“Solutions Properties and Techniques for Implicit Systems”

by

Athanasios A Pantelous

Thesis submitted for the award of the degree of

Doctor of Philosophy

in

Systems and Control Theory

CONTROL ENGINEERING RESEARCH CENTER

SCHOOL OF ENGINEERING AND MATHEMATICAL SCIENCES

CITY UNIVERSITY

LONDON EC1V 0HB

OCTOBER 2013

“Solutions Properties and Techniques for Implicit Systems”

Στον Πατέρα μου, τη Μητέρα μου …

Declaration

The University Librarian of the City University may allow this thesis to be copied in whole or in part without any reference to the author. This permission covers only single copies, made for study purposes, subject to normal conditions of acknowledgements.

Table of Contents

Frequently Used Notations

∞:	infinity
ε:	absolute error
Ω:	sample space
ω:	event, outcome
$a \in A$:	a element of set A
$I_A(\cdot)$:	indicator function of event A
$\mathcal{P}$:	probability measure
$\mathcal{A}$:	$\sigma-$ algebra
$(\Omega, \mathcal{F}, \mathcal{P})$:	probability space
$\mathbb{R}^n$:	n-dimensional Euclidean space
$\mathbb{R}_+$:	positive real line
$\mathbb{F}$:	Field
$\mathcal{M}(n \times m; \mathbb{F})$:	algebra of $n \times m$-matrices with elements over $\mathbb{F}$
$\mathcal{D}'_{n-1}$:	space of Dirac distribution having derivatives up to an order $n-1$
$\mathcal{D}$:	space of infinitely differentiable complex-valued functions on $\mathbb{F}$
$\mathcal{B}(\mathcal{D})$:	Borel σ-field
$C^\infty(\cdot)$:	set of smooth functions
$L^2(\cdot)$:	space of quadratically integrable functions
a.s.:	almost surely
$[a,b]$:	closed interval from a to b
(a,b):	open interval from a to b
LTI :	Linear Time Invariant
$\underline{x}(t)$:	state vector parameter
$\underline{u}(t)$:	input vector parameter
$\underline{y}(t)$:	output vector parameter

$\underline{a}$: input signal

$\|\cdot\|_2$: Euclidean distance

$\delta(t)$: Dirac function

$\delta^{(k)}(t)$: k^{th} derivative of the Dirac δ-function

$\delta_a(t)$: *nascent* delta function

$\varphi(t)$: test function

$A=\left[a_{ij}\right]_{1\le i,j\le n}$: constant matrix in $\mathbb{R}^{n\times n}$

$diag\{\cdot\}$: diagonal matrix

$\det A$: determinant of matrix A

J: Jordan canonical form of matrix A.

$\mathbb{O}$: zero matrix

$V_{m,n}\equiv V_n(\lambda_1,\lambda_2,...,\lambda_m)$: Vandermonde matrix, which is defined in terms of scalars $\lambda_1,\lambda_2,...,\lambda_m\in\mathbb{R}$ (where $m\neq n$)

A^{-1}: inverse matrix A

$A^{\dagger}$: *Moore-Penrose* inverse of a matrix A

A^{D}: *Drazin inverse* of square matrix A

$A^{\{1,2,3\}}$: {1, 2, 3}- generalized inverse of matrix A.

H_q: nilpotent matrix

q^*: the annihilation index of H_q

$Ind(A)$: the smallest non-negative integer such as $rank\left(A^{Ind(A)}\right)=rank\left(A^{Ind(A)+1}\right)$

$C_n([\vdots])$: the n-order compound matrix of $[\vdots]$

$\mathrm{Re}(\lambda)$: real part of a complex number λ

$\mathrm{Im}(\lambda)$: imaginary part of a complex number λ

*: conjugate transpose index of the relevant matrix

$\circ$: action

$\coprod_{\cdot}$: order left multiplication of matrices

LU: "Lower Upper" factorization

$\rho_j = \max_{z=1,2,\ldots,d} \mu_z$: index of annihilation for the eigenvalue μ_z.

$\underset{\sim}{n}$: set $\{1,2,...,n\}$

ODE: ordinary differential equation

NBV: normalized bounded variation function

GDDS: generalized differential delay system

DDDS: differential systems with distributed delay

DDE: delay differential equation

DAS: differential-algebraic system

SVD: singular value decomposition

e.d.: elementary divisors

z.e.d.: zero elementary divisors

nz. f.e.d.: nonzero finite elementary divisors

i.e.d.: infinite elementary divisors

c.m.i.: column minimal indices

r.m.i.: row minimal indices

$\mathbb{F}[s,\hat{s}]$: ring of polynomials in s and $\hat{s}=1/s$ with coefficients on $\mathbb{F}$

$sF-G$: the pencil (F,G)

e_g: (I_n, I_n) identity element of the group $(g,*)$ on the set of $\mathcal{L}^r_{n,n}$

$\mathcal{E}_s$: a strict equivalence relation

$\mathcal{L}^r_{n,n}$: the set on $n \times n$ regular pencils

$\Delta(z)$: the characteristic matrix *defined by* $zI - A\int_{t_o}^{t_o+\tau} e^{-zt} d\mu(t)$.

μ: NBV function $[t_o, t_o+\tau] \to \mathbb{C}^{n\times n}$

sBm: standard Brownian motion

fBm: fructional Brownian motion

$\mathcal{H} \in (0,1)$: Hurst parameter

$W^{\mathcal{H}}(t)$: Representation of fBm of Hurst parameter

$\Gamma(a)$: Gamma function

F_X: distribution function of random variable X

$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$: the Gaussian probability density function

$\Phi(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$: the Gaussian cumulative distribution function

$\int_{-\infty}^{K(t,\sigma)} \phi(x) dx$: the *cumulative distribution function* (cdf) of a random variable $X \sim N(0,1)$ evaluated at the upper limit of the integral $K(t,\sigma)$, denoting the probability that $X \le K(t,\sigma)$.

$K(t,\sigma)$: t / σ

$\int ... \, dt$: (Lebesgue, Riemann) integral

$\int ... \, dW_t$: Ito integral

$\int_{L(\gamma)} ... \, dz$: so-called principal value integral $\lim_{\omega \to \infty} \int_{\gamma - i\omega}^{\gamma + i\omega} \ldots \, dz$

$W(t)$: Wiener process at time t

$\mathrm{I}\Gamma_{\mathcal{H}}$: a transformation which transforms the white noise (the derivative of sBm) to fractional noise (the derivative of fBm)

$\langle \xi, \varphi \rangle = \int_{t_o}^{T} \varphi(s) \, dW(s)$: generalized stochastic (random) process

$\langle \sigma\delta_S, \varphi \rangle = \int_S \varphi(\xi) \sigma(\xi) \delta S$: the linear continuous functional $\sigma\delta_S$ on the space $\mathcal{D}$ of infinitely differentiable complex-valued functions on $\mathbb{F}$ with compact support

List of Figures

Acknowledgements

These few lines seem painfully inadequate for me to thank all those who have helped me and supported me through all the years of my research.

I hope to continue supporting me for many more years.

Definitely, I owe an immeasurable debt to my supervisor Prof. Nicos Karcanias, who has guided me through that time. His valuable comments and constructive criticism, his continuous support and understanding, became my cornerstone the last six years.

I am really very grateful to my external and internal examiner Prof. Gaynor Taylor and Prof. Martin Newby respectively, who have afforded me considerable assistance in enhancing both the quality of the findings and the clarity of my PhD thesis presentation.

Words cannot express my gratitude to my beloved family who gave me the opportunity to pursue my dreams in the UK. Your unconditional love and support is my safety blanket.

Thanks a lot!

I am really very grateful

Chapter 1

Introduction – Contribution

In economic theory, input-output analysis has been developed for the description of the production of a multi-sector economy. An input-output model is a quantitative economic technique that represents the interdependencies between different branches of a national economy or different regional economies. In the region of input-output economics, many models were established to describe the real economics (see for example, Leontief (1966) and R. O'Connor, E.W. Henry (1975)).

The economic traditional Leontief dynamic input-output model is described by

$$\underline{x}_k = A\underline{x}_k + L\left[\underline{x}_{k+1} - \underline{x}_k\right] + \underline{g}_k ,$$

where the vector $\underline{x}_k = \left[x_{1,k} \quad x_{2,k} \quad \cdots \quad x_{n,k} \right]^T$ is the total output vector and $x_{i,k}$ is the total output from sector $1 \le i \le n$. The vector $\underline{g}_k$ is the *final net product* and $g_{i,k}$ denotes the *final net product* of sector $1 \le i \le n$. The matrix $A = \left[a_{ij} \right], 1 \le i, j \le n$, is the *direct consumption coefficient* matrix (also called the Leontief intput-output matrix) and $L = \left[l_{ij} \right], 1 \le i, j \le n$, is the *capital coefficient* matrix. Initially, this model has been studied in discrete-time where the matrices A and L have been assumed to be constant over time, i.e. that market and technology do not change under the considered time period. The discrete-time version of this input-output model has been used widely because of the nature of the problem (see for example Luenberger and Arbel (1977), Szyld (1985) and references therein). However, as it is true, the production of a nation (or a factory) in real economic terms is in fact continuous. Thus, an analogous continuous in time dynamic input-output model of the form

$$\underline{x}(t) = A\underline{x}(t) + L\dot{\underline{x}}(t) + \underline{g}(t),\ t > 0,$$

has been also proposed and studied in the literature of economic modelling (see Fleissner (1990), Jodar and Merello (2010), Zhao and Jiang (2009) and references therein). In this input-output model, the capital coefficient matrix L is not always invertible, since the product of some sectors can not be treated as a capital product or/and utilized by others (for example, agriculture, service sectors also do not produce durable goods etc.). In fact, the element l_{ij} of matrix L represents the amount of stock of commodity i, as a capital good, that sector j must have on hand for each unit of production. Since not every sector produces significant capital goods, it is common for some rows of the matrix L to contain only zero elements. System above, which can be formally written as

$$L\dot{\underline{x}}(t) = M\underline{x}(t) + \underline{f}(t),\ t > 0,$$

where $M = I - A$, $\underline{f}(t) = -\underline{g}(t)$ and L is a non-invertible constant matrix, is a linear time invariant (LTI) singular system and it is often called degenerate (or of descriptor type). It is useful here to emphasize that the parameter $\underline{f}(t),\ t \geq 0$ can be considered either as just a (regular or irregular) disturbance or as the Leontief dynamic input-output model's control vector, as long as the quantity of the final net product can be affected by various ways.

However, in Engineering now, very recently, in the very interesting paper by Karcanias (2008), we can see that the always challenging problem of integrated engineering design, which is strongly linked to systems and control theory (and their applications), is revealed as a typical structure evolution process. Such processes emerge in many application domains and in the engineering context in problems such as integrated system design, integrated operations, re-engineering, lifecycle design issues, networks etc. Thus, it has been shown that the formation of the system, which is finally used for control design, evolves during the earlier design stages. The process synthesis and the overall instrumentation are also critical stages of the evolutionary process that shapes the final system structure and thus the potential for control design. Karcanias (2008) aims at

revealing the control theory context of the evolutionary mechanism in overall system design.

Familiarizing with the proposed results by Karcanias (2008), we can also claim that the characteristics and the nature of the process synthesis and the global instrumentation depend on the type of available models. Thus, there are models where some of the internal variables are classified into potential inputs, outputs, internal variables and referred to as oriented models, or models where no classification has been made of the internal variables these are called *implicit* models. All such models may be used for selection of effective sets of inputs and outputs, they are referred to as progenitor models and they may be classified as: (a) Internal Models, (b) External Models and (c) Internal–External Models.

As we will see later, in this PhD thesis we are mostly interested in *internal* models. These models, see also Lewis (1989), have a very long history and are primarily described in terms of first order ordinary nonlinear equations and they are the standard state-space descriptions of the implicit type

$$F\left(\underline{x}(t), \dot{\underline{x}}(t)\right) = \underline{0} \ \text{(or } F\left(\underline{x}_k, \underline{x}_{k+1}\right) = \underline{0}\text{)},$$

where $\underline{x}(t)$ is the vector of all internal model variables. In the linear case, the above reduces to matrix pencil model can be defined by

$$E\dot{\underline{x}}(t) = A\underline{x}(t) \ \text{(or } E\underline{x}_{k+1} = A\underline{x}_k\text{)}.$$

When the inputs $\underline{u}(t)$, outputs $\underline{y}(t)$ have been defined, then the nonlinear control model is defined by

$$F\left(\underline{x}(t), \dot{\underline{x}}(t), \underline{u}(t)\right) = \underline{0}, \ \underline{y}(t) = G\left(\underline{x}(t), \dot{\underline{x}}(t), \underline{u}(t)\right)$$

$$\text{(or } F\left(\underline{x}_k, \underline{x}_{k+1}, \underline{u}_k\right) = \underline{0}, \ \underline{y}_k = G\left(\underline{x}_k, \underline{x}_{k+1}, \underline{u}_k\right)\text{)},$$

and in the linear case is expressed by the singular model

$$E\underline{\dot{x}}(t) = A\underline{x}(t) + B\underline{u}(t), \quad \underline{y}(t) = C\underline{x}(t) \text{ (or } \quad E\underline{x}_{k+1} = A\underline{x}_k + B\underline{u}_k, \quad \underline{y}_{k+1} = C\underline{x}_k \text{).}$$

In the literature, linear internal models are called Descriptor (differential/difference) systems (or generalized systems or differential algebraic systems), and they have a key role in the modelling and simulation process of constrained dynamical systems in many *applications*. Thus, such systems have been intensively studied, theoretically as well as numerically, in the last decades. For a systematic and comprehensive exposition of important aspects regarding the theory, the numerical treatment and many applications of first order descriptor differential/difference systems, see for instance Campbell (1980, 1982), Karcanias and Hayton (1982), Griepentrog and März (1986), Lewis (1986), Dai (1989), Hairer, Lubich and Roche (1989), Willems (1989), Brenan, Campbell and Petzold (1996), Eich-Soellner and Führer (1998), Kunkel and Mehrmann (2006), Karcanias (2008), Pantelous, Zimbidis and Kalogeropoulos (2010) and the references therein.

The strong motivation behind this PhD thesis is based on the significant extension of the continuous in time Leontief model in order to bring it closer to reality and to make it as general as it is possible covering many interesting cases and phenomena. Thus, in the present PhD thesis, the study of the derived equations is being considered in order to cover differenct very general case that the total output, the total demand, as well as the entrances of the coefficient matrices to depend on different economic parameters such as the individual and cooperative decision processes, the resource limitations, the environmental and geographical constraints, the institutional and legal requirements and the purely random fluctuations. For this purpose, as it will become clearer with the next paragraphs and sections, different types of implicit systems will be proposed, considered and developed, In most cases, the existence and the solvability will be investigated. Our task is motivated theoretically, as we are not providing numerical algorithms.

Analytically, this PhD thesis deals with the following 5 interesting topics:

A) Impulsive Control: Change the Initial State in Zero Time

In the 2nd Chapter, a solid methodology has been proposed for approximating the distributional trajectory that transfers the state of a linear differential system in (almost) zero time by using the impulse-function and its derivatives. The motivation behind this section is related to investigate the change of the status of a economical system almost instantly, i.e. in zero time (for instance, the change of the nominal interest rate from Central Banks).

The new results are based on the research work proposed by Gupta and Hasdorff in 1963. As a first step, using some basic elements of measure theory, we show that the input vector has to be a linear combination of the δ-function of Dirac and its derivatives, i.e.

$$u_o(t) = \sum_{k=0}^{n} a_k \delta^{(k)}(t).$$

Our approach is based on the approximation of the Dirac function using the Gaussian (Normal) function. However, since the methodology is quite general, the present results can be further modified and extended using other different kinds of approximations of the Dirac function, for instance Airy functions. Concluding, the present work has involved the following three distinct problems:

(i) We have started with the impulsive trajectory that transfers the origin to a point in the state space and used this as the central point motivating the need to approximate distributions by smooth functions.

