed in the matrices defining these equations.

Remark 1.1. *While the discussion to follow is restricted to discrete time, most results apply to continuous time in an obvious way. Continuous time is not explicitly pursued in this work.*

To set the problem in perspective and provide motivation, we first recall some well known facts from linear system theory. There, state space equations of the form

$$(1.1) \qquad \begin{aligned} x_{n+1} &= Ax_n + Bu_n, \quad n = 0, 1, \ldots \\ y_n &= Cx_n + Du_n, \end{aligned}$$

play an important role. In (1.1), $A \in \mathbb{C}^{N \times N}$, $B \in \mathbb{C}^{N \times q}$, $C \in \mathbb{C}^{p \times N}$, $D \in \mathbb{C}^{p \times q}$, the states x_n take values in $\mathbb{C}^N$, the inputs u_n in $\mathbb{C}^q$ and the outputs y_n in $\mathbb{C}^p$. Taking the Z transform, assuming $x_0 = 0$, (1.1) leads to

$$(1.2) \qquad \begin{aligned} \mathscr{X}(\zeta) &= \zeta A \mathscr{X}(\zeta) + \zeta B \mathscr{U}(\zeta) \\ \mathscr{Y}(\zeta) &= C \mathscr{X}(\zeta) + D \mathscr{U}(\zeta), \end{aligned}$$

where the Z transform variable is denoted by ζ, so that

$$(1.3) \qquad \mathscr{Y}(\zeta) = \mathscr{H}(\zeta) \mathscr{U}(\zeta),$$

where

$$\mathscr{Y}(\zeta) = \sum_{n=0}^{\infty} y_n \zeta^n, \quad \mathscr{U}(\zeta) = \sum_{n=0}^{\infty} u_n \zeta^n, \quad \mathscr{X}(\zeta) = \sum_{n=0}^{\infty} x_n \zeta^n,$$

and

$$(1.4) \qquad \mathscr{H}(\zeta) = D + \zeta C (I_N - \zeta A)^{-1} B.$$

The function $\mathscr{H}(\zeta)$ is called the transfer function of the system defined by (1.1). It is also possible to replace $\mathbb{C}$ by a field $\mathfrak{K}$ over the complex numbers. Matrices A, B, C and D then have their entries in $\mathfrak{K}$, and $\mathscr{H}(\zeta)$ makes sense for all ζ such that $(I_N - \zeta A)$ is invertible. The case where the components of A, B, C and D belong to some commutative ring is of special interest. See for instance [15], [17], [16], [11], [12], [5]. This ring is often assumed Noetherian, to enable to formulate results. Even when the ring contains the complex numbers, formula (1.4) does not make sense in general because elements in the ring need not be

invertible in the ring. It will make sense in a normed ring for ζ small enough, as is seen using the power expansion

$$(1.5) \qquad (I_N - \zeta A)^{-1} = \sum_{n=0}^{\infty} \zeta^n A^n.$$

The ring $\mathfrak{R}$ defined below is not a normed ring, but still it will be possible to define expansions of the form (1.5) in it.

As we have explained in our previous paper [1], a Gaussian input into a linear system with nonrandom coefficients, will result in a Gaussian output. In the present series of papers, and in particular in the present work, we aim to model linear Gaussian input-output relations when the underlying linear system is random. Here we allow Gaussian inputs and randomness in the matrices A, B, C and D, in such a way that the outputs remain Gaussian. While indeed a Gaussian input into a linear system with random coefficients cannot be expected to result in a Gaussian output, we will use the white noise space setting (see [9], [10] and Section 2 below), and replace the pointwise product with the Wick product, enabling Gaussian input-output relations when the underlying system has random coefficients. This framework will preserve Gaussian input-output relation, while allowing uncertainty in the form of randomness in the linear system under study. Such a setting may prove useful so as to study a linear state space system with *nonrandom* uncertainties, a system that indeed maintains Gaussian input-output relation. This, by utilizing the Bayesian embedding approach to solve problems associated with a system subjected to a deterministic uncertainty, through solutions to corresponding problems associated with systems with random uncertainties, see e.g. [13].

In the white noise space setting, the space of complex numbers $\mathbb{C}$ is replaced by a space of stochastic distributions called the Kondratiev space, denoted by S_{-1}. This space contains $\mathbb{C}$, is the inductive limit of a certain family of Hilbert spaces (see (2.2) below), and is nuclear; see [10, Definition 2.3.2 (b) p. 30, Lemma 2.8.2 p. 74]. A key element in our formulation is a product defined on S_{-1}, namely the Wick product, denoted by $u \lozenge v$, which reduces to multiplication by a constant when one of the elements u or v is non random. We thus replace the equations (1.1) by

$$(1.6) \qquad \begin{aligned} x_{n+1} &= A \lozenge x_n + B \lozenge u_n \\ y_n &= C \lozenge x_n + D \lozenge u_n \end{aligned}$$

where $A \in (S_{-1})^{N \times N}$, $B \in (S_{-1})^{N \times q}$, $C \in (S_{-1})^{p \times N}$, and $D \in (S_{-1})^{p \times q}$. The states x_n take values in $(S_{-1})^N$, the input u_n in $(S_{-1})^q$ the output in $(S_{-1})^p$.

A fundamental tool in white noise analysis is the Hermite transform

$$F \mapsto \mathbf{I}(F)$$

(see below), which associates to every element in S_{-1}, a power series in a countable number of complex variables,

(1.7) $z = (z_1, z_2, z_3, \ldots)$

and transforms the Wick product into a point-wise product:

(1.8) $\mathbf{I}(F \diamond G)(z) = (\mathbf{I}(F)(z))(\mathbf{I}(G)(z)), \quad \forall F, G \in S_{-1}.$

The image of the Kontradiev space under the Hermite transform is a commutative ring without divisors of zeros (that is, a domain), which we will denote by $\mathfrak{R}$. It is not Noetherian, so most results in system theory on commutative rings cannot be applied. Still, it has a very important property, which allows us to proceed. An $\mathbf{F} \in \mathfrak{R}$ is invertible in $\mathfrak{R}$ if and only if its constant coefficient is non zero (recall that $\mathbf{F}$ is a power series). More generally, for $p \in \mathbb{N}$, an $\mathbf{F} \in \mathfrak{R}^{p \times p}$ will be invertible in $\mathfrak{R}^{p \times p}$ if and only if the matrix $\mathbf{F}(0)$ (which belongs to $\mathbb{C}^{p \times p}$) is invertible; see Theorem 3.1 below. This theorem follows from a non trivial result on the characterization of the range of the Hermite transform, given in [10, Theorem 2.6.11, p. 62].