(ii) After that, we have examined the family of Gaussian functions, which may be used to approximate distributions and we have defined an appropriate Euclidean metric to measure how good the approximation is and investigates the link of the σ parameter of Gauss functions to the time and, inevitably, to the distance from the desired initial state.

(iii) We have pre-determined the minimal time required for achieving a solution to the above standard controllability problem in terms of approximations to the distributional solutions, by using Gaussian families for the approximation. Finally, the CIZT algorithm has been proposed for the calculation of the coefficients of our input function.

B) Generalized Inverses: Vandermonde and Special Matrix

In the 3rd Chapter, three main results have been proposed and discussed: First, we have provided a (quasi) LU factorization, and secondly we have calculated analytically the generalized inverses of the rectangular (and square) Vandermonde matrix, which is defined in terms of scalars $\lambda_1, \lambda_2, ..., \lambda_m \in \mathbb{R}$ (where $m \neq n$) by the following expression:

$$V_{m,n} \equiv V_n\left(\lambda_1, \lambda_2, ..., \lambda_m\right) \triangleq \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_m & \cdots & \lambda_m^{n-1} \end{bmatrix}.$$

Finally, similar results with the Vandermonde matrix have been presented for a special structure matrix, i.e.

$$\begin{bmatrix} 1 & \mu & \mu^2 & \mu^3 & * & \cdots & * & * & \cdots & \mu^{n-1} \\ 1 & \lambda & \lambda^2 & \lambda^3 & * & \cdots & * & * & \cdots & \lambda^{n-1} \\ 0 & 1 & 2\lambda & 3\lambda^2 & * & \cdots & * & * & \cdots & (n-1)\lambda^{n-2} \\ 0 & 0 & 1 & 3\lambda & * & \cdots & * & * & \cdots & (n-1)(n-2)\lambda^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & * & \cdots & \frac{1}{(m-1)}\frac{d^{m-1}}{d\lambda_j^{m-1}}\left(\lambda^{n-1}\right) \end{bmatrix}.$$

Both matrices have appeared recently in control and system theory's literature, where the change of the initial state of a linear system in zero time is required, see also 2nd Chapter. This is a complementary to the 2nd chapter as it considers the case that the economical system might be descriptor, see for more details Pantelous et al. (2010).

C) Descriptor Delay Differential Systems: Solutions Properties

In the 4th Chapter, a special class of generalized regular differential delay systems with constant coefficients, i.e.

$$E\underline{x}'(t) = A\int_{t_o}^{t_o+\tau} \underline{x}(t-s)d\mu(s) + B\underline{u}(t)$$

is extensively studied, where E, $A \in \mathbb{C}^{n\times n}$, $\det E = 0$ and $B \in \mathbb{C}^{n\times l}$ are constant matrices, $\underline{u} \in C\left([t_o,\infty),\mathbb{C}^l\right)$ is a control (column vector function of dimension l), and $t \geq t_o$, where $\tau > 0$ is constant. Furthermore, there exists a unique *normalized bounded variation (**NBV**) function (or distribution)* $\mu : [t_o, t_o+\tau] \to \mathbb{C}$.

In practice, these kinds of systems can model the size of a population or the value of an investment. By considering the regular Matrix Pencil approach, we finally decompose it into two subsystems, whose solutions are obtained. Moreover, since the initial function is given, the corresponding initial value problem is uniquely solvable.

Finally, an illustrative application is presented using **dde23** MatLab (m–) file based on the explicit Runge - Kutta method.

D) Generalized Neutral Differential Multi-Delay Systems: Solutions Properties

In the 5th Chapter, the generalized singular neutral differential multi-delay system with constant coefficients, i.e.

$$E\underline{x}'(t) = A\underline{x}(t) - \sum_{i=1}^{\rho} B_i\underline{x}'(t-\tau_i) + \sum_{i=1}^{\rho} C_i\underline{x}(t-\tau_i) + D\underline{u}(t)$$

where, E, A and $B_i, C_i \in \mathbb{C}^{n\times n}$ for $i = 1,2,\ldots,\rho$ are constant matrices, with $\det E = 0$, and the input function $\underline{u} \in C^1[t_o,\infty)$ (column vector function of dimension l) is assumed to consist of all differentiable functions whose derivative is continuous (*continuously differentiable*), and $t \geq t_o$, $0 < \tau_1 < \tau_2 < \ldots < \tau_\rho$ are constants, is studied.

These kinds of systems are inherent in many economical, physical and engineering phenomena. Using the Matrix Pencil theory we decompose it into five subsystems, whose solutions are obtained. Moreover, the form of the initial function is given, so the corresponding initial value problem is uniquely solvable.

E) Generalized Stochastic Differential Delay Systems: Generalized Random Processes

In the last Chapter, we consider the generalized linear regular stochastic differential delay system with constant coefficients and two simultaneous external differentiable and non differentiable perturbations, i.e.

$$E\underline{x}'(t) = A\underline{x}(t) + B\underline{x}(t-\tau) + C\underline{u}(t) + D\underline{f}(t) + R\underline{w}(t)$$

where $\underline{w}$ is a (fractional) white noise of dimension s, $\underline{f} \in C^{\infty}[t_o, \infty)$ is a smooth input (column vector function of dimension k), and $\underline{u} \in C[t_o, \infty)$ is a control (column vector function of dimension l). The $E, A, B \in \mathbb{C}^{n\times n}$, with $\det E = 0$, $C \in \mathbb{C}^{n\times l}$, $D \in \mathbb{C}^{n\times k}$, and $R \in \mathbb{C}^{n\times s}$ are constant matrices.

These kinds of systems are inherent in many application fields; among them we mention fluid dynamics, the modelling of multi body mechanisms, economics and the problem of protein folding. Using regular Matrix Pencil theory, we decompose it into two subsystems, whose solutions are obtained as generalized processes.

Moreover, the form of the initial function is given, so the corresponding initial value problem is uniquely solvable. Finally, two illustrative applications are presented using white noise and fractional white noise, respectively.

Analytically, we use standard Brownian motion (sBm), $\{W(t), t \geq 0\}$, on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Moreover, if $\varphi \in C^{\infty}(U)$ is as a test function, then

$$\langle \xi, \varphi \rangle = \int_{t_o}^{T} \varphi(s)\, dW(s)$$

in the sense of equality in law. More precisely, the Wiener integral is defined as the extension to $L^2(\mathbb{R}_+)$ of white noise, see Kuo (1975) and Borodin and Salminen (2002) for more details about the construction of the Wiener integral as the extension of white noise.

Moreover, we show a way to adapt the traditional white noise calculus to the fractional white noise case. Firstly, we recall that if $\{W(t), t \geq 0\}$ is a standard Brownian motion (sBm) on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$, then it is defined

$$W^{\mathcal{H}}(t) = \int_{t_o}^{T} Z_{\mathcal{H}}(t,s) dW(s),\ t \geq 0$$

which is the representation of fBm of Hurst parameter $\mathcal{H} \in (0,1)$ on the same probability space (see Hu, 2005, for more details) , where

$$Z_{\mathcal{H}}(t,s) = \begin{cases} k_{\mathcal{H}} \left[\left(\frac{t}{s}\right)^{\mathcal{H}-\frac{1}{2}} (t-s)^{\mathcal{H}-\frac{1}{2}} - \left(\mathcal{H}-\frac{1}{2}\right) s^{\frac{1}{2}-\mathcal{H}} \int_s^t u^{\mathcal{H}-\frac{3}{2}} (u-s)^{\mathcal{H}-\frac{1}{2}} du \right], & \text{if } 0 < \mathcal{H} < 1/2 \\ \left(\mathcal{H}-\frac{1}{2}\right) k_H s^{\frac{1}{2}-\mathcal{H}} \int_s^t u^{\mathcal{H}-\frac{1}{2}} (u-s)^{\mathcal{H}-\frac{3}{2}} du, & \text{if } 1/2 < \mathcal{H} < 1 \end{cases}$$

Also $k_{\mathcal{H}} = \sqrt{\dfrac{2\mathcal{H}\Gamma\left(\frac{3}{2}-\mathcal{H}\right)}{\Gamma\left(\mathcal{H}+\frac{1}{2}\right)\Gamma(2-2\mathcal{H})}}$, $\Gamma(a) = \int_o^{\infty} s^{a-1} e^{-s} ds$ is the gamma function.

Chapter 2

Approximating Distributional Behaviour of Linear Systems Using Gaussian Function and its Derivatives

2.1 Introduction

The use of Dirac $\delta-$distributions in the study of LTI differential system problems is a well-established subject going back to Gupta and Hasdorff (1963), Zadeh and Desoer (1963), Verghese (1979), Verghese and Kailath (1979), Karcanias and Kouvaritakis (1979), Campbell (1980, 1982), Willems (1981), Jaffe and Karcanias (1981), Cobb (1982, 1983), Karcanias and Hayton (1982), Karcanias and Kalogeropoulos (1989), Willems (1991), and references there in. The work so far has dealt with the characterisation of basic system properties such as infinite poles and zeros Verghese (1979), Verghese and Kailath (1979) for regular and singular (implicit) systems, as well as the study of fundamental control problems where the solution is expressed in terms of Dirac $\delta-$ distributions. Typical problems are those dealing with the notions of almost (A,B)-invariance and almost controllability subspaces Willems (1981), Jaffe and Karcanias (1981).

In particular, the study of distributional solutions plays a key role in many areas in systems and control such as:

(i) Controllability, Observability.

(ii) Infinite zero characteristic behaviour.

(iii) Almost invariant subspaces, almost controllability spaces.

(iv) Dynamics of singular systems etc.

The distributional characterization is also linked to solution of a number of control problems. The solution of such problems have theoretical significance, given that distributions cannot be constructed and only smooth functions can be constructed and implemented. The idea of approximating distributional inputs with smooth functions that achieve a similar control objective was first introduced by Gupta and Hasdorff (1963), Gupta (1966).

In the present section, which actually extends and provides a *rigorous* reformulation of the early ideas presented in Gupta and Hasdorff (1963), we consider the problem of approximating Dirac distributions with smooth functions of infinite support, and more specifically using the Gaussian density and its derivatives. Thus, a *new* methodology is proposed for approximating the distributional trajectory that transfers the state of a LTI differential system in (almost-) zero time by using an impulsive input. So, with the new approach, the following three distinct problems are addressed:

(i) First, we determine the (unique) impulsive input signal (and its smooth approximation) which transfers the state of the system from the origin to an arbitrary point in state space in (almost-) zero time, subject to appropriate controllability assumptions.

(ii) Then, we calculate the approximation error in the state-trajectories of the system resulting from substituting impulsive input signals by smooth signals. Thus, for the very first time (according to the author's knowledge), the optimal choice of two significant parameters of the Gaussian distribution and its derivatives, i.e. time t and volatility σ, characterising the family of all smooth approximating functions, is considered and eventually an elegance formula combining them is derived.

(iii) Finally, we solve two state-space maximum-distance problems in the context of (almost) zero-time state-transition. These correspond to two different types of constraints on the coefficients of the impulsive input signal and its smooth approximation, involving the Euclidian and infinity norms of the vector of coefficients. It is interested for further consideration that we can prove that both problems are

tractable and can be solved via an SVD and the solution of a quadratic programming problem with box constraints.

More specifically, in sub-section 2.2, we present the problem formulation for a LTI differential system. In sub-section 2.3, we provide a brief review of the different types of approximations of distributions by smooth functions and explain their significance in characterizing system properties. In sub-section 2.4, we assume that the system is controllable, and under this assumption we establish an interesting connection between a time-parameter t and a volatility parameter σ of the Gaussian density function used in the approximation. It turns out that the fraction t/σ can be fixed (to a sufficiently large value) and in this case parameter t (or σ) parameter controls the state-transition time and the accuracy of the approximation (which can be interpreted probabilistically). A new algorithm is proposed for calculating the smooth input signal that approximates the distributional input which transfers the origin of the state-space to an arbitrary target point (subject to a controllability assumption) and the distance (Euclidean norm) between the actual terminal state and the target state; this distance is subsequently minimized subject to magnitude constraints imposed on the coefficients of the control signal. Finally, in sub-section 2.5 we define the distance from the origin using the Euclidean norm. Moreover, we consider the problem of maximising the distance from the origin with constrained input. Sub-section 2.6 concludes the paper.

2.2 Problem Definition

We consider the linear time invariant (LTI) system

$$\underline{x}'(t) = A\underline{x}(t) + \underline{b}u_o(t), \tag{2.2.1}$$

where $\underline{x}(t) \in \mathcal{C}^{\infty}\left(\mathbb{F}, \mathcal{M}(n\times 1;\mathbb{F})\right)$ (*smooth* function over the field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, whose elements belong to the algebra $\mathcal{M}(n\times 1;\mathbb{F})$), and $u_o(t) \in \mathcal{D}'_{n-1}$ (where $\mathcal{D}'_{n-1}$ is the *space* of Dirac distribution having derivatives up to an order $n-1$) are the state vector, and the impulsive input, respectively and $A \in \mathcal{M}(n\times n;R)$ and $\underline{b} \in \mathcal{M}(n\times 1;R)$. Following also Gupta and Hasdorff (1963), we assume that is simple and expressed as

$$A = diag\{\lambda_1, \lambda_2, ..., \lambda_n\}, \tag{2.2.2}$$

where $\lambda_i \neq \lambda_j \neq 0$ for every $i \in \underset{\sim}{n}$ ( ). This assumption can be further re-laxed; see for more details Remark 2.4.1.

This chapter deals with the following *key* question: "*Can we develop an approximation to impulsive behaviour with a respective approximation of the related system and control properties?*"

The answer to this question underpins, the development of a smooth approximation of impulsive trajectories and thus also of the related system and control properties. A number of control problems involving distributional solutions relate to the adjustment of initial conditions with distributional inputs, resulting to distributional state trajectories; these imply changing the given state of a linear system to a desired state in minimum time. The important questions that arise are:

(i) How can we approximate distributions and their derivatives by different families of smooth functions and their derivatives?

(ii) What are the different types of approximation?

(iii) What is the impact of the approximation on the properties of the associated control problem and on the nature of the resulting transition, when smooth functions are used?

It is assumed that the input to the LTI is a *linear* combination of the Dirac δ-function and its first $n-1$ derivatives, i.e.

$$u_o(t) = \sum_{k=0}^{n} a_k \delta^{(k)}(t). \tag{2.2.3}$$

which is a linear combination of Dirac δ-function and its first $n-1$ derivatives, where $\delta^{(k)}$ or $\frac{d^k \delta}{dt^k}$ is the k^{th} derivative of the Dirac δ-function, and a_k for $i \in \underset{\sim}{n}_o$ ($\underset{\sim}{n}_o \triangleq \{0,1,2,...,n-1\}$) are the magnitudes of the delta function and its derivatives. We shall denote the state of the system at time $t = 0^-$ as $\underline{x}(0^-)$ and at time $t \geq 0^+$ as $\underline{x}(0^+)$.

Now, practically speaking, we assume that $\underline{x}(0^-) = [0 \quad 0 \quad \dots \quad 0]^T$ at $t = 0^-$ and $\underline{x}(0^+) = [x_1 \quad x_2 \quad \dots \quad x_n]^T$ at $t \geq 0^+$. Furthermore, we assume that the system is *controllable* and thus we can transfer the state to any desired point of the state space.

Furthermore, we assume that our system is *controllable*, i.e. we can transfer the state to any desired point. Let the state of the system at time $t = 0^-$ be $\underline{x}(0^-) = \underline{0}$ and at time $t = 0^+$, $\underline{x}(0^+)$. The existence of an input that transfers the state of the system (2.2.1) from $\underline{x}(0^-) = \underline{0}$ to $\underline{x}(0^+)$ requires that the vector $\underline{x}(0^+)$ belongs to the controllable subspace of the pair, see Antsaklis and Michel (2009). The necessary and sufficient condition for the state of a system (2.2.1) to be transferred from $\underline{x}(0^-) = \underline{0}$ at time $t = 0^-$ to some $\underline{x}(0^+) \in \{A | \underline{b}\}$ at $t = 0^+$ by the action of a control input of type (2.2.3) is that the resulting trajectory $\underline{x}(t)$ is expressed as $\underline{x}(t) = \sum_{k=0}^{n-1} \beta_k \delta^{(k)}(t)$ where the coef-

ficients β_k for $k \in \underset{\sim}{n}_o$ are chosen to be the components of $\underline{x}\left(0^+\right)$ along the subspace $\{\underline{b}, A\underline{b}, A^2\underline{b}, \ldots, A^{n-1}\underline{b}\}$, respectively according to some projections law.

In the next sub-section, we consider some background results on the approximation of Dirac delta function are presented.

2.3 Approximations of Dirac Delta Function

The approximation of distributions by smooth functions is a problem which has been considered in the literature. In this section, we review the main results in this area and suggest a systematic and rigorous procedure for approximating distributions and their derivatives. If the standard approximating technique of the Dirac δ-function is followed, (see Gupta and Hasdorff, (1963), Gupta, (1966), Zemanian, (1987), Cohen and Kirchner, (1991), Estrada and Kanwal, (2000), Kanwal, (2004) etc) the change of the state in some minimum practical time depends mainly on the accuracy of the approximations that have been generated. The relation between the type of approximation used and the duration of the resulting state-transition is one of the important issues considered in this section.

The Dirac δ-function can be viewed as the limit of the sequence function

$$\delta(t) = \lim_{a \to 0} \delta_a(t), \tag{2.3.1}$$

where $\delta_a(t) \in \mathcal{C}^\infty\left(\mathbb{F}, \mathcal{M}(1\times 1;\mathbb{F})\right)$ is called a *nascent* delta function. This limit is in the sense that

$$\lim_{a \to 0} \int_{-\infty}^{+\infty} \delta_a(t) f(t)\, dt = f(0). \tag{2.3.2}$$

These properties can often be simulated by using a smooth, finite approximation of the Dirac distribution. Such approximations have additional advantages. In fact, approximating the Dirac distribution by a smooth function may actually be a better representation of the solution sought in the particular problem, especially if the width of the approximation function can be coupled to the physics of the problem. Following the ideas of Cohen and Kirschner (1991), a suitable approximating function, which is convenient for computations, should satisfy the following important properties everywhere on the domain under consideration:

1. Its limit with some defining parameter is the Dirac distribution (see eq. (2.3.1)).

2. It is positive, decreases monotonically from a finite maximum at the source point, and tends to zero at the domain extremes.

3. Its derivative exists and is a continuous function.

4. It is symmetric about the source point, for instance 0 (see eq. (2.3.1) and (2.3.2)).

5. It can be represented by a simple Fourier integral (for infinite domains) or Fourier series (for finite domains).

Next, we discuss the appropriate approximation of Dirac function based on the finiteness or infiniteness of the time domain.

2.3.1 Infinite Time-Support Approximations

We first point out that the best nascent delta function depends on the particular application. Some well known (and very useful in applications) nascent delta functions are the Gaussian and Cauchy distributions, the rectangular function, the derivative of the sigmoid (or Fermi-Dirac) function, the Airy function etc; see for instance Gupta (1966), Zemanian (1987), Estrada and Kanwal (2000), Kanwal (2004) et al. and recently the use of a finite difference formula which is converted into an appropriate sequence; see Boykin (2003). A short review of such approximations is given next.

Nascent delta functions very useful in applications are:

● The Cauchy function,

$$\delta_a(t) = \frac{1}{\pi}\frac{a}{a^2+t^2} = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{ikt-|ak|}dk\,,$$

● The rectangular function,

$$\delta_a(t) = \frac{rect(t/a)}{a} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \sin c\left(\frac{ak}{2\pi}\right)e^{ikt}dk\,,$$

where

$$rect(t)=\begin{cases}1, & -1\le t\le 1\\ 0, & |t|>1\end{cases}.$$

● The derivative of the sigmoid (or Fermi-Dirac) function,

$$\delta_a(t)=\partial_t\frac{1}{1+e^{-t/a}}=-\partial_t\frac{1}{1+e^{t/a}},$$

● The Airy function

$$\delta_a(t)=\frac{1}{a}A_i\left(\frac{t}{a}\right).$$

Following Boykin (2003), the finite difference formula may be easily converted into a sequence that approaches a derivative of the Dirac delta function in one dimension. Thus, we obtain

$$\delta_a(t)=\begin{cases}\frac{1}{a}, & -\frac{a}{2}<t<\frac{a}{2}\\ 0, & |t|>\frac{a}{2}\end{cases}, \tag{2.3.3}$$

which approaches $\delta(t)$ as $a\to 0$. An expression for the derivatives of $\delta(t)$ is given by,

$$\frac{d^k}{dx^k}\delta(x)=\lim_{\substack{a\to 0\\ k\to 0}}\left[\left(\frac{1}{h}\right)^k\sum_{j=0}^{k}a_j\delta_a\left(x+b_jh\right)\right], \tag{2.3.4}$$

where $x=t_o-t$ and the a_j are appropriate constants defining the finite differences Boykin (2003), and

$$\frac{d^k}{du^k}\delta(u)|_{t-t_o}=(-1)^k\frac{d^k}{du^k}\delta(u)|_x.$$

The expression (2.3.4) is exactly what we would obtain by making the substitution $f(t) \to \delta_a(t)$ in the following finite difference approximation for the k^{th} derivative of a smooth test function f evaluated at t_o :

$$\frac{d^k}{dt^k} f(t)|_{t=t_o} \approx \left(\frac{1}{h}\right)^k \sum_{j=0}^{k} a_j f\left(t_o + b_j h\right). \tag{2.3.5}$$

Note that a_j and b_j are suitable chosen constants and (2.3.5) becomes exact in the limit $h \to 0$. Furthermore, due to the fact that f is sampled at discrete points, we can write

$$\frac{d^k}{dt^k} f(t)|_{t=t_o} = \lim_{h\to 0}\left\{\left(\frac{1}{h}\right)^k \sum_{j=0}^{k} a_j \int_{-\infty}^{+\infty} \delta\left(t-\left(t_o + b_j h\right)\right) f(t)\, dt\right\} \tag{2.3.6}$$

2.3.2 Finite Time-Support Approximations

Unfortunately, the Gaussian function is not a good approximation of the Dirac distribution on a finite domain, namely that the first derivative (which is important in this paper) can be discontinuous at a special point. Thus, recently, a different approximation has been proposed by Cohen and Kirschner (1991), which satisfies all the properties (**1**) through (**5**). This is the β-function of classical probability theory. This function has the expression

$$\beta_\pi(\theta) = \begin{cases} \dfrac{(\pi+\theta)^{a-1}(\pi-\theta)^{b-1}}{(2\pi)^{2a-1} B(a,b)}, \forall \theta \in \mathcal{J} \\ 0 \qquad\qquad\qquad\qquad , otherwise \end{cases} \tag{2.3.7}$$

where $\mathcal{J}$ is a finite interval and

$$B(a,b) \triangleq \int_{\mathcal{J}} (\pi+\theta)^{a-1}(\pi-\theta)^{b-1}\, d\theta = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)},$$

where $\Gamma(x)$ is the well-known Gamma function. Since, in the next few lines of the present paper, the infinite time domain is used, the interested reader may consult Cohen and Kirschner (1991) for further details.

2.3.3 Why a Sum of Dirac Delta Functions?

However, in our approach, our time domain is *infinite* and the classical Gaussian function, i.e.

$$\delta(t) = \lim_{\sigma \to 0} \frac{1}{\sigma\sqrt{2\pi}} e^{-t^2/2\sigma^2} = \lim_{\sigma \to 0} \frac{1}{\sigma} \phi\left(\frac{t}{\sigma}\right), \tag{2.3.8}$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ is being used.

Consequently, the approximate expression for the controller (2.2.3) is given by

$$u_\sigma(t) = \sum_{k=0}^{n-1} a_k \frac{1}{\sigma^{k+1}} \phi^{(k)}\left(\frac{t}{\sigma}\right), \tag{2.3.9}$$

where $\phi^{(i)}\left(\frac{t}{\sigma}\right) = \left(\frac{d^i}{dt^i}\left(\frac{t}{\sigma}\right)^i\right)\phi\left(\frac{t}{\sigma}\right)$.

Then, we take the limit

$$u_o(t) = \lim_{\sigma \to 0} u_\sigma(t). \tag{2.3.10}$$

Moreover, at the end of this section, we are answering another significant question: "*why a sum of Dirac delta functions*?"

Considering the results of 2nd sub-section and the whole discussion till that part of the 3nd section, generally speaking, we should point out that the input for the linear differential system (2.2.1) should be given by a *single-layer distribution;* see Zemanian (1987), Estrada and Kanwal (2000) and Kanwal (2004). This kind of distributions has a huge importance in many applications.

Lemma 2.3.1 *If $\mathcal{U}$ is a bounded closed set in $\mathbb{F}$ and $\mathcal{Y}$ is a neighbourhood of $\mathcal{U}$, then there exists a function such that $n=1$ on $\mathcal{U}$, $n=0$ outside $\mathcal{Y}$, and $0\le n\le 1$ over $\mathbb{F}$.* □

Definition 2.3.1 *Let $\mathcal{S}$ be a piecewise regular curve in $\mathbb{F}$ and σ is a locally integrable function defined on $\mathcal{S}$. The linear continuous functional $\sigma\delta_S$ on the space $\mathcal{D}$ of infinitely differentiable complex-valued functions on $\mathbb{F}$ with compact support is defined as*

$$\langle \sigma\delta_S, \varphi\rangle = \int_S \varphi(\xi)\sigma(\xi)\delta S$$

$\forall\ \varphi\in\mathcal{D}$ and is called single (or simple) layer on $\mathcal{S}$ with density σ. □

Note that $\sigma\delta_S(x) = \int_S \delta(x-\xi)\sigma(\xi)\delta S$.

Definition 2.3.2 *Let $\mathcal{S}$ be a piecewise regular curve in $\mathbb{F}$ and $\mu\delta_S$. The linear continuous functional $-d/dt(\mu\delta_S)$ on the space $\mathcal{D}$ of infinitely differentiable complex-valued functions on $\mathbb{F}$ with bounded support is defined as*

$$\langle -d/dt(\sigma\delta_S), \varphi\rangle = \int_S \sigma(\xi)\frac{d\varphi(x-\xi)}{dt}\delta S \quad \forall\ \varphi\in\mathcal{D}.$$ □

Consequently, it can be easily shown that every distribution $\sigma\delta_S(x)$ that has compact support is of finite order, see Zemanian (1987) Estrada and Kanwal (2000). Thus, it is deduced that every distribution $\sigma\delta_S(x)$ whose support is the point $x=\tau$ has the form $\sum_{k=0}^{n-1} c_k\delta^{(k)}(t-\tau)$, i.e. a linear independent combination of Dirac δ-function and its first $n-1$ derivatives. Consequently, since we are interested in transfering the state of system (2.2.1) at time $t=0^-$ from the initial point $\underline{x}(0^-)$ and at time $t\ge 0^+$ to achieve $\underline{x}(0^+)$, (2.2.3) is appropriate, when the support point is $\tau=0$.

2.4. Design of Approximate Input Signal

In this section, we will try to answer to the following questions: "*if we wish to achieve state* $\underline{x}\left(0^+\right)$ *at time* $t \geq 0^+$ *what are the necessary coefficients* a_k *for* $k \in \underset{\sim}{n}$ *and what is the optimal choice of* σ *that it takes the state there at time* $t \geq 0^+$ *?*" In this direction, the following known results are significant.

Lemma 2.4.1 *The solution of system (2.2.1) is given by*

$$\underline{x}(t) = e^{At} \int_{-\infty}^{t} e^{-A\tau} \underline{b} u_o(\tau)\, d\tau, \tag{2.4.1}$$

where A *is diagonal and* $u_o(\tau)$ *is given by combining (2.3.9) and (2.3.10).* □

Remark 2.4.1 Recall that for simplicity it is assumed that matrix A is *diagonal, i.e. (2.2.2)* , with distinct eigenvalues; as Gupta and Hasdorff (1963) have also assumed in their work. This reduces the complexity of various mathematical expressions and the number of technicalities involved, without introducing any real loss of generality. The general case can be tackled by defining a $n \times n$ non-singular similarity transformation $Q = [v_1, v_2, \ldots, v_n] \in \mathcal{M}(n \times n; \mathbb{F})$ that takes A into the Jordan canonical form.

In the next lines, we present briefly the more essential part. Further details are omitted, since they are far beyond the scope of the present version of the PhD thesis.

Thus, there exists an invertible matrix $Q \in \mathcal{M}(n \times n; \mathbb{F})$ such as $J = Q^{-1} A Q$, where $J \in \mathcal{M}(n \times n; \mathbb{F})$ is the Jordan canonical form of matrix A. Analytically,

$$J = block\ diag\left\{J_o, J_{q+1}, J_{q+2}, \ldots, J_k\right\}$$

• The block diagonal matrix $J_o = block\ diag\left\{J_1, J_2, \ldots, J_q\right\}$, where

$$J_i = \begin{bmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & 1 & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{bmatrix} \in \mathcal{M}(\tau_i \times \tau_i; \mathbb{F})$$

is also a diagonal matrix with diagonal elements the eigenvalue λ_i, for $i = \underset{\sim}{q}$. Consequently, the dimension of J_o is $s \times s$, $s \triangleq \sum_{i=1}^{q} \tau_i$.

• Also, each block matrix $J_j = block\ diag\left\{J_{j,1}, J_{j,2}, \ldots, J_{j,d_j}\right\}$,

$$J_{j,z_j} = \begin{bmatrix} \lambda_j & 1 & & & \\ & \lambda_j & 1 & & \\ & & \lambda_j & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_j \end{bmatrix} \in \mathcal{M}\left(z_j \times z_j; \mathbb{F}\right)$$

for $j = q+1, q+2, \ldots, k$, and $z_j = \underset{\sim}{d}_j$. □

However, only for the simplicity of calculations, we have already assumed that the matrix A is in diagonal form. Consequently, the solution (2.4.1) is transposed into

$$\underline{x}(t) = \lim_{\sigma \to 0} \left\{ e^{At} \int_{-\infty}^{t} e^{-A\tau} \underline{b} u_\sigma(\tau)\, d\tau \right\} ,$$

or equivalently, $$\underline{x}(t) = e^{At} \lim_{\sigma \to 0} \left[\int_{-\infty}^{t} e^{-A\tau} \underline{b} \sum_{k=0}^{n-1} a_k \frac{1}{\sigma^{k+1}} \phi^{(k)}\left(\frac{\tau}{\sigma}\right) d\tau \right].$$

As $\sigma \to 0$, the energy of the input signal "concentrates" around $\tau = 0$. Hence the zero-time state-transition problem involves setting $t = 0^+$ and selecting the coefficients a_k so that (an arbitrary) $\underline{x}(0^+) \in \mathbb{R}^n$ is reached (recall that controllability of the pair $(A, \underline{b})$ is assumed).

Remark 2.4.2 To reduce the complexity of the solution (due to the large number of terms involved), see the following Lemma and its discussion, we exploit the fact that

$$\phi\left(t/\sigma\right)=\frac{1}{\sqrt{2\pi}}\ e^{-\frac{1}{2}\left(\frac{t}{\sigma}\right)^2},$$

and its derivatives tend to zero very *strongly* with $t/\sigma\to\infty$, see similar statements by Gupta and Hasdorff (1963). Define $t/\sigma\triangleq K\left(t,\sigma\right)$ and assume that *t* is fixed to a positive value, so that $K\left(t,\sigma\right)\to\infty$ as $\sigma\to 0$. Then,

$$\phi\left(t/\sigma\right)\triangleq\phi\left(K\left(t,\sigma\right)\right)\overset{K(t,\sigma)\to\infty}{\to}0,$$

and its derivatives $\phi^{(k)}\left(t/\sigma\right)\triangleq\phi^{(k)}\left(K\left(t,\sigma\right)\right)\overset{K(t,\sigma)\to\infty}{\to}0,\quad k\in\underset{\sim}{n}_o$.

where $\phi^{(0)}\left(t/\sigma\right)\triangleq\phi\left(t/\sigma\right)$. □

A suitable choice of $K\left(t,\sigma\right)$ depends on the choice of the transition time-variable t and the volatility-parameter σ. In practice, t can be fixed, since we can pre-define the duration of the (almost) zero- transition between the initial and final (target) state of the system when solving the (almost-) zero- time state transition problem (e.g., we can select *t* to be of the order of $t\propto 10^{-6}$ seconds, say). This is the approximate version of the exact problem and can be formulated as follows:

For a fixed value of the time parameter $t=t^*$ and a fixed $\varepsilon>0$ determine

$$\sigma^*=\sup\left\{\sigma\in R_+:\left\|\underline{x}\left(t^*\right)-\underline{\hat{x}}\left(t^*\right)\right\|<\varepsilon\right\},\tag{2.4.3}$$

where $\underline{x}\left(t^*\right)$ is the target state and $\underline{\hat{x}}\left(t^*\right)$ is the actual terminal state resulting from the approximation of the input signal, see equation (2.4.1).