We now take the Z transform and the Hermite transform of (1.6). The Z transform of the series $(\mathbf{I}(u_n))_{n=0,1,\ldots}$ is denoted by $\mathscr{U}(\zeta, z)$

$$\mathscr{U}(\zeta, z) = \sum_{n=0}^{\infty} \zeta^n (\mathbf{I}(u_n))(z),$$

and similarly for $\mathscr{Y}(\zeta, z)$ and $\mathscr{X}(\zeta, z)$. We obtain:

(1.9)
$$(I_N - \zeta \mathbf{A}(z))\mathscr{X}(\zeta, z) = \mathbf{B}(z)\mathscr{U}(\zeta, z),$$
$$\mathscr{Y}(\zeta, z) = \mathbf{C}(z)\mathscr{X}(\zeta, z) + \mathbf{D}(z)\mathscr{U}(\zeta, z),$$

where

$$\mathbf{A}(z) = \mathbf{I}(A)(z), \ \mathbf{B}(z) = \mathbf{I}(B)(z), \ \mathbf{C}(z) = \mathbf{I}(C)(z), \ \mathbf{D}(z) = \mathbf{I}(D)(z).$$

Since $\mathbf{A}(z)$ is bounded in a certain neighborhood of the origin (see Theorem 3.1 below), the matrix $(I_N - \zeta \mathbf{A}(z))$ is invertible for ζ small enough, and we can write

$$\mathscr{Y}(\zeta, z) = \mathscr{H}(\zeta, z)\mathscr{U}(\zeta, z),$$

where

$$(1.10) \qquad \mathscr{H}(\zeta, z) = \mathbf{D}(z) + \zeta \mathbf{C}(z)(I_N - \zeta \mathbf{A}(z))^{-1} \mathbf{B}(z),$$

is the transfer function of the system defined by the equations (1.6).

Remark 1.2. *When we set $z = 0$ in (1.9) we retrieve (1.4), that is, we are back within the deterministic setting.*

We now view (1.10) as a $\mathfrak{R}$-valued function. Consider ζ such that

$$(1.11) \qquad \det(I_N - \zeta \mathbf{A}(0)) \neq 0.$$

It follows from Theorem 3.3 below that $(I_N - \zeta \mathbf{A})$ is invertible in $\mathfrak{R}$ for such ζ. Therefore the $\mathfrak{R}$-valued function $\mathscr{H}(\zeta)$ given by

$$(\mathscr{H}(\zeta))(z) = \mathscr{H}(\zeta, z),$$

that is,

$$(1.12) \qquad \mathscr{H}(\zeta) = \mathbf{D} + \zeta \mathbf{C}(I_N - \zeta \mathbf{A})^{-1} \mathbf{B}$$

is well defined for ζ satisfying (1.11).

Functions of the form (1.12) will be called *rational functions associated with the white noise space*. We note that in [2] another approach to rational functions, with emphasis on rationality with respect to a finite number of the variables z_k is considered. The purpose of this paper is to give a number of equivalent characterizations of functions of the form (1.12) and to study the notions of controllability, observability and minimality in the setting of the ring $\mathfrak{R}$.

A major non-trivial feature of this work is the observation that key characteristics of a linear, time invariant, stochastic system (e.g. invertibility), are determined by the corresponding characteristics associated with the deterministic part of the system under study, namely its average behavior. For instance, a realization of the perturbed system (1.12) will be observable (see Definition 5.1 below) if the corresponding realization of the unperturbed system, namely with $z = 0$ is observable. See Theorem 5.3.

The paper consists of six sections besides the introduction, and its outline is as follows. We review, as already mentioned, white noise space theory in Section 2. We study the ring $\mathfrak{R}$ in Section 3. Equivalent characterizations of rational functions are given in Section 4. Observable pairs are studied in Section 6. In Section 6 we consider controllable

pairs, and briefly discuss minimal realizations. The last section considers the case where the functions take values in one of the Hilbert spaces which make $\mathfrak{R}$.

2. A BRIEF SURVEY OF WHITE NOISE SPACE ANALYSIS

The starting point to construct the white noise space is the Schwartz space $\mathcal{S}$ of *real-valued* smooth functions which, together with their derivatives, decrease rapidly to zero at infinity. For $s \in \mathcal{S}$, let $\|s\|$ denote its $\mathbf{L}_2(\mathbb{R})$ norm. The function

$$K(s_1 - s_2) = e^{-\frac{\|s_1 - s_2\|^2}{2}}$$

is positive (in the sense of reproducing kernels) for s_1, s_2 running in $\mathcal{S}$. The space is nuclear. By an extension of Bochner's theorem to nuclear spaces due to Minlos (see [14], [7, Théorème 3, p. 311]), there exists a probability measure P on $\mathcal{S}'$ such that

$$K(s) = \int_{\mathcal{S}'} e^{-i\langle s', s\rangle} dP(s'),$$

where we have denoted by $\langle s', s\rangle$ the duality between $\mathcal{S}$ and $\mathcal{S}'$. The real Hilbert space $\mathbf{L}_2(\mathcal{S}', \mathcal{F}, dP)$, where $\mathcal{F}$ is the Borelian σ-algebra, is called the white noise space. We will denote it by $\mathcal{W}$, and its elements by ω, by setting $\Omega = \mathcal{S}'$.

Among all orthogonal Hilbert bases of the white noise space, one plays a special role. It is constructed in terms of Hermite functions and its elements are denoted by H_α, where the index α runs through the set ℓ of sequences $(\alpha_1, \alpha_2, \ldots)$, whose entries are in

$$\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\},$$

and $\alpha_k \neq 0$ for only but a finite number of indices k. Furthermore, with the multi-index notation

$$\alpha! = \alpha_1! \alpha_2! \cdots,$$

we have

(2.1) $$\|H_\alpha\|_{\mathcal{W}}^2 = \alpha!.$$

In view of (2.1), the map

$$H_\alpha \mapsto z^\alpha = z_1^{\alpha_1} z_2^{\alpha_2} z_3^{\alpha_2} \cdots$$

extends to a unitary map between $\mathcal{W}$ and the reproducing kernel Hilbert space with reproducing kernel

$$k(z, w) = e^{\langle z, w \rangle_{\ell_2}} = \sum_{\alpha \in \ell} \frac{z^\alpha w^{*\alpha}}{\alpha!},$$

where z, w run through ℓ_2.

The Wick product in $\mathcal{W}$ is defined by the formula

$$H\alpha \Diamond H_\beta = H_{\alpha + \beta}, \quad \alpha, \beta \in \ell,$$

and the Hermite transform is defined through linearity as

$$(\mathbf{I}(H_\alpha))(z) = z^\alpha.$$

The space $\mathcal{W}$ is too small to be stable under the Wick product, and one defines the Kondratiev space S_{-1}, within which the Wick product is stable. More precisely, S_{-1} is a nuclear space, and is defined as the inductive limit of the increasing family of Hilbert spaces $\mathcal{H}_k, k = 1, 2, \ldots$ of formal series $\sum_{\alpha \in \ell} f_\alpha H_\alpha$ such that

$$(2.2) \qquad \|f\|_k \stackrel{\text{def.}}{=} \left(\sum_{\alpha \in \ell} |f_\alpha|^2 (2\mathbb{N})^{-k\alpha} \right)^{1/2} < \infty,$$

where, for $\beta \in \ell$,

$$(2\mathbb{N})^\beta = 2^{\beta_1} (2 \times 2)^{\beta_2} (2 \times 3)^{\beta_3} \cdots .$$

That the Wick product is stable within S_{-1} is made more precise by Våge's inequality (see [10, Proposition 3.3.2, p. 118]), which we now recall. Let l and k be natural numbers such that $k > l + 1$. Let $h \in \mathcal{H}_l$ and $u \in \mathcal{H}_k$. Then,

$$(2.3) \qquad \|h \Diamond u\|_k \leq A(k - l) \|h\|_l \|u\|_k,$$

where

$$(2.4) \qquad A(k - l) = \sum_{\alpha \in \ell} (2\mathbb{N})^{(l-k)\alpha}.$$

For a proof that $A(k - l)$ is finite, see [10, Proposition 2.3.3, p. 31].