This is in the form of a distance-approximation problem. Roughly, for a fixed state-transition time-duration, we seek the "smoothest" input signal for which the error tolerance of the distance between the target and actual terminal state is kept within a pre-defined level ε. Note, that since this distance tends to zero as $\sigma\to 0$ and the only source of error arises from the approximation of the Dirac delta function and its deriva-

tives, an alternative equivalent formulation of the problem is to determine (for a fixed value $t = t^*$),

$$\sigma^* = \sup\left\{\sigma \in R_+ : \left|\phi^{(k)}\left(K\left(t^*, \sigma\right)\right)\right| < \varepsilon_k, k \in \underset{\sim}{n}_o\right\},$$

where the ε_k are suitable positive constants.

The following lemma is required for subsequent developments. The objective is to develop approximation bounds for the terminal state when the impulsive inputs in equation (2.4.1) are substituted by their smooth approximations.

Lemma 2.4.2 *Consider* $u_\sigma(t)$ *defined in equations (3.8). Then*

$$\int_{-\infty}^{t} e^{-\lambda_i \tau} u_\sigma(\tau)\, d\tau = \sum_{k=0}^{n-1} a_k \left\{ e^{-\lambda_i t} \sum_{m=1}^{k} \left(\frac{\lambda_i^{m-k+1}}{\sigma^{k-m+1}} \phi^{(k-m)}\left(\frac{t}{\sigma}\right)\right) + \lambda_i^k e^{\frac{1}{2}\lambda_i^2\sigma^2} \phi^{-1}\left(\frac{t}{\sigma} + \lambda_i \sigma\right)\right\}, \tag{2.4.4}$$

where $\phi^{(0)}(x) \triangleq \phi(x)$, $\phi^{(-1)}(x) \triangleq \int_{-\infty}^{x} \phi(y)\, dy = \sqrt{2}\ \mathrm{erf}^{-1}(2x-1)$, $x \in (0,1)$.

Proof. Substituting the expression (2.3.9) into the integral $\int_{-\infty}^{t} e^{-\lambda_i \tau} u_\sigma(\tau)\, d\tau$, and we obtain

$$\int_{-\infty}^{t} e^{-\lambda_i \tau} \sum_{k=0}^{n-1} \frac{a_k}{\sigma^{k+1}} \phi^{(k)}(\tau/\sigma) d\tau = \sum_{k=0}^{n-1} a_k \int_{-\infty}^{t} e^{-\lambda_i \tau} \frac{\phi^{(k)}(\tau/\sigma)}{\sigma^{k+1}} d\tau.$$

Consider first the term corresponding to $k = 0$,

$$\int_{-\infty}^{t} e^{-\lambda_i \tau} \frac{\phi(\tau/\sigma)}{\sigma} d\tau = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}\lambda_i^2\sigma^2} \int_{-\infty}^{t} e^{-\frac{1}{2}\left(\frac{\tau}{\sigma} + \lambda_i \sigma\right)^2} d\tau = e^{\frac{1}{2}\lambda_i^2\sigma^2} \phi^{-1}\left(\frac{t}{\sigma} + \lambda_i \sigma\right).$$

Consider first the term corresponding to $k = 0$,

$$\int_{-\infty}^{t} e^{-\lambda_i \tau} \frac{\phi(\tau/\sigma)}{\sigma} d\tau = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}\lambda_i^2\sigma^2} \int_{-\infty}^{t} e^{-\frac{1}{2}\left(\frac{\tau}{\sigma}+\lambda_i\sigma\right)^2} d\tau = e^{\frac{1}{2}\lambda_i^2\sigma^2} \phi^{-1}\left(\frac{t}{\sigma}+\lambda_i\sigma\right).$$

Consider next the term corresponding to $k=1$. Integration by parts and using the equation above gives

$$\int_{-\infty}^{t} e^{-\lambda_i \tau} \frac{\phi'(\tau/\sigma)}{\sigma^2} d\tau = e^{-\lambda_i \tau} \frac{1}{\sigma} \phi(\tau/\sigma)\Big|_{-\infty}^{t} + \lambda_i \int_{-\infty}^{t} e^{-\lambda_i \tau} \frac{\phi(\tau/\sigma)}{\sigma} d\tau$$

$$= e^{-\lambda_i t} \frac{1}{\sigma} \phi(t/\sigma) + \lambda_i e^{\frac{1}{2}\lambda_i^2\sigma^2} \phi^{-1}\left(\frac{t}{\sigma}+\lambda_i\sigma\right).$$

Similarly,

$$\int_{-\infty}^{t} e^{-\lambda_i \tau} \frac{\phi''(\tau/\sigma)}{\sigma^3} d\tau = e^{-\lambda_i \tau} \frac{1}{\sigma^2} \phi'(\tau/\sigma)\Big|_{-\infty}^{t} + \lambda_i \int_{-\infty}^{t} e^{-\lambda_i \tau} \frac{\phi'(\tau/\sigma)}{\sigma^2} d\tau$$

$$= e^{-\lambda_i t}\left[\lambda_i \frac{1}{\sigma} \phi(t/\sigma) + \frac{1}{\sigma^2} \phi'(t/\sigma)\right] + \lambda_i^2 e^{\frac{1}{2}\lambda_i^2\sigma^2} \phi^{-1}\left(\frac{t}{\sigma}+\lambda_i\sigma\right).$$

A recursive application of this procedure gives

$$\int_{-\infty}^{t} e^{-\lambda_i \tau} \frac{\phi^{(k)}(\tau/\sigma)}{\sigma^{k+1}} d\tau = e^{-\lambda_i t} \sum_{m=1}^{k} \lambda_i^{m-k+1} \frac{1}{\sigma^{k-m+1}} \phi^{(k-m)}\left(\frac{t}{\sigma}\right) + \lambda_i^k e^{\frac{1}{2}\lambda_i^2\sigma^2} \phi^{-1}\left(\frac{t}{\sigma}+\lambda_i\sigma\right),$$

from which the result follows. □

Choose $0^+/\sigma \triangleq K(0^+,\sigma)$ sufficiently large so that $\phi^{(k)}(0^+/\sigma) \triangleq \phi^{(k)}(K(0^+,\sigma)) \approx 0$, $k \in \underset{\sim}{n}_o$. Then the following approximation is valid

$$\int_{-\infty}^{0^+} e^{-\lambda_i \tau} \frac{\phi^{(k)}(\tau/\sigma)}{\sigma^{k+1}} d\tau \approx \lambda_i^k e^{\frac{1}{2}\lambda_i^2\sigma^2} \phi^{-1}\left(K(0^+,\sigma)+\lambda_i\sigma\right).$$

Combining expressions (2.4.2) and (2.4.4) then gives

$$\underline{x}_i\left(K(0^+,\sigma)\sigma\right) \approx b_i e^{\lambda_i K(0^+,\sigma)\sigma + \frac{1}{2}\lambda_i^2\sigma^2} \phi^{-1}\left(K(0^+,\sigma)+\lambda_i\sigma\right) \sum_{k=0}^{n-1} a_k \lambda_i^k, \qquad (2.4.5)$$

for $i=1,2,\ldots,n$.

The *approximate almost zero-time state-transfer problem* can now be defined as follows: Suppose that parameters $(0^+,\sigma)$ have been chosen so that $\phi^{(k)}\left(0^+/\sigma\right)\triangleq \phi^{(k)}\left(K\left(0^+,\sigma\right)\right)\approx 0,\ k\in \underset{\sim}{n}_o$. Then, given $\underline{x}\left(0^+\right)\in\mathbb{R}^n$ determine real scalars $a_k,\ k\in \underset{\sim}{n}_o$ such that (2.4.5) are satisfied with equality for all $i\in\{1,2,\ldots,n\}$.

Note that the impulsive response is recovered as $\sigma\to 0$ in which case the approximation in the above equation becomes exact; in this case we also have that $\underline{x}_i(0^+)\to\underline{\hat{x}}_i(0^+),\ \phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)\to 1$, and

$$\underline{\hat{x}}_i\left(0^+\right)=b_i e^{\lambda_i 0^+}\sum_{k=0}^{n-1}a_k\lambda_i^k,\quad i=1,2,\ldots,n$$

so that

$$\underline{x}_i\left(0^+\right)\approx e^{\frac{1}{2}\lambda_i^2\sigma^2}\phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)\underline{\hat{x}}_i\left(0^+\right),\quad i=1,2,\ldots,n.$$

Theorem 2.4.1 now follows.

Theorem 2.4.1 *Let* $A=diag\{\lambda_1,\lambda_2,\ldots\lambda_n\}$ *with* $\lambda_i\neq\lambda_j$ *for* $i\neq j$, $\underline{b}=\begin{pmatrix}b_1 & b_2 & \ldots & b_n\end{pmatrix}^T$ *and assume that the pair* $(A,\underline{b})$ *is controllable. Let also* $\hat{B}=diag\{1/b_1,1/b_2,\ldots,1/b_n\}$ *and denote by* $V\triangleq V\left(\lambda_1,\lambda_2,...,\lambda_n\right)$ *the Vandermonde matrix*

$$V\triangleq V\left(\lambda_1,\lambda_2,...,\lambda_n\right)=\begin{bmatrix}1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{n-1}\\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{n-1}\\ & & \vdots & & \\ 1 & \lambda_n & \lambda_n^2 & \ldots & \lambda_n^{n-1}\end{bmatrix}.$$

Then the coefficient vector $\underline{a}=\begin{bmatrix}a_o & a_1 & \cdots & a_{n-1}\end{bmatrix}^T$ *of the input signal defined in (2.3.9) which solves the almost zero-time state-transfer problem is given by*

$$\underline{a} = V^{-1} e^{-A \cdot 0^+} \hat{B}^{-1} \underline{\hat{x}}\left(0^+\right), \tag{2.4.6}$$

where

$$\underline{\hat{x}}_i\left(0^+\right) \triangleq \frac{\underline{x}_i\left(K\left(0^+,\sigma\right)\sigma\right)}{e^{\frac{1}{2}\lambda_i^2\sigma^2}\phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)}\ ,\ i \in \underset{\sim}{n}. \tag{2.4.7}$$

Proof. Expression (2.4.4) can be re-written as

$$\hat{x}_i(0^+) = b_i e^{\lambda_i 0^+} \sum_{k=0}^{n-1} a_k \lambda^k \cong \frac{x_i(0^+)}{e^{\lambda_i^2\sigma^2/2}\phi^{-1}(0^+ + \lambda_i\sigma)} \text{ for } i \in \underset{\sim}{n}.$$

Thus we can write $\underline{\hat{x}}\left(0^+\right) = \hat{B} e^{A \cdot 0^+} V \underline{a}$ or equivalently (2.4.6). Note that the indicated inverses V^{-1} and $\hat{B}^{-1}$ exist due the assumption that the eigenvalues of *A* are distinct, and the assumed controllability of $(A, \underline{b})$, respectively. □

Ideally the parameters $t^* = 0^+$ and σ should be chosen so that the distance

$$\left\| \underline{x}\left(t^*\right) - \underline{\hat{x}}\left(t^*\right) \right\|_2 = \sqrt{\sum_{i=1}^{n} \left[\underline{x}_i\left(K\left(t,\sigma\right)\sigma\right) - \underline{\hat{x}}_i\left(K\left(t,\sigma\right)\sigma\right) \right]^2}$$

is "small". Clearly the distance is zero provided that $K\left(t,\sigma\right)$ is selected so that

$$\phi^{-1}\left(K\left(t,\sigma\right)+\lambda_i\sigma\right) - e^{-\frac{1}{2}\lambda_i^2\sigma^2} = 0 \tag{2.4.8}$$

for all *i* which requires $\sigma \to 0$, in which case (2.4.8) implies that

$$\lim_{\sigma\to 0} \phi^{-1}\left(K\left(t,\sigma\right)\right) = 1 \Leftrightarrow \lim_{\sigma\to 0} \int_{-\infty}^{K(t,\sigma)} \phi\left(x\right) dx = 1 \Leftrightarrow K(t,\sigma) \to \infty. \tag{2.4.9}$$

In probability theory and statistics, the normal or Gaussian function $\phi\left(x\right)$ is widely used. The graph of $\phi\left(x\right)$ is bell-shaped and is known as the Gaussian function or bell curve. Actually, in this case we are interested in

$$\int_{-\infty}^{K(t,\sigma)} \phi(x)dx,$$

which is the *cumulative distribution function* (cdf) of a random variable $X \sim N(0,1)$ evaluated at the upper limit of the integral $K(t,\sigma)$, denoting the probability that $X \le K(t,\sigma)$. In practice, if $|\lambda_i \sigma| << 1$ for all *i*, we can assume that equation (2.4.8) is approximately satisfied if $K_0 \triangleq K(t,\sigma) \ge 3.9$ (in which case $\phi^{-1}(K_0) > 1 - 10^{-4}$, see relevant table value for the Standard Normal Distribution which represents area to the left of Z score). Thus, a reasonable choice for the volatility parameter is $\sigma^* = K_0^{-1} t^* \approx 0.256 t^*$.

The results of the section are summarized in the following algorithm.

Algorithm TIAZT (**T**ransfer **I**n **A**lmost **Z**ero **T**ime)

1st Step: Define the terminal (target) state of the transition $\underline{x}(0^+)$.

2nd Step: Using the required transition time $t^*(\equiv 0^+)$ define the optimal volatility parameter $\sigma^* = 0.256 t^*$.

3rd Step: Finally, the coefficients of the input signal $\underline{a} = [a_o \quad a_1 \quad \cdots \quad a_{n-1}]^T$ defined in equation (2.3.9) are obtained by (2.4.6), i.e. $\underline{a} = V^{-1} e^{-A \cdot 0^+} \hat{B}^{-1} \underline{\hat{x}}(0^+)$ where all variables are defined in Theorem 2.4.1.

Remark 2.4.4 From the control viewpoint it is important to choose appropriate time duration for the state transition. This ultimately depends on the type of application, e.g. due to control signal magnitude or "slew-rate" limitations. It is clear from the imposed proportionality $\sigma^* = K_0^{-1} t^*$ that increasing the duration of the state-transition results is "smoother" input signals, which is often desirable. For example, if the system operates

in a feedback loop (in which case the input signal is generated by a feedback controller), highly discontinuous signals typically correspond to system overdesign (e.g. excessive closed-loop bandwidth) and may have detrimental effects on the stability and performance characteristics, e.g. in terms of reduced robust stability margins and sensor noise amplification. □

Example 2.4.1 (See Gupta, 1966) Consider the system

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u_o(t),$$

where $\underline{x}(t)$ and $u_o(t)$ are the state and the input signals, respectively. Suppose we wish to transfer the state of the system from $\underline{x}(0) = (0 \quad 0)^T$ to $\underline{x}(0^+) = (3 \quad 4)^T$ at time $0^+ = 1\mu s$ (1 microsecond). Application of the TIAZT algorithm gives

1st Step: Here the desired state is $\underline{x}(0^+) = (3 \quad 4)^T$.

2nd Step: The transition duration has been pre-determined as $0^+ = 10^{-6}$ s, so the optimal volatility parameter is $\sigma^* = 2.56 \cdot 10^{-7}$ (taking $K_0 = 3.9$).

3rd Step: Here, $\hat{\underline{x}}_1(10^{-6}) \approx \underline{x}_1(10^{-6}) = 3$ and $\hat{\underline{x}}_2(10^{-6}) \approx \underline{x}_2(10^{-6}) = 4$. The inverse of the Vandermonde matrix is:

$$V^{-1} = V^{-1}(-2,-3) = \begin{bmatrix} 1 & -2 \\ 1 & -3 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -2 \\ 1 & -1 \end{bmatrix}.$$

Thus, the coefficient vector $\underline{a} = [a_0 \quad a_1]^T$ is calculated as:

$$\underline{a} = \begin{bmatrix} a_o \\ a_1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \exp(2\times10^{-6}) & 0 \\ 0 & \exp(3\times10^{-6}) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \approx \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

2.5 Distance Problems

2.5.1. Distance from the origin in state-space

In this section, we define the distance from the origin corresponding to a state transition of the system (2.2.1) from the zero (or ground) state, $\underline{x}\left(0^-\right)=\left[0 \quad 0 \quad \cdots \quad 0\right]^T$. Using the Euclidean norm this is defined as

$$r^2 \triangleq \left\|\underline{x}\left(0^+\right)-\underline{x}\left(0^-\right)\right\|^2 = \underline{x}^T\left(0^+\right)\underline{x}\left(0^+\right)=\sum_{i=1}^{n} x_i^2\left(0^+\right), \tag{2.5.1}$$

(see Fig 2.5.1). The time interval of the transition has been defined in previous sections as 0^+ (t^*) and the target state is $\underline{\hat{x}}\left(0^+\right)$.

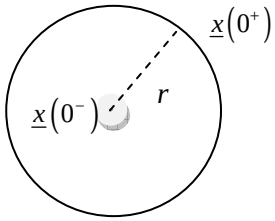

Fig. 2.5.1: 2-ball with centre $\underline{x}\left(0^-\right)$ and radius $\hat{r}$

However, if the Dirac delta function and its derivatives are replaced by smooth signals (Gaussian function and its derivatives), this target state will not be reached exactly, in general. The distance in terms of the target state $\underline{\hat{x}}\left(0^+\right)$ is defined as

$$\hat{r}^2 \triangleq \sum_{i=1}^{n} \hat{x}_i^2\left(0^+\right) = \sum_{i=1}^{n} \frac{x_i^2\left(K\left(0^+,\sigma\right)\sigma\right)}{e^{\lambda_i^2\sigma^2}\left[\phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)\right]^2},$$

where (2.4.7) has been used. Note that fixing $K\left(t,\sigma\right)$ and taking $\sigma\to 0$, we get $\hat{r}\to r$.

Example 2.5.1 Consider the system:

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u_o(t),$$

where $\underline{x}(t) \in \mathcal{C}^{\infty}\left(R, \mathcal{M}(2\times 1; R)\right)$ and $u_o(t)$ are the state vector and the input, respectively. Let $\underline{x}(0^-) = 0$ and $\underline{x}(0^+) = [3 \quad 4]^T$. Then

$$\hat{r} = \left\| \underline{\hat{x}}(0^+) - \underline{x}(0^-) \right\| = \sqrt{\sum_{i=1}^{2} \frac{x_i^2\left(K(0^+,\sigma)\sigma\right)}{\left[\underbrace{e^{\lambda_i K(0^+,\sigma)\sigma + \frac{1}{2}\lambda_i^2\sigma^2}\phi^{-1}\left(K(0^+,\sigma)+\lambda_i\sigma\right)}_{\beta_i}\right]^2}} = \sqrt{\frac{9}{\beta_1^2} + \frac{16}{\beta_2^2}}.$$

As $\beta_1, \beta_2 \to 1$, $\hat{r} \to r = 5$.

2.5.2 Maximum distance from the origin with constrained input

Here we assume that the system (2.2.1) starts from the zero state at time $t = 0^-$ and consider the problem of maximizing the distance to the terminal state in an (almost) zero-time state transition. This problem of course makes sense if the input signal is constrained in some sense, see Gupta (1964). Thus, here we also impose constraints on the coefficient vector of the input signal $\underline{a} = [a_0 \quad a_1 \quad \cdots \quad a_{n-1}]^T$ in terms of the Euclidian and the infinity norms (alternatively, you can consider bounded energy, instead of bounded gain). Again, our approach reformulates, extends and supports the preliminary ideas proposed by Gupta (1964), as we can prove that both problems are tractable and can be solved via an SVD and the solution of a quadratic programming problem with box constraints, respectively. Especially, the connection of our problem with the literature of quadratic programming is very fruitful for further future consideration.

Lemma 2.5.1 *Let* $\lambda_i \neq 0$, $i = 1, 2, \ldots, n$. *Then* $\sum_{i=1}^{n} |\lambda_i|^{p-1} \leq \max\left\{n, \sum_{i=1}^{n} |\lambda_i|^{n-1}\right\}$ *for all* $p = 1, 2, \ldots, n$.

Proof. Define function $f(x) = \sum_{i=1}^{n} |\lambda_i|^{x-1}$ which can be written as $f(x) = \sum_{i=1}^{n} e^{m_i(x-1)}$ by setting $m_i = \ln|\lambda_i|$. Since $f''(x) = \sum_{i=1}^{n} m_i^2 e^{m_i(x-1)} > 0$ for all $x \in \mathbb{R}$, function is convex

for all $x \in \mathbb{R}$ and specifically in the interval $1 \le x \le n$. Thus $f(x)$ attains its maximum at an edge of the interval $1 \le x \le n$, i.e.

$$\sum_{i=1}^{n} |\lambda_i|^{p-1} \le \max_{1 \le x \le n} f(x) = \max\{ f(1), f(n)\} = \max\left\{n, \sum_{i=1}^{n} |\lambda_i|^{n-1}\right\},$$

for every $p = 1, 2, \ldots, n$ as required. □

Under this framework, the following Theorem can be characterized as a useful complementary result of Theorem 2.4.1, where an interesting upper bound is given for the maximum distance of the zero-time state-transition problem when we have imposed constraints on the coefficient vector of the input signal $\underline{a}$.

Theorem 2.5.1 *Let* $A = diag\{\lambda_1, \lambda_2, \ldots \lambda_n\}$, $\underline{b} = \begin{pmatrix} b_1 & b_2 & \ldots & b_n \end{pmatrix}^T$ *and assume that the pair* $(A, \underline{b})$ *is controllable. Define* $\hat{B} = diag\{1/b_1, 1/b_2, \ldots, 1/b_n\}$ *and denote by* $V \triangleq V(\lambda_1, \lambda_2, ..., \lambda_n)$ *the Vandermonde matrix*

$$V = V(\lambda_1, \lambda_2, ..., \lambda_n) = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{n-1} \\ & & \vdots & & \\ 1 & \lambda_n & \lambda_n^2 & \ldots & \lambda_n^{n-1} \end{bmatrix}.$$

Let $\underline{a} = \begin{bmatrix} a_o & a_1 & \cdots & a_{n-1} \end{bmatrix}^T$ *be the coefficient vector of the input signal* $u_o(t) = \sum_{i=0}^{n-1} a_i \delta^{(i)}(t)$ *defined in (2.3.9). Then, if* $\underline{\hat{x}}(0^+)$ *denotes the terminal state of the zero-time state-transition problem with* $\underline{\hat{x}}(0^-) = \underline{0}$,

$$\max_{\|a\|=1} \left\| \underline{x}(t^+) \right\| = \left\| \hat{B} e^{A \cdot 0^+} V \right\| \le \frac{t^+ \rho(A) \sqrt{n}}{\min_{i \in n} |b_i|} \max\left\{n, \sum_{i=1}^{n} |\lambda_i|^{n-1}\right\}, \qquad (2.5.2)$$

where the indicated matrix norm denotes the largest singular value (spectral norm) and $\rho(A)$ *denotes the spectral radius of A.*

Proof. In the notation of Theorem 2.4.1 the terminal state of the transition is $\underline{\hat{x}}\left(0^+\right) = \hat{B}e^{A\cdot 0^+}V\underline{a}$. Thus, $\max_{\|a\|=1}\left\|\underline{\hat{x}}\left(0^+\right)\right\| = \left\|\hat{B}e^{A\cdot 0^+}V\right\|$, while the maximizing coefficient vector $\underline{a}$ is the (normalized) singular vector of $\hat{B}e^{A\cdot 0^+}V$ corresponding to the largest singular value. (If the largest singular value is repeated we can choose any linear combination of unit length of the singular vectors corresponding to the repeated largest singular value). Note also that

$$\left\|\hat{B}e^{A\cdot 0^+}V\right\| \le \left\|\hat{B}\right\|\left\|e^{A\cdot 0^+}\right\|\|V\| = \frac{t^* \max_{i\in n}|\lambda_i(A)|}{\min_{i\in n}|b_i|}\|V\| = \frac{t^*\rho(A)}{\min_{i\in n}|b_i|}\|V\|. \tag{2.5.3}$$

Now,

$$\|V\| \le \sqrt{n}\|V\|_1 = \sqrt{n}\left\|V^T\right\|_\infty = \sqrt{n}\max_{p=1,2,\ldots,n}\sum_{i=1}^n|\lambda_i|^{p-1} = \sqrt{n}\max\left\{n, \sum_{i=1}^n|\lambda_i|^{n-1}\right\}, \tag{2.5.4}$$

see Lemma 2.5.1 and Gupta and Hasdorff (1963), where $\|\cdot\|_1$ and $\|\cdot\|_\infty$ denote the induced 1 and ∞-matrix norms, respectively. Equation (2.5.2) follows by combining (2.5.3) and (2.5.4). □

Remark 2.5.1 Consider the almost zero-time state transition problem in which $K\left(t^+,\sigma\right) = t^+/\sigma$ has been fixed and σ has been chosen sufficiently small so that $|\lambda_i\sigma| << 1$ for all *I* and approximation Gautshi (1975) is valid.

Then we have

$$\underline{x}\left(0^+\right) = \Gamma Be^{A0^+}V\underline{a},$$

where $\Gamma = diag\left\{\lambda_i^2\sigma^2\phi^{-1}\left(K(0^+,\sigma)+\lambda_i\sigma\right)/2\right\}$.

It follows that in this case

$$\max_{\|a\|=1}\left\|\underline{x}(0^+)\right\| = \left\|\Gamma Be^{A0^+}V\right\| \le \psi(n)\max_{i\in\underline{n}}\left\{|b_i|\lambda_i^2\, e^{\lambda_i 0^+}\,\phi^{-1}\left(K(0^+,\sigma)+\lambda_i\sigma\right)\right\},$$

where

$$\psi(n) = \frac{\sqrt{n}\sigma^2}{2} \max\left\{n, \sum_{i=1}^{n} |\lambda_i|^{n-1}\right\},$$

while the maximizing coefficient vector $\underline{a}$ is the (normalized) singular vector of $\Gamma B e^{A0^+} V$ corresponding to the largest singular value.

Next, we impose magnitude constraints on the coefficients defining the distributional input signal. Again we assume that $\underline{\hat{x}}\left(0^-\right) = \underline{x}(0^-) = \underline{0}$ and seek to maximize $\left\|\underline{\hat{x}}\left(0^+\right)\right\|$ using the impulsive input $u_0(t)$ in equation (2.3.10) (or $\left\|\underline{x}\left(0^+\right)\right\|$ using its smooth approximation $u_\sigma(t)$ in (2.3.9)) subject to the constraint:

$$|a_i| \le c_i,\ c_i > 0, \text{ for } i \in \underset{\sim}{n} \tag{2.5.5}$$

(see also Gupta and Hasdorff (1963)). Geometrically, we seek constants a_i for $i \in \underset{\sim}{n}$ in the ranges defined by (2.5.5) such as the radius $\hat{r}$ depicted in fig. 2.5.2 is maximized, (starting from $\underline{\hat{x}}\left(0^-\right) = \underline{0}$) where

$$\hat{r}^2 = \left\|\underline{\hat{x}}\left(0^+\right)\right\|^2 = \sum_{i=1}^{n} \hat{x}_i^2\left(0^+\right) = \sum_{i=1}^{n} b_i^2 e^{2\lambda_i 0^+} \sum_{j=1}^{n} \sum_{s=1}^{n} \lambda_i^{j+s-2} a_{j-1} a_{s-1} \tag{2.5.6}$$

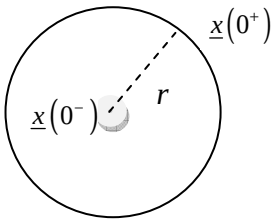

Fig. 2.5.2: *n*-ball with centre $\underline{x}\left(0^-\right)$ and radius r

Again, if the smooth approximation signal $u_\sigma(t)$ is applied, equation (2.4.6) should be used; substitution into equation (2.5.6) shows that in this case we seek to maximize:

$$r^2 = \sum_{i=1}^{n} x_i^2\left(0^+\right) = \sum_{i=1}^{n} \hat{x}_i^2\left(0^+\right)\left[e^{\frac{1}{2}\lambda_i^2\sigma^2} \phi^{-1}\left(K\left(0^+, \sigma\right) + \lambda_i \sigma\right)\right]^2.$$

Next note that equation (2.4.5) gives:

$$\hat{x}_i\left(0^+\right) = b_i e^{\lambda_i 0+} \sum_{j=1}^{n} \lambda_i^{j-1} a_{j-1},$$

and hence

$$\hat{x}_i^2\left(0^+\right) = b_i^2 e^{2\lambda_i 0+} \sum_{j=1}^{n} \sum_{s=1}^{n} \lambda_i^{j+s-2} a_{j-1} a_{s-1}, \quad i \in \underset{\sim}{n} \tag{2.5.7}$$

Substituting, (2.5.7) into (2.5.6), gives

$$r^2 = \left\| \underline{x}\left(0^+\right) \right\|^2 = \sum_{i=1}^{n} \left[b_i e^{\lambda_i 0^+ + \frac{1}{2}\lambda_i^2 \sigma^2} \phi^{-1}\left(K\left(0^+, \sigma\right) + \lambda_i \sigma\right) \right]^2 \sum_{j=1}^{n} \sum_{s=1}^{n} \lambda_i^{j+s-2} a_{j-1} a_{s-1}.$$

(2.5.8)

Define the symmetric matrix

$$Q(\sigma) = V^T D^2(\sigma) V,\ D = diag\left(b_i e^{\lambda_i 0^+ + \frac{1}{2}\lambda_i^2 \sigma^2} \phi^{-1}\left(K\left(0^+, \sigma\right) + \lambda_i \sigma\right) \right).$$

Note that due to the assumed controllability of $(A, \underline{b})$ (which implies that $b_i \neq 0$, $i \in \underset{\sim}{n}$) and the assumption that the eigenvalues of A are distinct (which implies that $\det(V) \neq 0$), we have that $Q(\sigma) = Q^T(\sigma) > 0$. The two distance maximization problems now have the form

$$\max\ r^2 = \left\| \underline{x}\left(0^+\right) \right\|^2 = a^T Q(\sigma) a \quad s.t.\ -c_i \leq a_i \leq c_i,\ \ i \in \underset{\sim}{n}$$

and

$$\max\ \hat{r}^2 = \left\| \underline{\hat{x}}\left(0^+\right) \right\|^2 = a^T Q(0) a \quad s.t.\ -c_i \leq a_i \leq c_i,\ \ i \in \underset{\sim}{n},$$

which are *Quadratic Programming* optimization problems with "box" constraints. Since the cost function ($f(a) = a^T Q(\sigma) a$) which is maximized is convex, the constrained maximum is achieved in a vertex of a hyper-cube $|a_i| = c_i,\ i \in \underline{n}$.