The series $\sum_{\alpha \in \ell} s_\alpha z^\alpha$ will be said to be convergent at z if

$$\sum_{\alpha \in \ell} |s_\alpha| |z|^\alpha < \infty,$$

that is, if it is absolutely convergent, see [10, p. 60]. The following easy lemma will be used below.

Lemma 2.1. *Assume that $f(z) = \sum_{\alpha \in \ell} f_\alpha z^\alpha$ and $g(z) = \sum_{\alpha \in \ell} g_\alpha z^\alpha$ are absolutely convergent power series at z. Then*

$$(2.5) \quad |f(z)g(z)| \leq \sum_{\gamma \in \ell} |z|^\gamma \cdot \Big| \sum_{\substack{\alpha+\beta=\gamma \\ \alpha,\beta \in \ell}} f_\alpha g_\beta \Big| \leq \Big(\sum_{\alpha \in \ell} |f_\alpha||z|^\alpha\Big)\Big(\sum_{\alpha \in \ell} |g_\alpha||z|^\alpha\Big),$$

and in particular, the product fg is an absolutely convergent power series at z, and it holds that

$$f(z)g(z) = \sum_{\gamma \in \ell} z^\gamma \left(\sum_{\substack{\alpha+\beta=\gamma \\ \alpha,\beta \in \ell}} f_\alpha g_\beta \right).$$

Furthermore, $(f(z))^n$ is an absolutely convergent power series at z for all $n \in \mathbb{N}$.

Proof: The power series

$$\sum_{\gamma \in \ell} z^\gamma \left(\sum_{\substack{\alpha+\beta=\gamma \\ \alpha,\beta \in \ell}} f_\alpha g_\beta \right)$$

is absolutely convergent since

$$\sum_{\gamma \in \ell} |z|^\gamma \cdot \Big| \sum_{\substack{\alpha+\beta=\gamma \\ \alpha,\beta \in \ell}} f_\alpha g_\beta \Big| \leq \sum_{\gamma \in \ell} |z|^\gamma \left(\sum_{\substack{\alpha+\beta=\gamma \\ \alpha,\beta \in \ell}} |f_\alpha| \cdot |g_\beta| \right)$$

$$= \Big(\sum_{\alpha \in \ell} |f_\alpha||z|^\alpha\Big)\Big(\sum_{\alpha \in \ell} |g_\alpha||z|^\alpha\Big).$$

$\square$

3. THE RING $\mathbf{I}(S_{-1})$

Consider the image $\mathfrak{R} \overset{\text{def.}}{=} \mathbf{I}(S_{-1})$ under the Hermite transform of the Kondratiev space. This is a space of power series which has been characterized in [10, Theorem 2.6.11, p. 62]. In that statement, $(\mathbb{C}^\mathbb{N})_c$ denotes the space of finite sequences of complex numbers indexed by the integers, and the set $K_q(\delta)$ is defined by

$$(3.1) \qquad K_q(\delta) = \{z \in \mathbb{C}^\mathbb{N} : \sum_{\substack{\alpha \in \ell \\ \alpha \neq (0,0,...)}} |z^\alpha|^2 (2\mathbb{N})^{q\alpha} < \delta^2\}.$$

Note that $\alpha = (0, 0, \ldots)$ is excluded from the sum. See [10, Definition 2.6.4, p. 59].

Theorem 3.1. [10, Theorem 2.6.11, p. 62]

(1) *If* $F(\omega) = \sum_\alpha a_\alpha H_\alpha(\omega) \in S_{-1}$, *then there exist* $q < \infty$, $M_q < \infty$ *such that*

$$(3.2) \qquad |\mathbf{I}(F)(z)| \le \sum_{\alpha \in \ell} |a_\alpha| \, |z^\alpha| \le M_q \left(\sum_{\alpha \in \ell} (2\mathbb{N})^{q\alpha} \, |z^\alpha|^2 \right)^{\frac{1}{2}}$$

for all $z \in (\mathbb{C}^{\mathbb{N}})_c$. *In particular,* $\mathbf{I}(F)$ *is a bounded analytic function on* $K_q(\delta)$ *for all* $\delta < \infty$.

(2) *Conversely, suppose* $g(z) = \sum_\alpha b_\alpha z^\alpha$ *is a given power series of* $z \in (\mathbb{C}^{\mathbb{N}})_c$ *with* $b_\alpha \in \mathbb{C}$, *with* $q < \infty$ *and* $\delta > 0$ *such that* $g(z)$ *is absolutely convergent for* $z \in K_q(\delta)$ *and*

$$\sup_{z \in K_q(\delta)} |g(z)| < \infty.$$

Then there exists a unique $G \in S_{-1}$ *such that* $\mathbf{I}(G) = g$, *namely*

$$G(\omega) = \sum_{\alpha \in \ell} b_\alpha H_\alpha(\omega).$$

A characterization of convergent sequences in S_{-1} is given in the following theorem proved in [10]. It will be used in particular in the proof of Proposition 5.2.

Theorem 3.2. [10, Theorem 2.8.1, p. 74] *A sequence of elements* $F^{(n)}$ *in the Kondratiev space* S_{-1} *converges to* $F \in S_{-1}$ *if there exist* $\delta > 0$ *and* $q < \infty$ *such that* $\mathbf{I}(F^{(n)})$ *converges to* $\mathbf{I}(F)$ *pointwise boundedly, or equivalently, uniformly, in* $K_q(\delta)$.

The main result of this section is:

Theorem 3.3. $\mathfrak{R}$ *is a commutative ring, which contains* $\mathbb{C}$ *and has no divisors of zero. Furthermore, let* $x(t) = \sum_{n=0}^{\infty} x_n t^n$ *be a power series, with strictly positive radius of convergence. Let* $p \in \mathbb{N}$. *Then for every* $\mathbf{r} \in \mathfrak{R}^{p \times p}$ *such that* $\mathbf{r}(0) = 0_{p \times p}$, *the series*

$$(x(\mathbf{r}))(z) = \sum_{n=0}^{\infty} x_n (\mathbf{r}(z))^n$$

converges to a limit in $\mathfrak{R}^{p \times p}$. *If* $y(t) = \sum_{n=0}^{\infty} y_n t^n$ *is another such power series, then*

$$(3.3) \qquad (xy)(\mathbf{r}) = x(\mathbf{r})y(\mathbf{r}), \quad \forall \mathbf{r} \in \mathfrak{R}.$$

In particular, an element $\mathbf{s}$ *is invertible in* $\mathfrak{R}^{p \times p}$ *if and only if* $\mathbf{s}(0)$ *is invertible.*

Proof: To simplify the notation we give a proof for $p = 1$. The fact that we have a ring follows from the formula (1.8). The way to prove the second claim is to use Theorem 3.1 to show that the *a-priori* formal power series

$$\sum_{n=0}^{\infty} x_n (\mathbf{r}(z))^n$$

is in fact the image under the Hermite transform of an element in S_{-1}. Since $\mathbf{r}$ is the image of an element of S_{-1} under the Hermite transform, it satisfies (3.2) for some $q \in \mathbb{N}$ and a constant $M_q > 0$. Let r_x be the radius of convergence of the power series defining x (and similarly for r_y below). We choose δ such that

$$M_q \delta \overset{\text{def.}}{=} \rho < r_x,$$

Then, by (3.2), we have for $z \in K_q(\delta)$,

$$|\mathbf{r}(z)| \leq \rho,$$

hence

$$|\sum_{n=1}^{\infty} x_n (\mathbf{r}(z))^n| \leq \sum_{n=1}^{\infty} |x_n| \rho^n, \quad z \in K_q(\delta).$$

We conclude the proof by using Theorem 3.1. We first prove (3.3). By the preceding arguments we know that $x(\mathbf{r})$, $y(\mathbf{r})$, and $(xy)(\mathbf{r})$, are well defined. On the other hand, for $|\mathbf{r}(z)| < \min(r_x, r_y)$ we have:

$$(x(\mathbf{r}(z))(y(\mathbf{r}(z)) = \sum_{n=0}^{\infty} (\sum_{p=0}^{n} x_p y_{n-p})(\mathbf{r}(z))^n$$
$$= ((xy)(\mathbf{r}))(z).$$

We now turn to the last statement. Assume that $\mathbf{s}$ is invertible in $\mathfrak{R}$, and let $\mathbf{u} \in \mathbf{R}$ be such that $\mathbf{su} = 1$. Then, in particular, $\mathbf{s}(0)\mathbf{u}(0) = 1$, so $\mathbf{s}(0) \neq 0$. Conversely, we can assume without loss of generality, that $\mathbf{s}(0) = 1$. It suffices then to take in (3.3) $x(t) = 1 - t$, $y(t) = (1 - t)^{-1}$, and $\mathbf{r} = \mathbf{u} - 1$. (Note that $\mathbf{r}(0) = 0$.) $\qquad \square$

4. RATIONAL FUNCTIONS

Let $f(\zeta, z) = \sum_{n=0}^{\infty} f_n(z) \zeta^n \in \mathfrak{R}^{p \times p}(\zeta)$ be a power series with coefficients in $\mathfrak{R}^{p \times q}$. Define

$$R_0 f(\zeta, z) = \frac{f(\zeta, z) - f(0, z)}{\zeta}.$$

Theorem 4.1. *Let $\mathcal{H}(\zeta) = \sum_{n=0}^{\infty} \mathbf{f}_n \zeta^n \in \mathfrak{R}^{p \times q}((\zeta)))$ be a formal power series. Then the following are equivalent:*

(1) Components of $\mathcal{H}$ are obtained by adding, multiplying and dividing polynomials of $\mathfrak{R}[\zeta]$, with division being performed only when the constant coefficient is invertible in $\mathfrak{R}$.

(2) $\mathcal{H}$ admits a realization in the form of (1.12), with coefficients matrices having entries in $\mathfrak{R}$.

(3) The formal power series converges in a neighborhood of the origin, and there exists a finite number M such that for every $n \geq M$ the function $R_0^n \mathcal{H}$ is a linear combination of $R_0 \mathcal{H}, \ldots R_0^{M-1} \mathcal{H}$ with coefficients in $\mathfrak{R}$.

Proof: We first note that elements of the form (1.12) are convergent power series in $\mathbb{R}$ and not only formal power series, as follows from Theorem 3.3. Elements $\mathcal{H} \in \mathfrak{R}^{p \times q}(\zeta)$ of the form

$$(4.4) \qquad \mathcal{H}(\zeta) = \mathbf{D} \quad \text{or} \quad \mathcal{H}(\zeta) = \zeta \mathbf{C},$$

where $\mathbf{C}, \mathbf{D} \in \mathfrak{R}^{p \times q}$ are clearly in the form (1.12). Furthermore, as is well known, if $p = q$ and

$$\mathscr{H}(\zeta) = \mathbf{D} + \zeta \mathbf{C}(I_N - \zeta \mathbf{A})^{-1} \mathbf{B} \in \mathfrak{R}^{p \times p}(\zeta)$$

with $\mathbf{D}$ invertible, then we have:

$$(\mathscr{H}(\zeta))^{-1} = \mathbf{D}^{-1} - \zeta \mathbf{D}^{-1} \mathbf{C}(I_N - \zeta \mathbf{A}^{\times})^{-1} \mathbf{B} \mathbf{D}^{-1},$$

where

$$\mathbf{A}^{\times} = \mathbf{A} - \mathbf{B} \mathbf{D}^{-1} \mathbf{C}.$$

Furthermore, if

$$\mathscr{H}_1(\zeta) = \mathbf{D}_1 + \zeta \mathbf{C}_1 (I_{N_1} - \zeta \mathbf{A}_1)^{-1} \mathbf{B}_1 \in \mathfrak{R}^{p_1 \times s}(\zeta)$$

and

$$\mathscr{H}_2(\zeta) = \mathbf{D}_2 + \zeta \mathbf{C}_2 (I_{N_2} - \zeta \mathbf{A}_2)^{-1} \mathbf{B}_2 \in \mathfrak{R}^{s \times q_2}(\zeta),$$

then

$$(\mathscr{H}_1 \mathscr{H}_2)(\zeta) = \mathbf{D} + \zeta \mathbf{C}(I_N - \zeta \mathbf{A})^{-1} \mathbf{B}, \quad N = N_1 + N_2,$$

with $\mathbf{D} = \mathbf{D}_1 \mathbf{D}_2$ and

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{B}_1 \mathbf{C}_2 \\ 0 & \mathbf{A}_2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \mathbf{B}_1 \mathbf{D}_2 \\ \mathbf{B}_2 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \mathbf{C}_1 & \mathbf{D}_1 \mathbf{C}_2 \end{pmatrix}.$$

A sum of matrices is a special case of a product, as follows from the formula

$$\mathbf{M}_1 + \mathbf{M}_2 = \begin{pmatrix} \mathbf{M}_1 & I_p \end{pmatrix} \begin{pmatrix} I_q \\ \mathbf{M}_2 \end{pmatrix},$$

where $\mathbf{M}_1$ and $\mathbf{M}_2 \in \mathfrak{R}^{p \times q}$.

See for instance [3] for some of these formulas when the coefficients are complex. It follows from these formulas that any $\mathfrak{R}$-valued function (that is, when $p = q = 1$) which is obtained by addition, multiplication and, when defined, inversion, of functions of the form (4.4), is of the form (1.12). The matrix-valued case is obtained by concatenation using the formulas

$$\begin{pmatrix} \mathscr{H}_1 & \mathscr{H}_2 \end{pmatrix} (\zeta) =$$

$$= \begin{pmatrix} \mathbf{D}_1 & \mathbf{D}_2 \end{pmatrix} + \zeta \begin{pmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{pmatrix} \left(I_{N_1+N_2} - \zeta \begin{pmatrix} \mathbf{A}_1 & 0 \\ 0 & \mathbf{A}_2 \end{pmatrix} \right)^{-1} \begin{pmatrix} \mathbf{B}_1 & 0 \\ 0 & \mathbf{B}_2 \end{pmatrix},$$

and

$$\begin{pmatrix} \mathscr{H}_1 \\ \mathscr{H}_2 \end{pmatrix} (\zeta) =$$

$$= \begin{pmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \end{pmatrix} + \zeta \begin{pmatrix} \mathbf{C}_1 & 0 \\ 0 & \mathbf{C}_2 \end{pmatrix} \left(I_{N_1+N_2} - \zeta \begin{pmatrix} \mathbf{A}_1 & 0 \\ 0 & \mathbf{A}_2 \end{pmatrix} \right)^{-1} \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix}$$

for two functions $\mathscr{H}_1$ and $\mathscr{H}_2$ of appropriate dimensions which admit a realization. Thus, (1) implies (2). We now prove that (2) implies (3). Let $\mathscr{H}$ be of the form (1.12). Then,

$$R_0^n \mathscr{H}(\zeta) = \mathbf{C}(I_N - \zeta\mathbf{A})^{-1}\mathbf{A}^{n-1}\mathbf{B}, \quad n = 1, 2, \ldots$$