Thus, as Gupta and Hasdorff (1963) have mentioned, we can also prove

$$\left\{(-\lambda_i)^{j+s-2}\operatorname{sgn} a_{j-1}\operatorname{sgn} a_{s-1}\right\}>0 \text{ for all } j \text{ and } k.$$

This can be easily derived if we assume that

$$\operatorname{sgn} a_{j-1}=(-1)^{j-1} \text{ and } \operatorname{sgn} a_{s-1}=(-1)^{s-1},$$

so we obtain

$$\operatorname{sgn} a_{j-1}\operatorname{sgn} a_{s-1}=(-1)^{j-1}(-1)^{s-1}=(-1)^{j+s-2}.$$

So, the maximum distance is given by

$$r \triangleq \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{s=1}^{n}|\lambda_i|^{j+s-2}c_{j-1}c_{s-1}\left[e^{\lambda_i K(0^+,\sigma)\sigma+\frac{1}{2}\lambda_i^2\sigma^2}\phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)\right]^2. \quad (2.5.9)$$

Finally, again if we assume that $t^*=K\left(t^*,\sigma^*\right)\sigma^*\to 0$, and $K\left(t^*,\sigma^*\right)$ to be equal or greater to 3.90, we obtain

$$r \triangleq \left\|\underline{x}\left(t^*\right)-\underline{x}\left(0^-\right)\right\|^2=\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{s=1}^{n}|\lambda_i|^{j+s-2}c_{j-1}c_{s-1}. \quad (2.5.10)$$

The following numerical example illustrates some of the results of this section.

Example 2.5.2 Consider the (almost) zero state transition problem for the system defined in example 2.5.1 with $\underline{x}\left(0^-\right)=\underline{0}$. Suppose that the following constraints are imposed on the coefficients of the input signal

$$|a_0|\le c_0=1, \text{ and } |a_1|\le c_1=2.$$

Subject to these constraints, the maximum distance from the zero state is:

$$\left\|\underline{x}\left(0^+\right)-\underline{x}\left(0^-\right)\right\|^2=\sum_{i=1}^{2}x_i^2\left(0^+\right)$$

$$=\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{s=1}^{2}|\lambda_i|^{j+s-2}c_{j-1}c_{s-1}\left[e^{\lambda_i K(0^+,\sigma)\sigma+\frac{1}{2}\lambda_i^2\sigma^2}\phi^{-1}\left(K\left(0^+,\sigma\right)+\lambda_i\sigma\right)\right]^2.$$

$$= \sum_{j=1}^{2}\sum_{s=1}^{2} |\lambda_1|^{j+s-2} c_{j-1}c_{s-1} \left[\underbrace{e^{\lambda_1 K(0^+,\sigma)\sigma + \frac{1}{2}\lambda_1^2\sigma^2} \phi^{-1}\left(K(0^+,\sigma)+\lambda_1\sigma\right)}_{\beta_1} \right]^2$$

$$+ \sum_{j=1}^{2}\sum_{s=1}^{2} |\lambda_2|^{j+s-2} c_{j-1}c_{s-1} \left[\underbrace{e^{\lambda_2 K(0^+,\sigma)\sigma + \frac{1}{2}\lambda_2^2\sigma^2} \phi^{-1}\left(K(0^+,\sigma)+\lambda_2\sigma\right)}_{\beta_2} \right]^2 .$$

$$= \sum_{s=1}^{2} |\lambda_1|^{s-1} c_0 c_{s-1}\beta_1^2 + \sum_{s=1}^{2} |\lambda_1|^{s} c_1 c_{s-1}\beta_1^2 + \sum_{s=1}^{2} |\lambda_2|^{s-1} c_0 c_{s-1}\beta_2^2 + \sum_{s=1}^{2} |\lambda_2|^{s} c_1 c_{s-1}\beta_2^2$$

$$= 2\left(\beta_1^2+\beta_2^2\right)c_0c_0 + 2\left(|\lambda_1|\beta_1^2 + |\lambda_2|\beta_2^2\right)c_0c_1 + \left(|\lambda_1|^2\beta_1^2 + |\lambda_2|^2\beta_1^2\right)c_1c_1$$

In this example, $c_0 = 1,\ c_1 = 2$ and $|\lambda_1| = 2,\ |\lambda_2| = 3$.

So, the maximum radius is given by

$$r \triangleq \left\| \underline{x}(0^+) - \underline{x}(0^-) \right\| = \sqrt{4\left(\beta_1^2+\beta_2^2\right) + 4\left(2\beta_1^2+3\beta_2^2\right) + 2\left(4\beta_1^2+9\beta_1^2\right)} = \sqrt{20\beta_1^2+34\beta_2^2} .$$

Now, for the case that $t^* = K\left(t^*,\sigma^*\right)\sigma^* \to 0$, we have $\beta_1^2, \beta_2^2 \to 1$ and

$$r \triangleq \left\| \underline{x}(t^*) - \underline{x}(0^-) \right\| = \sqrt{4+4(2+3)+2(4+9)} = \sqrt{54} \approx 7.35.$$

2.6. Conclusions – Further Research

In this chapter, a novel methodology has been proposed for approximating the distributional trajectory that transfers the state of a LTI differential system in (almost) zero time by using an impulsive input. It has been shown that no loss of generality is introduced if the impulsive input signal is chosen as a linear combination of the Dirac δ-function and its first $n-1$ derivatives, where n is the order of the system. Approximations of the impulsive input signal were considered using the Gaussian (Normal) function, and the resulting response of the system was analysed. The work has addressed the following three distinct problems:

(i) We have determined the (unique) impulsive input signal (and its smooth approximation) which transfers the state of the system from the origin to an arbitrary point in state space in zero (almost-zero) time, subject to appropriate controllability assumptions. To simplify our presentation, the simplest set of assumptions has been selected (full system controllability, single control input, distinct set of eigenvalues in the system matrix); however, extension to the general case is straightforward at the expense of possible loss of uniqueness and considerable additional complexity in the resulting mathematical expressions.

(ii) A Euclidean metric has been defined to quantify the approximation error in the state-trajectories of the system resulting from substituting impulsive input signals by smooth signals. The optimal choice of two parameters (time and volatility) characterising the family of all smooth approximating functions has been obtained, along with an interesting probabilistic interpretation.

(iii) The solution of two state-space maximum-distance problems in the context of (almost) zero-time state-transition has been presented for the case of system (2.2.1). These correspond to two dif-ferent types of constraints on the coefficients of the impulsive input signal and its smooth approximation, involving the Euclidian and infinity norms of the vector of coefficients. Both problems are tractable and can be solved via an SVD and the solution of a quadratic programming problem with box constraints, respectively.

Future work will attempt to: (i) extend the results of this paper to more general classes of systems (e.g. descriptor, singular), (ii) investigate the numerical properties of simulating impulsive trajectories and their smooth approximation, and (iii) develop alternative energy-based approximation techniques of impulsive behaviour especially in the context of large-scale systems and model reduction.

Chapter 3

Generalized Inverses of Structural Matrices (Vandermonde and a Special Matrix) Appearing in Control

3.1 Introduction

The square and rectangular Vandermonde matrices have been appeared several times in many distinct areas of numerical analysis, in control and system theory; see for instance: Wertz (1965), Klinger (1967), Björck and Pereyra (1970), Tang and Golub (1981), Martinez and Peña (1998 a, b), Eisinberg, Franzé and Salerno (2001), Eisinberg and Fedele (2006), Karageorgos, Pantelous and Kalogeropoulos (2009) and the 2nd Chapter of this PhD thesis. Moreover, different kind of approaches and algorithms for the representation of the (generalized) inverse of a square Vandermonde matrix have been proposed, see Wertz (1965), Klinger (1967), Kaufman (1969), Bork and Pereira (1970), Tang and Glob (1981), Martinez and Pena (1998a, b), Orcus and Phillips (2000), Deisenberg and Fidel (2006), and of a rectangular Vandermonde matrix, see Deisenberg, Franz and Salerno (2001), using different factorization methods and numerical functions.

Recently, in the literature of control and system theory, see characteristically Karageorgos, Pantelous and Kalogeropoulos (2009), the transfer of the initial state of an

open loop, linear higher-order descriptor (regular) differential system in (almost) zero-time has been fully investigated, i.e.

$$F\underline{x}^{(r)}(t) = G\underline{x}(t) + \underline{b}u(t),$$

with known initial conditions

$$\underline{x}(t_o),\ \underline{x}'(t_o),....,\ \underline{x}^{(r-1)}(t_o),$$

where $F, G \in \mathcal{M}(n\times n;\mathbb{F})$, and $\underline{b} \in \mathcal{M}(n\times 1;\mathbb{F})$ (i.e. $\mathcal{M}$ is the algebra of $n\times m$ matrices with elements in the field $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$) with $\det F = 0$ (0 is the zero element of $\mathcal{M}(n=1,\mathbb{F})$), $\underline{x}(t)\in \mathcal{C}^{\infty}(\mathbb{F}, \mathcal{M}(n\times 1;\mathbb{F}))$ and $u(t)\in \mathcal{D}'_{n-1}$ (where $\mathcal{D}'_{n-1}$ is the *space* of Dirac distribution having derivatives up to an order $n-1$). For the sake of simplicity, we set in the sequel $\mathcal{M}_n \triangleq \mathcal{M}(n\times n;\mathbb{F})$ and $\mathcal{M}_{n,m} \triangleq \mathcal{M}(n\times m;\mathbb{F})$.

In order to solve this problem, the appropriate input vector has to be made up as a linear combination of the Dirac δ-function and its derivatives, for more details see Karageorgos, Pantelous and Kalogeropoulos (2009) and references therein, i.e.

$$u_o(t) = \sum_{k=1}^{n-1} a_k \delta^{(k)}(t), \tag{3.1.1}$$

where $\delta^{(k)}(t)$ or $\dfrac{d^k\delta(t)}{dt^k}$ is the k^{th}-derivative of the Dirac δ-function, and a_i for $i = 0,1,\ldots,n-1$ are the magnitudes of the delta function and its derivatives. Furthermore, we assume that the state of the system at time 0^- is

$$\underline{x}(0^-) = \underline{x}'(0^-) = \cdots = \underline{x}^{(r-1)}(0^-) = [0 \quad 0 \quad \ldots \quad 0]^t,$$

and at time 0^+, it achieves

$$\underline{x}(0^+) = \left[x_1^0 \quad x_2^0 \quad \ldots \quad x_n^0\right]^t,\ \underline{x}'(0^+) = \left[x_1^1 \quad x_2^1 \quad \ldots \quad x_n^1\right]^t, \ldots,$$

$$\underline{x}^{(r-1)}(0^+) = \left[x_1^{r-1} \quad x_2^{r-1} \quad \ldots \quad x_n^{r-1}\right]^t.$$

Obviously, such an input which can be expressed by the summation (3.1.1), is very hard to imagine *physically*. However, we can think of it approximately as a combination of small pulses of very high magnitude and infinitely small duration.

In the paper by Karageorgos, Pantelous and Kalogeropoulos (2009), a classical approximated expression for the controller (3.1.1), which is based on the Gaussian (Normal) function, is used. Thus, by considering what are the Dirac δ- function and the Gaussian (Normal) function we obtain:

$$\delta(t) = \lim_{\sigma \to 0} \frac{1}{\sigma\sqrt{2\pi}} e^{-t^2/2\sigma^2} = \lim_{\sigma \to 0} \frac{1}{\sigma} \phi\left(\frac{t}{\sigma}\right),$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

So, the approximate expression for the impulsive-input (3.1.1) is given by

$$u(t) = \sum_{k=0}^{n-1} \frac{1}{\sigma^{k+1}} \phi^{(k)}\left(\frac{t}{\sigma}\right) a_k .$$

Then, we can take the limit $u_o(t) = \lim_{\sigma \to 0} u(t)$.

Thus, in the paper proposed by Karageorgos, Pantelous and Kalogeropoulos (2009), the unknown vector-coefficient $\underline{a} = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-1} \end{bmatrix}^t$, where $a_i \in \mathbb{F}$ for $i = 0,1,\ldots,$ $n-1$ has been analytically calculated by solving the system (3.1.2).

$$\begin{bmatrix} V_l \\ V_{(l+1)_{z_{l+1}}} \\ V_{(l+2)_{z_{l+2}}} \\ \vdots \\ V_{\kappa_{d_\kappa}} \end{bmatrix} \underline{a} = \begin{bmatrix} \underline{z}_l\left(0^+\right) \\ \underline{z}_{(l+1)_{z_{l+1}}}\left(0^+\right) \\ \underline{z}_{(l+2)_{z2}}\left(0^+\right) \\ \vdots \\ \underline{z}_{\kappa_{z_\kappa}}\left(0^+\right) \end{bmatrix}, \tag{3.1.2}$$

where the vector $\left[\underline{z}_l^t\left(0^+\right)\ \underline{z}_{(l+1)_{z_{l+1}}}^t\left(0^+\right)\ \underline{z}_{(l+2)_{z2}}^t\left(0^+\right)\cdots\ \underline{z}_{\kappa_{z_\kappa}}^t\left(0^+\right)\right]^t$ is constant, $V_l \in \mathcal{M}_{s,n}$ is a *rectangular* $s\times n$ *-Vandermonde* matrix and $V_{j_{z_j}} \in \mathcal{M}_{\mu_{z_j},n}$, with $j=l+1,\ l+2,\ldots,\ \kappa$ and $z_j=1,2,\ldots,d_j$, is a *special* matrix.

Obviously, the system (3.1.2) can be further transposed to a more convenient system. Analytically, if we multiply the 1st row of *Vandermonde* matrix V_l, i.e. $\left[1\ \lambda\ \lambda^2\ \ldots\ \lambda^{n-1}\right]$ with the number (-1) and we added it to the 1st row of each of $V_{(l+1)_{z_{l+1}}}$, $V_{(l+2)_{z_{l+2}}}$,..., $V_{\kappa_{z_\kappa}}$, then $V_{j_{z_j}}$ is given by

$$V_{j_{z_j}}=\begin{bmatrix} 0 & \mu_j-\lambda & \mu_j^2-\lambda^2 & \cdots & \cdots & \mu_j^{n-1}-\lambda^{n-1} \\ 0 & 1 & 2\mu_j & \cdots & \cdots & (n-1)\mu_j^{n-2} \\ 0 & 0 & 1 & \cdots & \cdots & (n-1)(n-2)\mu_j^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 1 & \dfrac{1}{(\rho_j-1)!}\dfrac{d^{\rho_j-1}\mu_j^{\rho_{j+1}}}{d\mu_j^{\rho_j-1}}\cdots\dfrac{1}{(\rho_j-1)!}\dfrac{d^{\rho_j-1}\mu_j^{n-1}}{d\mu_j^{\rho_j-1}} \end{bmatrix}. \quad (3.1.3)$$

(Note that we have shown that the matrices $V_{j_{z_j}}$, for $j=l+1,\ l+2,\ldots,\ \kappa$, do not contain zero rows, see also *Comment*, Karageorgos, Pantelous and Kalogeropoulos, 2009)

We can easily see that the 1st row of matrix (3.1.3) can be re-written as below, i.e. the element $\mu_j^\kappa+\mu_j^{\kappa-1}\lambda+\ldots+\mu_j\lambda^{\kappa-1}+\lambda^\kappa=\sum_{\substack{k_1,k_2=0\\ \sum_{i=1}^{2}k_i=\kappa}}^{\kappa}\mu_j^{k_1}\lambda^{k_2}$.

Thus, the first row is presented as

$$(\mu_j-\lambda)\left[0\ 1\ \sum_{\substack{k_1,k_2=0\\ \sum_{i=1}^{2}k_i=1}}^{1}\mu_j^{k_1}\lambda^{k_2}\ \sum_{\substack{k_1,k_2=0\\ \sum_{i=1}^{2}k_i=2}}^{2}\mu_j^{k_1}\lambda^{k_2}\ \ldots\ \sum_{\substack{k_1,k_2=0\\ \sum_{i=1}^{2}k_i=n-2}}^{n-2}\mu_j^{k_1}\lambda^{k_2}\right].$$

Since, the element $\mu_j - \lambda \neq 0$, we can multiply by left the eq. (3.1.3) with a properly chosen transformation matrix, so as to obtain

$$S_{j_{z_j}} \triangleq \begin{bmatrix} 0 & 1 & \sum_{\substack{k_1,k_2=0 \\ \sum_{i=1}^{2} k_i = 1}}^{1} \mu_j^{k_1} \lambda^{k_2} & \cdots & \cdots & \cdots & \sum_{\substack{k_1,k_2=0 \\ \sum_{i=1}^{2} k_i = n-2}}^{n-2} \mu_j^{k_1} \lambda^{k_2} \\ 0 & 1 & 2\mu_j & \cdots & \cdots & \cdots & (n-1)\mu_j^{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 1 & \frac{1}{(\rho_j - 1)!} \frac{d^{\rho_j - 1} \mu_j^{\rho_j + 1}}{d\mu_j^{\rho_j - 1}} & \cdots & \frac{1}{(\rho_j - 1)!} \frac{d^{\rho_j - 1} \mu_j^{n-1}}{d\mu_j^{\rho_j - 1}} \end{bmatrix}. \quad (3.1.4)$$

Finally, the system (3.1.5) is derived, where the matrices $S_{j_{z_j}}$ for $j = l+1, l+2, \ldots, \kappa$ are derived by taking into account a properly chosen transformation left-matrix Z, as follows

$$Z \begin{bmatrix} V_l \\ V_{(l+1)_{z_{l+1}}} \\ V_{(l+2)_{z_{l+2}}} \\ \vdots \\ V_{\kappa_{d_\kappa}} \end{bmatrix} \underline{a} = Z \begin{bmatrix} \underline{z}_l^*(0^+) \\ \underline{z}_{(l+1)_{z_{l+1}}}(0^+) \\ \underline{z}_{(l+2)_{z2}}(0^+) \\ \vdots \\ \underline{z}_{\kappa_{z_\kappa}}(0^+) \end{bmatrix} \Leftrightarrow \begin{bmatrix} V_l \\ S_{l+1} \\ S_{l+2} \\ \vdots \\ S_\kappa \end{bmatrix} \underline{a} = \begin{bmatrix} \underline{z}_l(0^+) \\ \underline{d}_{l+1}(0^+) \\ \underline{d}_{l+2}(0^+) \\ \vdots \\ \underline{d}_\kappa(0^+) \end{bmatrix}, \quad (3.1.5)$$

where $V_l \in \mathcal{M}_{l,n}$, $S_j \in \mathcal{M}_{j,n}$, $\underline{z}_l(0^+) \in \mathcal{M}_{l,1}$ and $\underline{d}_{l+2}(0^+) \in \mathcal{M}_{j,1}$, for $j = l+1, l+2, \ldots, \kappa$. Note that $\rho_j = \max_{z_i = 1,2,\ldots,d_j} \mu_{z_j}$ is the index of annihilation for the eigenvalue μ_{z_j}.

Consequently, the system (3.1.5) contains the following sub-systems.

$$\left.\begin{array}{c} V_l \underline{a} = \underline{z}_l(0^+) \\ S_{l+1}\underline{a} = \underline{d}_{l+1}(0^+) \\ \cdots \\ S_\kappa \underline{a} = \underline{d}_\kappa(0^+) \end{array}\right\},$$

where V_l, S_j for $j = l+1, l+2, \ldots, \kappa$ are non-square matrices.

Thus, for the analytic solution of the above system, i.e. for the determination of the coefficients $\underline{a}$ of the input (3.1.1), some elements of the generalized inverse theory are needed.

More analytically, in the sub-section 3.2, we investigate the generalized inverses of the rectangular Vandermonde matrix, V. According to the number of rows and columns, different types of generalized inverses derive. In the sub-section 3.3, we investigate the {1, 2, 3}- generalized inverse of a very special rectangular matrix S. For the better understanding of the presented results, some numerical examples are considered. The 3.4 sub-section concludes the whole chapter. Further directions for research are also presented.

As a last part of this introduction, the following basic definitions for different kind of generalized inverses are simply repeated; see for more details Campbell and Meyer, Jr (1979).

Definition 3.1.1 *Denote the square matrix* $A \in \mathcal{M}_n$. *We say that the non-negative integer* k *is the index of* A, $Ind(A) = k$, *if* k *is the smallest non-negative integer such as*

$$rank\left(A^k\right) = rank\left(A^{k+1}\right). \qquad \square$$

Definition 3.1.2 *The Moore-Penrose inverse of a rectangular matrix* $A \in \mathcal{M}_{m,n}$ *is the matrix* $A^\dagger \in \mathcal{M}_{m,n}$ *such that*

$$(1)\ AA^\dagger A = A,$$

$$(2)\ A^\dagger AA^\dagger = A^\dagger,$$

$$(3)\ \left(AA^\dagger\right)^* = AA^\dagger,$$

$$(4)\ \left(A^\dagger A\right)^* = A^\dagger A,$$

where $*$ *the conjugate transpose index of the relevant matrix.* $\square$

Moreover, the *Drazin inverse* of square matrix $A \in \mathcal{M}_n$, $Ind(A) = k$ is the matrix A^D satisfying

$$\text{(i)}\ A^D A A^D = A^D,$$

$$\text{(ii)}\ A A^D = A^D A,$$

$$\text{(iii)}\ A^{l+1} A^D = A^l,$$

for $l \geq k = Ind(A)$.

Note that if A is non-singular, then $A^\dagger \equiv A^D \equiv A^{-1}$.

3.2 The Generalized Inverses of the Vandermonde Matrix

In this section, we study three different cases of the $V_{m,n} \equiv V_n(\lambda_1, \lambda_2, ..., \lambda_m)$, Vandermonde matrix. The first two cases, where $n > m$ and $n < m$, create the rectangular Vandermonde matrix with different number of rows and columns and the third one creates the more classical square Vandermonde matrix $V_n = V_n(\lambda_1, \lambda_2, ..., \lambda_n)$.

For all cases, our wish is to transform the Vandermonde matrix

$$V_{m,n} \equiv V_n(\lambda_1, \lambda_2, ..., \lambda_m) \triangleq \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_m & \cdots & \lambda_m^{n-1} \end{bmatrix} \in \mathcal{M}_{m,n},$$

into an equivalent matrix of the following form:

I) For $n > m$ (rectangular case with more rows than columns), we obtain

$$\begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \end{bmatrix}.$$

II) For $n < m$ (rectangular case with more columns than rows), we obtain

$$\begin{bmatrix} I_n \\ \mathbb{O}_{m-n,n} \end{bmatrix}.$$

III) For $n = m$ (square case, equal number of rows and columns), we obtain

$$I_n.$$

Definition 3.2.1 *Consider the following matrices:*

a) Let $P_i(a)$ be a $m \times m$ -matrix which has a non-zero element a in the i^{th}-row and the j^{th} - column, i.e.

$$P_i(a)=\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & \mathbb{O} & \\ & & 1 & & & & \\ & & & a & & & \\ & & & & 1 & & \\ & \mathbb{O} & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}. \qquad (3.2.1)$$

Thus, whenever a matrix A is multiplied from the left by $P_i(a)$ then the i^{th}-row of it is multiplied by the non-zero number a.

b) Let $P_i(j,a)$ be a $m\times m$-matrix which has a non-zero element a in the i^{th}-row and the j^{th}-column, i.e.

$$P_i(j,a)=\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & \mathbb{O} & \\ & & 1 & \cdots & a & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & \mathbb{O} & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}. \qquad (3.2.2)$$

Thus, whenever a matrix A is multiplied from the left by $P_i(j,a)$ then the j^{th}-row of it is multiplied by the non-zero number a and it is added to the j^{th}-row of A.

c) Let $Q_i(a)$ be a $n\times n$-matrix which has a non-zero element a in the j^{th}-row and the i^{th}-column, i.e.

$$Q_i(a)=\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & \mathbb{O} & \\ & & 1 & & & & \\ & & & a & & & \\ & & & & 1 & & \\ & \mathbb{O} & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}. \tag{3.2.3}$$

Thus, whenever a matrix A is multiplied from the right by $Q_i(a)$ then the i^{th}-column of it is multiplied by the non-zero number a.

d) Let $Q_i(j,a)$ be a $n\times n$-matrix which has a non-zero element a in the j^{th}-row and the i^{th}-column, i.e.

$$Q_i(j,a)=\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & \mathbb{O} & \\ & & 1 & & & & \\ & & \vdots & \ddots & & & \\ & & a & \cdots & 1 & & \\ & \mathbb{O} & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}. \tag{3.2.4}$$

Thus, whenever a matrix A is multiplied from the right by $Q_i(j,a)$ then the i^{th}-column of it is multiplied by the non-zero number a and it is added to the j^{th}-column of A. □

Definition 3.2.2 *Let us define with the $\coprod\cdot$ symbol the order left multiplication of matrices as it is given by* $\coprod_{m}^{j=1} P_j = P_m P_{m-1}\cdots P_2 P_1$. □

Proposition 3.2.1 (*Vandermonde parameterization*)

a) For the (I) case, there are invertible matrices $\mathbf{P}_1\in\mathcal{M}_m$ and $\mathbf{Q}_1\in\mathcal{M}_n$ such that

$$\mathbf{P}_1 V_{m,n} \mathbf{Q}_1 = \mathbf{P}_1 \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_m & \cdots & \lambda_m^{n-1} \end{bmatrix} \mathbf{Q}_1 = \begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \end{bmatrix}, \tag{3.2.5}$$

where the permutated matrices are given analytically by the following expressions (3.2.6) and (3.2.7), i.e.

$$\mathbf{P}_1 = \coprod_{m-1}^{s=1} \coprod_{m}^{j=s+1} P_j\left(\frac{1}{\lambda_j - \lambda_s}\right) P_j(s,-1) \in \mathcal{M}_m, \tag{3.2.6}$$

where $P_i(a)$ *and* $P_i(j,a)$ *are given by (3.2.1) and (3.2.2) respectively, and*

$$\mathbf{Q}_1 = \prod_{s=1}^{m} \prod_{r=s+1}^{n} Q_r\left(s, -\sum_{\substack{k_1,\ldots,k_s=0 \\ k_1+\ldots+k_s=r-s}}^{r-s} \prod_{l=1}^{s} \lambda_l^{k_l}\right) \in \mathcal{M}_n \ , \tag{3.2.7}$$

where $Q_i(j,a)$ *is given by (3.2.4).*

b) For the (II) case, there are invertible matrices $\mathbf{P}_2 \in \mathcal{M}_m$ *and* $\mathbf{Q}_2 \in \mathcal{M}_n$ *such that*

$$\mathbf{P}_2 V_{m,n} \mathbf{Q}_2 = \mathbf{P}_2 \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_m & \cdots & \lambda_m^{n-1} \end{bmatrix} \mathbf{Q}_2 = \begin{bmatrix} I_n \\ \mathbb{O}_{m-n,n} \end{bmatrix}, \tag{3.2.8}$$

where the permutated matrices are given analytically by the following expressions (3.2.9) and (3.2.10), i.e.

$$\mathbf{P}_2 = \prod_{z=n+1}^{m} P_z(n,-1) \coprod_{n-1}^{s=1} \coprod_{m}^{j=s+1} P_j\left(\frac{1}{\lambda_j - \lambda_s}\right) P_j(s,-1) \in \mathcal{M}_m, \tag{3.2.9}$$

where $P_i(a)$ *and* $P_i(j,a)$ *are given by (3.2.1) and (3.2.2) respectively, and*

$$\mathbf{Q}_2 = \prod_{s=1}^{n-1} \prod_{r=s+1}^{n} Q_r \left(s, - \sum_{\substack{k_1,\ldots,k_s=0 \\ k_1+\ldots+k_s=r-s}}^{r-s} \prod_{l=1}^{s} \lambda_l^{k_l} \right) \in \mathcal{M}_n, \tag{3.2.10}$$

where $Q_i(j,a)$ *is given by (3.2.4).*

c) For the (III) case, there are invertible matrices $\mathbf{P}_3 \in \mathcal{M}_n$ *and* $\mathbf{Q}_3 \in \mathcal{M}_n$ *such that*

$$\mathbf{P}_3 V_n \mathbf{Q}_3 = \mathbf{P}_3 \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{bmatrix} \mathbf{Q}_3 = I_n, \tag{3.2.11}$$

where the permutated matrices are given analytically by the following expression (3.2.12) and (3.2.13), i.e.

$$\mathbf{P}_3 = \coprod_{n-1}^{s=1} \coprod_{n}^{j=s+1} P_j \left(\frac{1}{\lambda_j - \lambda_s} \right) P_j(s,-1) \in \mathcal{M}_n, \tag{3.2.12}$$

where $P_i(a)$ *and* $P_i(j,a)$ *are given by (3.2.1) and (3.2.2) respectively, and*

$$\mathbf{Q}_3 = \prod_{s=1}^{n-1} \prod_{r=s+1}^{n} Q_r \left(s, - \sum_{\substack{k_1,\ldots,k_s=0 \\ k_1+\ldots+k_s=r-s}}^{r-s} \prod_{l=1}^{s} \lambda_l^{k_l} \right) \in \mathcal{M}_n \ , \tag{3.2.13}$$

where $Q_i(j,a)$ *is given by (3.2.4)* .

Proof. (I) For $n > m$. We start with the rectangular Vandermonde matrix

$$V_{m,n} \equiv V_n(\lambda_1, \lambda_2, ..., \lambda_m) \in \mathcal{M}_{m,n}$$

and we work as follows

$$P_2(1,-1)V_{m,n} \to P_2\left(\frac{1}{\lambda_2 - \lambda_1}\right) P_2(1,-1)V_{m,n} \to P_3(1,-1)\, P_2\left(\frac{1}{\lambda_2 - \lambda_1}\right) P_2(1,-1)V_{m,n}$$
$$\to P_3\left(\frac{1}{\lambda_3 - \lambda_1}\right) P_3(1,-1)\, P_2\left(\frac{1}{\lambda_2 - \lambda_1}\right) P_2(1,-1)V_{m,n} \to \ldots$$

$$\to P_m\left(\frac{1}{\lambda_m-\lambda_{m-1}}\right)P_m\left(m-1,-1\right)P_m\left(\frac{1}{\lambda_m-\lambda_{m-2}}\right)P_m\left(m-2,-1\right)$$

$$P_{m-1}\left(\frac{1}{\lambda_{m-1}-\lambda_{m-2}}\right)P_{m-1}\left(m-2,-1\right)P_m\left(\frac{1}{\lambda_m-\lambda_{m-2}}\right)P_m\left(m-3,-1\right)\cdot$$

$$\cdot P_{m-1}\left(\frac{1}{\lambda_{m-1}-\lambda_{m-3}}\right)P_{m-1}\left(m-3,-1\right)\cdot$$

$$\cdot P_{m-2}\left(\frac{1}{\lambda_{m-2}-\lambda_{m-3}}\right)P_{m-2}\left(m-3,-1\right)\cdots P_m\left(\frac{1}{\lambda_m-\lambda_1}\right)P_m\left(1,-1\right)P_{m-1}\left(\frac{1}{\lambda_{m-1}-\lambda_1}\right)\cdot$$

$$\cdot P_{m-1}\left(1,-1\right)\cdots P_2\left(\frac{1}{\lambda_2-\lambda_1}\right)P_2\left(1,-1\right)V_{m,n}.$$

Then, the matrix V_{mn} is transformed into

$$\begin{bmatrix}
1 & \lambda_1 & \lambda_1^2 & \lambda_1^3 & \lambda_1^4 & \cdots & * & * & \cdots & \lambda_1^{n-1} \\
0 & 1 & \lambda_1+\lambda_2 & \lambda_1^2+\lambda_1\lambda_2+\lambda_2^2 & \sum\limits_{\substack{k_1,k_2=0\\ k_1+k_2=3}}^{3}\lambda_1^{k_1}\lambda_2^{k_2} & \cdots & * & * & \cdots & \sum\limits_{\substack{k_1,k_2=0\\ k_1+k_2=n-2}}^{n-2}\lambda_1^{k_1}\lambda_2^{k_2} \\
0 & 0 & 1 & \sum\limits_{\substack{k_1,k_2,k_3=0\\ k_1+k_2+k_3=1}}^{1}\lambda_1^{k_1}\lambda_2^{k_2}\lambda_3^{k_3} & \sum\limits_{\substack{k_1,k_2,k_3=0\\ k_1+k_2+k_3=2}}^{2}\lambda_1^{k_1}\lambda_2^{k_2}\lambda_3^{k_3} & \cdots & * & * & \cdots & \sum\limits_{\substack{k_1,k_2,k_3=0\\ k_1+k_2+k_3=n-3}}^{n-3}\lambda_1^{k_1}\lambda_2^{k_2}\lambda_3^{k_3} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 1 & \sum\limits_{\substack{k_1,k_2,\ldots,k_m=0\\ k_1+k_2+\ldots+k_m=1}}^{1}\prod\limits_{l=1}^{m}\lambda_l^{k_l} & \cdots & \sum\limits_{\substack{k_1,k_2,\ldots,k_q=0\\ k_1+k_2+\ldots+k_q=n-1-(m-1)}}^{n-1-(m-1)}\prod\limits_{l=1}^{m}\lambda_l^{k_l}
\end{bmatrix}.$$

(Note that $\sum_{\substack{k_1,k_2,\ldots,k_m=0\\ k_1+k_2+\ldots+k_m=n-1-(m-1)}}^{n-1-(m-1)}\prod_{l=1}^{m}\lambda_l^{k_l}$ is a sum from 0 to $n-1-(m-1)$ such as the

$$\sum_{z=0}^{n-1-(m-1)}k_z=n-m\in\mathbb{N}.)$$

Thus, we conclude to the determination of the transformation matrix $\mathbf{P}_1$

$$\mathbf{P}_1=P_m\left(\frac{1}{\lambda_m-\lambda_{m-1}}\right)P_m\left(m-1,-1\right)\cdots P_2\left(\frac{1}{\lambda_2-\lambda_1}\right)P_2\left(1,-1\right)\triangleq\prod_{m-1}^{s=1}\prod_{m}^{j=s+1}P_j\left(\frac{1}{\lambda_j-\lambda_s}\right)P_j\left(s,-1\right)$$

As we can see the multiplication between matrices counts in reverse order, starting from $m\to m-1\to\ldots$ etc, see also Definition 3.2.2.