But the Cayley-Hamilton theorem holds in any commutative ring (see for instance [5, p. 14], [6, Theorem 4.3, p. 120], [4, p. A III.107]). Therefore there exist an $M \in \mathbb{N}$, a monic polynomial p of degree M with coefficients in $\mathfrak{R}$, such that $p(\mathbf{A}) = 0$. It follows that for $n \geq M$, the function $R_0^n \mathscr{H}$ is a linear combination of $1, R_0\mathscr{H}, \ldots R_0^{M-1}\mathscr{H}$ with coefficients in $\mathfrak{R}$.

We now assume that (3) is in force and prove that (1) holds. First, assume that $\mathcal{H}$ is $\mathfrak{R}$-valued (as opposed to $\mathfrak{R}^{p \times q}$-valued). By hypothesis there exists a matrix $\mathbf{A} \in \mathfrak{R}^{M \times M}$ such that

$$R_0 \begin{pmatrix} 1 & R_0\mathcal{H} & \cdots & R_0^{M-1}\mathcal{H} \end{pmatrix} = \begin{pmatrix} 1 & R_0\mathcal{H} & \cdots & R_0^{M-1}\mathcal{H} \end{pmatrix} \mathbf{A}.$$

Hence

$$\begin{pmatrix} 1 & R_0\mathcal{H} & \cdots & R_0^{M-1}\mathcal{H} \end{pmatrix} = \begin{pmatrix} 1 & R_0\mathcal{H} & \cdots & R_0^{M-1}\mathcal{H} \end{pmatrix} (0)(I_M - \zeta\mathbf{A})^{-1}.$$

Thus

$$R_0\mathcal{H} = \mathbf{C}(I_M - \zeta\mathbf{A})^{-1}\mathbf{B},$$

with

$$\mathbf{C} = \begin{pmatrix} 1 & R_0 \mathcal{H} & \cdots & R_0^{M-1} \mathcal{H} \end{pmatrix} (0) \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

and the result follows. The matrix-valued case is treated in much the same way. $\qquad\square$

5. Observable pairs

Consider an $\mathfrak{R}$-valued function of the complex variable ζ, of the form (1.12):

$$\mathcal{H}(\zeta) = \mathbf{D} + \zeta \mathbf{C}(I_N - \zeta \mathbf{A})^{-1} \mathbf{B}.$$

By Theorem 3.3, we know that $\mathcal{H}$ is well defined, in particular for ζ such that

$$\det(I_N - \zeta \mathbf{A}(0)) \neq 0.$$

Setting $z = 0$ in (1.12) we get the *unperturbed* transfer function, which motivates Theorem 5.3 below. We first give a definition and a proposition.

Definition 5.1. *The pair $(\mathbf{C}, \mathbf{A}) \in \mathfrak{R}^{p \times N} \times \mathfrak{R}^{N \times N}$ is called observable if the map*

$$\mathbf{f} \mapsto \begin{pmatrix} \mathbf{Cf} & \mathbf{CAf} & \mathbf{CA}^2 \mathbf{f} & \cdots \end{pmatrix}$$

is injective from $\mathfrak{R}^n$ into $(\mathfrak{R}^p)^{\mathbb{N}}$.

See [5, §2.2 p. 58].

Equivalently we have:

Proposition 5.2. *Realization (1.12) is observable if and only if (with $\mathbf{f} \in \mathfrak{R}^N$)*

$$\mathbf{C}(I_N - \zeta \mathbf{A})^{-1} \mathbf{f} \equiv 0_{\mathfrak{R}}^{p \times N} \implies \mathbf{f} = 0_{\mathfrak{R}}^N.$$

Proof: By Theorem 3.3 with $x(t) = (1 - \zeta t)^{-1}$, we have:

$$(I_N - \zeta \mathbf{A})^{-1} = \sum_{p=0}^{\infty} \zeta^p \mathbf{A}^p.$$

Therefore, and using Theorem 3.2,

$$\mathbf{C}(I_N - \zeta\mathbf{A})^{-1} = \mathbf{C}\sum_{n=0}^{\infty}\zeta^p\mathbf{A}^p.$$

$$= \sum_{n=0}^{\infty}\zeta^p\mathbf{C}\mathbf{A}^p.$$

$\square$

Theorem 5.3. *Assume that the realization*

(5.1) $\mathscr{H}(\zeta, 0) = \mathbf{D}(0) + \zeta\mathbf{C}(0)(I_N - \zeta\mathbf{A}(0))^{-1}\mathbf{B}(0)$

is observable. Then realization (1.12) *is observable.*

Proof: Assume first that realization (5.1) is observable, and let

$$\mathbf{f}(z) = \sum_{\alpha\in\ell} f_\alpha z^\alpha, \quad f_\alpha \in \mathbb{C}^N,$$

be such that

(5.2) $\mathbf{C}(I_N - \zeta\mathbf{A})^{-1}\mathbf{f} \equiv 0.$

To prove that the realization (1.12) is observable we need to show that all coefficients f_α in the expansion $\mathbf{f}(z) = \sum_{\alpha\in\ell} f_\alpha z^\alpha$, are identically zero. Since (5.1) is assumed observable, setting $z = 0$ in (5.2) leads to $f_0 = 0$. Let us now put $z = (z_1, z_2, \ldots, z_p, 0, 0, \ldots)$ in (5.2) with $p \geq 1$ and differentiate with respect to z_1. We obtain (with $'$ denoting differentiation with respect to z_1)

(5.3) $\mathbf{C}'(z)(I_N - \zeta\mathbf{A}(z))^{-1}\mathbf{f}(z) + \mathbf{C}(z)\left((I_N - \zeta\mathbf{A}(z))^{-1}\right)'\mathbf{f}(z)+$
$$+\mathbf{C}(z)(I_N - \zeta\mathbf{A}(z))^{-1}\mathbf{f}'(z) \equiv 0.$$

Setting $z_1 = z_2 = \cdots = z_p = 0$ we obtain that

$$\mathbf{C}(0)(I_N - \zeta\mathbf{A}(0))^{-1}f_{(1,0,0,\ldots)} \equiv 0,$$

and hence $f_{(1,0,0,\ldots)} = 0$ since the pair $(\mathbf{C}(0), \mathbf{A}(0))$ is observable. Differentiating in turn (5.3) with respect to z_1, we obtain an expression of the form

(5.4) $\mathbf{X}(z) + \mathbf{C}(z)(I_N - \zeta\mathbf{A}(z))^{-1}f''(z) \equiv 0,$

where $\mathbf{X}$ is a finite sum of the form

$$\mathbf{X}(z) = \sum_{j=1}^{M}\mathbf{U}_j(z)f^{(r_j)}(z)$$

where $r_j \in \{0, 1\}$ and $\mathbf{U}_j(z)$ is analytic in z_1 and may depend on ζ. Setting $z_1 = 0$ in (5.4) and taking into account that $f_0 = f_{(1,0,0,\dots)} = 0$, we obtain that

$$\mathbf{C}(0)(I_N - \zeta\mathbf{A}(0))^{-1}f_{(2,0,0,\dots)} \equiv 0,$$

and hence $f_{(2,0,0,\dots)} = 0$. More generally, an easy induction argument shows that the α_1-th derivative of (5.2) is of the form

(5.5) $$\mathbf{X}(z) + \mathbf{C}(z)(I_N - \zeta\mathbf{A}(z))^{-1}f^{(\alpha_1)}(z) \equiv 0,$$

where $\mathbf{X}$ is of the form

$$\mathbf{X}(z) = \sum_{j=1}^{M} \mathbf{U}_j(z)\mathbf{f}^{(n_j)}(z),$$

$\mathbf{U}_j$ being analytic in z_1 and $n_j \in \{0, \dots, \alpha_1 - 1\}$. Setting $z_1 = 0$ in (5.5) we obtain that $f_{(\alpha_1,0,0,\dots)} = 0$.

Similarly, by setting $z = (0, z_2, 0, \dots)$, and more generally

$$z = (0, 0, 0, \dots, z_j, 0, \dots)$$

in (5.2), and differentiating, we obtain that $f_\alpha = 0$ for all $\alpha \in \ell$ which have only one non-zero component. We now prove that $f_{(1,1,0,0,\dots)} = 0$. To that end, set $z = (z_1, z_2, z_3, \cdots, z_p, 0, \dots)$, with $p \geq 2$, in (5.2) and differentiate this equation with respect to z_1 and z_2. We obtain an equation of the form

(5.6) $$\mathbf{X}(z) + \mathbf{C}(z)(I_N - \zeta\mathbf{A}(z))^{-1}\frac{\partial^2 \mathbf{f}}{\partial z_1 \partial z_2}(z) \equiv 0,$$

where now $\mathbf{X}$ is a finite sum of elements of the form $\mathbf{U}(z)f(z)$ and $\mathbf{U}(z)\frac{\partial f}{\partial z_j}(z)$, with $j \in \{1, 2\}$ and $\mathbf{U}$ analytic in $z_1, z_2, \dots, z_p$. The fact that

$$f_0 = f_{(1,0,0,\dots)} = f_{(0,1,0,\dots)} = 0$$

implies that $\mathbf{X}(0) \equiv 0$. Setting $z_1 = z_2 = 0$ in (5.6) then leads to

$$\mathbf{C}(0)(I_N - \zeta\mathbf{A}(0))^{-1}f_{(1,1,0,\dots)} \equiv 0,$$

and hence $f_{(1,1,0,\dots)} = 0$, where we have used the observability of the pair $(\mathbf{C}(0), \mathbf{A}(0))$. By successive differentiation and setting $z = 0$ we obtain that $f_{(\alpha_1,\alpha_2,0,0,\dots)} = 0$ for every choice of natural integers α_1 and α_2. The fact that all other coordinates f_α are zero, and hence that the pair $(\mathbf{C}, \mathbf{A})$ is observable, is shown by induction as follows:

Induction hypothesis: *For $N \in \mathbb{N}$, it holds that*

(5.7) $$f_{(\alpha_1,\alpha_2,\dots,\alpha_N,0,0,\dots)} = 0, \quad \forall(\alpha_1, \alpha_2, \dots, \alpha_N) \in (\mathbb{N}_0)^N,$$

and

(5.8)
$$\frac{\partial^{\alpha_1+\cdots+\alpha_N}}{\partial z_1^{\alpha_1}\partial z_2^{\alpha_2}\cdots\partial z_N^{\alpha_N}}\mathbf{C}(z)(I_N - \zeta\mathbf{A}(z))^{-1} =$$

$$= \mathbf{X}_{\alpha_1,\ldots,\alpha_N}(z) + \mathbf{C}(z)(I_N - \zeta\mathbf{A}(z))^{-1}\frac{\partial^{\alpha_1+\cdots+\alpha_N}}{\partial z_1^{\alpha_1}\partial z_2^{\alpha_2}\cdots\partial z_N^{\alpha_N}}\mathbf{f}(z),$$

where

(5.9)
$$z = (z_1, z_2, \ldots, z_p, 0, 0, \ldots) \quad with \quad p \geq N,$$

and $\mathbf{X}_{\alpha_1,\ldots,\alpha_N}(z)$ is of the form

(5.10)
$$\mathbf{X}_{\alpha_1,\ldots,\alpha_N}(z) = \sum_{j=1}^{M}\mathbf{U}_j(z)\frac{\partial^{\beta_1^{(j)}+\cdots+\beta_N^{(j)}}}{\partial z_1^{\beta_1^{(j)}}\partial z_2^{\beta_2^{(j)}}\cdots\partial z_N^{\beta_N^{(j)}}}\mathbf{f}(z),$$

where $\beta_i^{(j)} \leq \alpha_i$ for $i = 1, \ldots, N$ and

(5.11)
$$\beta_1^{(j)} + \cdots + \beta_N^{(j)} < \alpha_1 + \cdots + \alpha_N,$$

with the functions $\mathbf{U}_j$ analytic in the variables $z_1, \ldots, z_p$.

The induction hypothesis holds for $N = 1$, as we have shown above. Assume that it holds at rank N. We take $p \geq N + 1$ in (5.9) and differentiate (5.8) with respect to z_{N+1}. Since

$$\frac{\partial}{\partial z_{N+1}}\mathbf{U}_j(z)\frac{\partial^{\beta_1^{(j)}+\cdots+\beta_N^{(j)}}}{\partial z_1^{\beta_1^{(j)}}\partial z_2^{\beta_2^{(j)}}\cdots\partial z_N^{\beta_N^{(j)}}}\mathbf{f}(z) =$$

$$= \left(\frac{\partial}{\partial z_{N+1}}\mathbf{U}_j(z)\right)\frac{\partial^{\beta_1^{(j)}+\cdots+\beta_N^{(j)}}}{\partial z_1^{\beta_1^{(j)}}\partial z_2^{\beta_2^{(j)}}\cdots\partial z_N^{\beta_N^{(j)}}}\mathbf{f}(z)+$$

$$+ \mathbf{U}_j(z)\left(\frac{\partial^{\beta_1^{(j)}+\cdots+\beta_N^{(j)}+1}}{\partial z_1^{\beta_1^{(j)}}\partial z_2^{\beta_2^{(j)}}\cdots\partial z_N^{\beta_N^{(j)}}z_{N+1}}\mathbf{f}(z)\right),$$

the term

$$\frac{\partial\mathbf{X}_{\alpha_1,\ldots,\alpha_N}(z)}{\partial z_{N+1}}$$

is of the form

(5.12)
$$\frac{\partial\mathbf{X}_{\alpha_1,\ldots,\alpha_N}(z)}{\partial z_{N+1}} = \sum_{j=1}^{P}\mathbf{V}_j(z)\frac{\partial^{\beta_1^{(j)}+\cdots+\beta_N^{(j)}+1}}{\partial z_1^{\beta_1^{(j)}}\partial z_2^{\beta_2^{(j)}}\cdots\partial z_N^{\beta_N^{(j)}}\partial z_{N+1}}\mathbf{f}(z),$$

where P is possibly different from M above, the $\mathbf{V}_j$ are analytic in the variables $z_1, \ldots, z_p$, and the $\beta_i^{(j)}$ are as above. Differentiating the term

$$\mathbf{C}(z)(I_N - \zeta \mathbf{A}(z))^{-1} \frac{\partial^{\alpha_1 + \cdots + \alpha_N}}{\partial z_1^{\alpha_1} \partial z_2^{\alpha_2} \cdots \partial z_N^{\alpha_N}} \mathbf{f}(z)$$

in (5.8) with respect to z_{N+1} we obtain a sum of two terms. The first is,

(5.13) $$\left(\frac{\partial}{\partial z_{N+1}} \mathbf{C}(z)(I_N - \zeta \mathbf{A}(z))^{-1} \right) \frac{\partial^{\alpha_1 + \cdots + \alpha_N}}{\partial z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_N^{\alpha_N}} \mathbf{f}(z),$$

while the second takes the form

$$\mathbf{C}(z)(I_N - \zeta \mathbf{A}(z))^{-1} \frac{\partial^{\alpha_1 + \cdots + \alpha_N + 1}}{\partial z_1^{\alpha_1} \partial z_2^{\alpha_2} \cdots \partial z_N^{\alpha_N} z_{N+1}} \mathbf{f}(z)$$

This proves (5.8) for $(\alpha_1, \ldots, \alpha_N, 1)$, with $\mathbf{X}_{\alpha_1, \ldots, \alpha_N, 1}$ being the sum of (5.13) and of (5.12), that is

(5.14)
$$\frac{\partial^{\alpha_1 + \cdots + \alpha_N + 1}}{\partial z_1^{\alpha_1} \partial z_2^{\alpha_2} \cdots \partial z_N^{\alpha_N} \partial z_{N+1}} \mathbf{C}(z)(I_N - \zeta \mathbf{A}(z))^{-1} =$$

$$= \mathbf{X}_{\alpha_1, \ldots, \alpha_N, 1}(z) + \mathbf{C}(z)(I_N - \zeta \mathbf{A}(z))^{-1} \frac{\partial^{\alpha_1 + \cdots + \alpha_N + 1}}{\partial z_1^{\alpha_1} \partial z_2^{\alpha_2} \cdots \partial z_N^{\alpha_N} \partial z_{N+1}} \mathbf{f}(z).$$

Setting $z_1 = \cdots = z_p = 0$ in that expression, we obtain

$$\mathbf{C}(0)(I_N - \zeta \mathbf{A}(0))^{-1} f_{(\alpha_1, \ldots, \alpha_N, 1, 0, 0 \ldots)} \equiv 0,$$

and hence $f_{(\alpha_1, \ldots, \alpha_N, 1, 0, 0 \ldots)} = 0$. Differentiating (5.8) a finite number of times with respect to z_{N+1}, a similar argument will show that

$$f_{(\alpha_1, \ldots, \alpha_N, \alpha_{N+1}, 0, 0 \ldots)} = 0, \quad \forall \alpha_{N+1} \in \mathbb{N}.$$

$\square$

Remark 5.4. *The converse to the previous theorem does not hold. That is, the observability of the pair $(\mathbf{C}, \mathbf{A})$ does not imply the observability of the pair $(\mathbf{C}(0), \mathbf{A}(0))$. As an example, take $N = 1$ and*

$$\mathbf{C}(z) = z_1, \quad \mathbf{A}(z) = 1.$$

The pair $(\mathbf{C}, \mathbf{A})$ is observable since, for $\mathbf{f} \in \mathfrak{R}$,

$$\frac{z_1}{1 - \zeta} \mathbf{f}(z) \equiv 0 \implies \mathbf{f} = 0.$$

But the pair

$$(\mathbf{C}(0), \mathbf{A}(0)) = (0, 1)$$

is not observable.

6. Controllable pairs and minimal realizations

In this section we study controllable pairs and minimal realizations within the setting of the ring $\mathfrak{R}$. We first recall that, given a commutative ring R, one of the characterization for a pair $(A, B) \in R^{N \times N} \times R^{N \times q}$ of matrices to be controllable (or reachable) is that the columns of the matrix

$$\begin{pmatrix} B & AB & \cdots & A^{N-1}B \end{pmatrix}$$

generate R^N; see [5, p. 55].

In the classical case (that is, for the complex numbers, or more generally, for the case of a field), it is well known that the pair (C, A) is observable if and only if the pair (A^T, C^T) is controllable (with ^{T} denoting transpose). This duality principle does not hold in general in the case of an arbitrary commutative ring. Only the following direction holds:

Theorem 6.1. [5, Theorem 2.7, p. 59] *Let R be a commutative ring and let $(C, A) \in R^{p \times N} \times R^{N \times N}$. Assume the pair (A^T, C^T) controllable. Then the pair (C, A) is observable.*

As explained in [5, p. 59], the lack of duality comes form the fact that an homomorphism of modules (say f, from the R-module M_1 into the R-module M_2) can be injective without being residually injective. Recall that residual injectivity means that, for every maximal ideal I of R, the induced map from M_1/IR into M_2/IR is injective when f is injective.

Theorem 6.1 does not help us to study controllability based on observability. Furthermore, in [8, Theorem 2.3 p. 178], it is shown that a necessary and sufficient condition on a commutative ring for the duality principle to hold for all pairs is that every finitely generated faithful ideal of the ring contains a unit. As a corollary, the authors of [8] state:

Proposition 6.2. [8, Corollary 2.4 p. 179]). *If the duality principle holds in a commutative ring, then the ring is a total quotient ring.*

For the purpose of the present paper we do not need to recall the definition of a faithful ideal (see [8, Theorem 1.5 (ii), p. 177]). The total quotient ring of a commutative ring R is the set of formal fractions associated with the set of elements of R which are not divisors of zero; see [5, p. 35]. Thus, in the case of a ring without divisors of zero (as is the case for the ring $\mathfrak{R}$) the total quotient ring is equal to the quotient field associated with the ring; see [5, p. 35]. Since $\mathfrak{R}$ is not a field, it

follows that the duality principle is not satisfied on it.

After these general preliminaries, let us study controllability and minimality in the setting of the ring $\mathfrak{R}$. Let us repeat the definition of controllability: The pair $(\mathbf{A}, \mathbf{B}) \in \mathfrak{R}^{N \times N} \times \mathfrak{R}^{N \times q}$ is said to be controllable (or reachable) if the columns of the matrix

$$\begin{pmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{N-1}\mathbf{B} \end{pmatrix}$$

generate $\mathfrak{R}^N$. See [5, p. 55]. We therefore have:

Proposition 6.3. *Assume the pair $(\mathbf{A}, \mathbf{B}) \in \mathfrak{R}^{N \times N} \times \mathfrak{R}^{N \times q}$ to be controllable. Then the pair $(\mathbf{A}(0), \mathbf{B}(0)) \in \mathbb{C}^{N \times N} \times \mathbb{C}^{N \times q}$ is controllable.*

Proof: Since $\mathbb{C}^N \subset \mathfrak{R}^N$, for every $f \in \mathbb{C}^N$ there exists $\mathbf{a} \in \mathbf{R}^{Nq}$ such that

$$f = \begin{pmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{N-1}\mathbf{B} \end{pmatrix} (z)\mathbf{a}(z).$$

Setting $z = 0$ in this equality we get the controllability of the pair $(\mathbf{A}(0), \mathbf{B}(0))$. $\qquad\square$

Remark 6.4. *We note the difference between Theorem 5.3 and Proposition 6.3. In the former, observability at $z = 0$ implies observability in $\mathfrak{R}$. In the latter, controllability in $\mathfrak{R}$ implies controllability at $z = 0$.*

The converse of Proposition 6.3 would be an analogue of Theorem 5.3 for the case of controllable pairs. But this is not possible for the ring $\mathfrak{R}$, in view of Proposition 6.2, since $\mathfrak{R}$ is different from its total quotient ring (which is in fact its quotient field since $\mathfrak{R}$ has no divisors of zero).

Still, we can give a counterpart of Theorem 5.3 for controllable and minimal realizations with the following ad-hoc definitions:

Definition 6.5. *Realization (1.12) will be called $\mathfrak{R}$-controllable if the following condition holds: Let $\mathbf{f} \in \mathfrak{R}^{1 \times N}$. Then:*

$$\mathbf{f}(I_N - \zeta\mathbf{A})^{-1}\mathbf{B} \equiv 0_{\mathfrak{R}}^{1 \times q} \Longrightarrow \mathbf{f} = 0_{\mathfrak{R}}^{1 \times N}.$$

A realization will be called $\mathfrak{R}$-minimal if it is both observable and $\mathfrak{R}$-controllable.

We can then state:

Theorem 6.6. *Assume that realization (5.1)*

$$\mathscr{H}(\zeta, 0) = \mathbf{D}(0) + \zeta\mathbf{C}(0)(I_N - \zeta\mathbf{A}(0))^{-1}\mathbf{B}(0)$$

is controllable (resp. minimal). Then realization (1.12) is $\mathfrak{R}$-controllable (resp. $\mathfrak{R}$-minimal).

Proof: The first statement is proved as Theorem 5.3. The second statement follows then from the definition of minimality. $\square$

7. HILBERT-SPACE VALUED TRANSFER FUNCTIONS

We recall that the Hilbert spaces $\mathcal{H}_k$ have been defined above by the finiteness of the norm (2.2). By $\mathbf{I}(\mathcal{H}_k)$ we mean the image of $\mathcal{H}_k$ under the Hermite transform. We note that

$$\mathfrak{R} = \cup_{k=1}^{\infty}\mathbf{I}(\mathcal{H}_k).$$

Theorem 7.1. *Let $\mathcal{H}$ be given by realization* (1.12), *and let l, k be natural numbers such that $k > l + 1$. Assume that, in the state space equations* (1.6), *the entries of A and C are in $\mathcal{H}_l$ and the entries of B and D are in $\mathcal{H}_k$. Then, the transfer function $\mathscr{H}$ is $\mathbf{I}(\mathcal{H}_k)$-valued.*

Proof: Inequality (2.3) expresses the fact that the multiplication operator

$$T_h \ : \ u \mapsto h \lozenge u$$

is a bounded map from the Hilbert space $\mathcal{H}_k$ into itself. Therefore the entries of the $\mathfrak{R}^{p \times q}$-valued function $\mathbf{C}\mathbf{A}^n\mathbf{B}$ are in $\mathbf{I}(\mathcal{H}_k)$. To conclude the proof, it remains to show that for every complex number ζ such that $(I_N - \zeta\mathbf{A})$ is invertible, the power series

$$\sum_{n=0}^{\infty} \zeta^n \mathbf{C}\mathbf{A}^n\mathbf{B}$$

converges in $\mathbf{I}(\mathcal{H}_k)$ to $\mathscr{H}(\zeta)$. But this is a consequence of Theorem 3.3. $\square$

Using once more Våge's inequality (2.3) we have:

Corollary 7.2. *Let now $m > k + 1$, where k is as in the previous theorem. Then, the operator of multiplication by $\mathscr{H}$ sends $\mathbf{I}(\mathcal{H}_m)$-valued signals into $\mathbf{I}(\mathcal{H}_m)$-valued signals.*

REFERENCES

[1] D. Alpay and D. Levanony. Linear stochastic systems: a white noise approach. To appear in *Acta Applicandae Mathematicae*, 2010.DOI 10.1007/s10440-009-9461-1.
[2] D. Alpay and D. Levanony. Rational functions associated with the white noise space and related topics. *Potential Analysis*, 29:195–220, 2008.
[3] H. Bart, I. Gohberg, and M.A. Kaashoek. *Minimal factorization of matrix and operator functions*, volume 1 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 1979.

[4] N. Bourbaki. *Éléments de mathématique. Algèbre. Chapitres 1 à 3*. Hermann, Paris, 1970.

[5] J. W. Brewer, J. W. Bunce, and F. S. Van Vleck. *Linear systems over commutative rings*, volume 104 of *Lecture Notes in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 1986.

[6] D. Eisenbud. *Commutative algebra*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. With a view toward algebraic geometry.

[7] I.M. Guelfand and N.Y. Vilenkin. *Les distributions. Tome 4: Applications de l'analyse harmonique*. Collection Universitaire de Mathématiques, No. 23. Dunod, Paris, 1967.

[8] J. Á. Hermida Alonso and T. Sánchez-Giralda. On the duality principle for linear dynamical systems over commutative rings. *Linear Algebra Appl.*, 139:175–180, 1990.

[9] T. Hida, H. Kuo, J. Potthoff, and L. Streit. *White noise*, volume 253 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1993. An infinite-dimensional calculus.

[10] H. Holden, B. Øksendal, J. Ubøe, and T. Zhang. *Stochastic partial differential equations*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, 1996.

[11] Naoharu Ito and Hiroshi Inaba. Dynamic feedback (A, B)-invariant submodules for linear systems over commutative Noetherian domains. *Linear Algebra Appl.*, 282(1-3):123–129, 1998.

[12] R. E. Kalman. Advanced theory of linear systems. In *Topics in Mathematical System Theory*, pages 237–339. McGraw-Hill, New York, 1969.

[13] D. Levanony and P. Caines. Stochastic Lagrangian adaptive LQG control. In *Stochastic theory and control (Lawrence, KS, 2001)*, volume 280 of *Lecture Notes in Control and Inform. Sci.*, pages 283–300. Springer, Berlin, 2002.

[14] R. A. Minlos. Generalized random processes and their extension to a measure. In *Selected Transl. Math. Statist. and Prob., Vol. 3*, pages 291–313. Amer. Math. Soc., Providence, R.I., 1963.

[15] Y. Rouchaleau and E.D. Sontag. On the existence of minimal realizations of linear dynamical systems over Noetherian integral domains. *J. Comput. System Sci.*, 18(1):65–75, 1979.

[16] E.D. Sontag. Linear systems over commutative rings: A survey. *Ricerche di Automatica*, 7:1–34, 1976.

[17] E.D. Sontag and Y. Rouchaleau. Sur les anneaux de Fatou forts. *C. R. Acad. Sci. Paris*, 284(5):A331–A333, 1977.

(DA) Department of mathematics, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 84105, Israel
E-mail address: dany@math.bgu.ac.il

(DL) Department of electrical engineering, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 84105, Israel
E-mail address: levanony@ee.bgu.ac.il

(AP) Department of mathematics, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 84105, Israel
E-mail address: arielp@bgu.ac.il