Now, we want to transfer the $\mathbf{P}_1 V_{m,n}$ into the desired matrix (3.2.5), so we act as follows,

$$\mathbf{P}_1 V_{m,n} Q_2\left(1,-\lambda_1\right) \to \mathbf{P}_1 V_{m,n} Q_2\left(1,-\lambda_1\right) P_3\left(1,-\lambda_1^2\right) \to \cdots$$

$$\to \mathbf{P}_1 V_{m,n} Q_2\left(1,-\lambda_1\right) Q_3\left(1,-\lambda_1^2\right) \cdots Q_n\left(1,-\lambda_1^{n-1}\right)$$

$$\to \mathbf{P}_1 V_{m,n} Q_2\left(1,-\lambda_1\right) Q_3\left(1,-\lambda_1^2\right) \cdots Q_n\left(1,-\lambda_1^{n-1}\right) Q_3\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{1}\prod_{l=1}^{2}\lambda_l^{k_l}\right)$$

$$\to \mathbf{P}_1 V_{m,n} Q_2\left(1,-\lambda_1\right) Q_3\left(1,-\lambda_1^2\right) \cdots Q_n\left(1,-\lambda_1^{n-1}\right) Q_3\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{1}\prod_{l=1}^{2}\lambda_l^{k_l}\right) Q_4\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=2}}^{2}\prod_{l=1}^{2}\lambda_l^{k_l}\right)$$

$$\to \cdots \to$$

$$\mathbf{P}_1 V_{m,n} Q_2\left(1,-\lambda_1\right) Q_3\left(1,-\lambda_1^2\right) \cdots Q_n\left(1,-\lambda_1^{n-1}\right) Q_3\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{1}\prod_{l=1}^{2}\lambda_l^{k_l}\right) \cdots Q_n\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=n-2}}^{n-2}\prod_{l=1}^{2}\lambda_l^{k_l}\right)$$

$$Q_4\left(3,-\sum_{\substack{k_1,k_2,k_3=0\\k_1+k_2+k_3=1}}^{1}\prod_{l=1}^{3}\lambda_l^{k_l}\right) \cdots Q_n\left(3,-\sum_{\substack{k_1,k_2,k_3=0\\k_1+k_2+k_3=n-3}}^{n-3}\prod_{l=1}^{3}\lambda_l^{k_l}\right) \cdots Q_{m+1}\left(m,-\sum_{\substack{k_1,k_2,\ldots,k_m=0\\k_1+k_2+\ldots+k_m=1}}^{1}\prod_{l=1}^{n}\lambda_l^{k_l}\right)$$

$$\cdots Q_n\left(2,-\sum_{\substack{k_1,k_2,\ldots,k_m=0\\k_1+k_2+\ldots+k_m=n-1-(m-1)}}^{n-1-(m-1)}\prod_{l=1}^{m}\lambda_l^{k_l}\right).$$

Thus, now we can define matrix the transformation matrix

$$\mathbf{Q}_1 = Q_2\left(1,-\lambda_1\right) Q_3\left(1,-\lambda_1^2\right) \cdots Q_n\left(1,-\lambda_1^{n-1}\right) Q_3\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{1}\prod_{l=1}^{2}\lambda_l^{k_l}\right) \cdots Q_n\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=n-2}}^{n-2}\prod_{l=1}^{2}\lambda_l^{k_l}\right)$$

$$Q_4\left(3,-\sum_{\substack{k_1,k_2,k_3=0\\k_1+k_2+k_3=1}}^{1}\prod_{l=1}^{3}\lambda_l^{k_l}\right) \cdots Q_n\left(3,-\sum_{\substack{k_1,k_2,k_3=0\\k_1+k_2+k_3=n-3}}^{n-3}\prod_{l=1}^{3}\lambda_l^{k_l}\right) \cdots Q_{m+1}\left(m,-\sum_{\substack{k_1,k_2,\ldots,k_m=0\\k_1+k_2+\ldots+k_m=1}}^{1}\prod_{l=1}^{n}\lambda_l^{k_l}\right)$$

$$\cdots Q_n\left(2,-\sum_{\substack{k_1,k_2,\ldots,k_m=0\\k_1+k_2+\ldots+k_m=n-1-(m-1)}}^{n-1-(q-1)}\prod_{l=1}^{m}\lambda_l^{k_l}\right) \triangleq \prod_{s=1}^{m}\prod_{r=s+1}^{n} Q_r\left(s,-\sum_{\substack{k_1,\ldots,k_s=0\\k_1+\ldots+k_s=r-s}}^{r-s}\prod_{l=1}^{s}\lambda_l^{k_l}\right).$$

Consequently, we have transposed the Vandermonde matrix into (3.2.5).

Similarly, we can work for the (II) and (III) cases, where the expression (3.2.8) - (3.2.13) derive.

So, the further details are omitted. □

In the next example, we illustrate the results of Proposition 3.2.1.

Example 3.2.1 Suppose that we have the 3×4 - Vandermonde matrix, i.e. (I) case,

$$V_{3,4}=V_4\left(\lambda_1=3,\lambda_2=10,\lambda_3=100\right)=\begin{bmatrix}1 & 3 & 3^2 & 3^3\\ 1 & 10 & 10^2 & 10^3\\ 1 & 100 & 100^2 & 100^3\end{bmatrix}\in \mathcal{M}_{3,4},$$

then by applying (3.2.6), we take

$$\mathbf{P}_1=\prod_{2}^{s=1}\prod_{3}^{j=s+1}P_j\left(\frac{1}{\lambda_j-\lambda_s}\right)P_j\left(s,-1\right)=P_3\left(\frac{1}{90}\right)P_3\left(2,-1\right)P_3\left(\frac{1}{97}\right)P_3\left(1,-1\right)P_2\left(\frac{1}{7}\right)P_2\left(1,-1\right)$$

$$=\begin{bmatrix}1 & 0 & 0\\ -1/7 & 1/7 & 0\\ 1/679 & -1/630 & 1/8730\end{bmatrix}.$$

The matrix $V_{3,4}$ is being transformed into the following

$$\mathbf{P}_1V_{3,4}=\begin{bmatrix}1 & \lambda_1 & \lambda_1^2 & \lambda_1^3\\ 0 & 1 & \lambda_1+\lambda_2 & \lambda_1^2+\lambda_1\lambda_2+\lambda_2^2\\ 0 & 0 & 1 & \lambda_1+\lambda_2+\lambda_3\end{bmatrix}=\begin{bmatrix}1 & 3 & 3^2 & 3^3\\ 0 & 1 & 3+10 & 3^2+3\cdot 10+10^2\\ 0 & 0 & 1 & 3+10+100\end{bmatrix}$$

$$=\begin{bmatrix}1 & 3 & 9 & 27\\ 0 & 1 & 13 & 139\\ 0 & 0 & 1 & 113\end{bmatrix}.$$

Now, we want to transfer the $\mathbf{P}_1V_{3,4}$ into the matrix $\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}$, so we apply (3.2.7),

i.e.

$$\mathbf{Q}_1 = \prod_{s=1}^{3}\prod_{r=s+1}^{4} Q_r\left(s, -\sum_{\substack{k_1,\ldots,k_s=0 \\ k_1+\ldots+k_s=r-s}}^{r-s}\prod_{l=1}^{s}\lambda_l^{k_l}\right)$$

$$= Q_2(1,-3)\,Q_3\left(1,-3^2\right)Q_4\left(1,-3^3\right)Q_3(2,-13)\,Q_4(2,-139)\,Q_4(3,-113)$$

$$= \begin{bmatrix} 1 & -3 & 30 & -3000 \\ 0 & 1 & -13 & 1330 \\ 0 & 0 & 1 & -113 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus, we take the parameterization

$$\mathbf{P}_1 V_{3,4} \mathbf{Q}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

where

$$\mathbf{P}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1/7 & 1/7 & 0 \\ 1/679 & -1/630 & 1/8730 \end{bmatrix} \quad \text{and} \quad \mathbf{Q}_1 = \begin{bmatrix} 1 & -3 & 30 & -3000 \\ 0 & 1 & -13 & 1330 \\ 0 & 0 & 1 & -113 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Remark 3.2.1 It is not difficult to verify (see also the above numerical example) that an explicit (quasi-) **LU** factorization of the rectangular Vandermonde matrix is obtained using non-singular matrices like (3.2.6), (3.2.7) (or (3.2.9), (3.2.10) or (3.2.12), (3.2.13)). Analytically, we have

(I) *Quasi* **LU** factorization:

$$V_{m,n} = \tilde{\mathbf{P}}_1 \begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \end{bmatrix} \tilde{\mathbf{Q}}_1,$$

where

$$\tilde{\mathbf{P}}_1 \triangleq \mathbf{P}_1^{-1} = \prod_{s=1}^{m-1}\prod_{j=s+1}^{m} P_j(s,1)\,P_j\left(\lambda_j - \lambda_s\right) \text{ (}\textit{Lower}\text{ Triangular Matrix)}$$

and

$$\tilde{\mathbf{Q}}_1 \triangleq \mathbf{Q}_1^{-1} = \coprod_{m}^{s=1} \coprod_{n}^{r=s+1} Q_r \left(s, \sum_{\substack{k_1,\ldots,k_s=0 \\ k_1+\ldots+k_s=r-s}}^{r-s} \prod_{l=1}^{s} \lambda_l^{k_l} \right) \text{ (}\textit{Upper}\text{ Triangular Matrix).}$$

(II) *Quasi* **LU** factorization:

$$V_{m,n} = \tilde{\mathbf{P}}_2 \begin{bmatrix} I_n \\ \mathbb{O}_{n-m,m} \end{bmatrix} \tilde{\mathbf{Q}}_2 ,$$

where

$$\tilde{\mathbf{P}}_2 \triangleq \mathbf{P}_2^{-1} = \prod_{s=1}^{n-1} \prod_{j=s+1}^{m} P_j(s,1) P_j\left(\lambda_j - \lambda_s\right) \prod_{z=n+1}^{m} P_z(n,1) \text{ (}\textit{Upper}\text{ Triangular Matrix)}$$

and

$$\tilde{\mathbf{Q}}_2 \triangleq \mathbf{Q}_2^{-1} = \coprod_{n-1}^{s=1} \coprod_{n}^{r=s+1} Q_r \left(s, \sum_{\substack{k_1,\ldots,k_s=0 \\ k_1+\ldots+k_s=r-s}}^{r-s} \prod_{l=1}^{s} \lambda_l^{k_l} \right) \text{ (}\textit{Lower}\text{ Triangular Matrix).}$$

(III) **LU** factorization:

$$V_n = \tilde{\mathbf{P}}_3 \tilde{\mathbf{Q}}_3 ,$$

where

$$\tilde{\mathbf{P}}_3 \triangleq \mathbf{P}_3^{-1} = \prod_{s=1}^{n-1} \prod_{j=s+1}^{n} P_j(s,1) P_j\left(\lambda_j - \lambda_s\right) \text{ (}\textit{Upper}\text{ Triangular Matrix)}$$

and

$$\tilde{\mathbf{Q}}_2 \triangleq \mathbf{Q}_2^{-1} = \coprod_{n-1}^{s=1} \coprod_{n}^{r=s+1} Q_r \left(s, \sum_{\substack{k_1,\ldots,k_s=0 \\ k_1+\ldots+k_s=r-s}}^{r-s} \prod_{l=1}^{s} \lambda_l^{k_l} \right) \text{ (}\textit{Lower}\text{ Triangular Matrix).}$$

Thus, for each (I) - (III) case, *upper* and *lower* triangular non-singular matrices are derived. The proposed results are compared with those derived from Kaufman (1969), Martinez and Peña (1998a) and Oruç and Phillips (2000) (see also Remark 3.2.2).

Characteristically, we remind that in Oruç and Phillips (2000), the **LU** factorization of a square Vandermonde matrix is obtained using complete symmetric functions. Our results are fully comparable to Oruç and Phillips (2000), since we can have also explicit formulae for the factorization matrices. □

Remark 3.2.2 As a further direction, but it is beyond the scopes of this chapter, it would be very interesting to compare the numerical results obtained by the LU factorization, especially for the (III) case –i.e. square Vandermonde matrix, with those derived in Björck and Pereyra (1970), Tang and Golub (1981), Oruç and Phillips (2000), Eisinberg, Franzé and Salerno (2001). □

In the next lines, we provide the main results of this section. The generalized inverses of the rectangular and square Vandermonde matrices are derived. Furthermore, it should be pointed out that analytical formulae for the calculation of the generalized inverses derive.

Theorem 3.2.1 *For the (I) case, the* $\{1,2,3\}$*-inverse of the rectangular Vandermonde matrix is given by*

$$V_{n,m}^{\{1,2,3\}} = \mathbf{Q}_1 \begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix} \mathbf{P}_1 \in \mathcal{M}_{n,m}, \tag{3.2.14}$$

where the permutated matrices $\mathbf{P}_1$ *and* $\mathbf{Q}_1$ *are given by the expressions (3.2.6) and (3.2.7), respectively.*

Proof. Consider the expression (3.2.5), i.e.

$$\mathbf{P}_1 V_{m,n} \mathbf{Q}_1 = \begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \end{bmatrix} \Leftrightarrow V_{m,n} = \mathbf{P}_1^{-1} \begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \end{bmatrix} \mathbf{Q}_1^{-1}.$$

In order the matrix $\mathbf{Q}_1 \begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix} \mathbf{P}_1$ to be the {1, 2, 3}-inverse of $V_{m,n}$, we have to prove the following three equalities,

$$(1)\quad V_{m,n}\mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1 V_{m,n} = V_{m,n},$$

$$(2)\quad \mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1 V_{m,n}\mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1 = \mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1$$

and

$$(3)\quad \left(V_{m,n}\mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1\right)^* = V_{m,n}\mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1.$$

Thus, the (1) holds since

$$\begin{aligned} V_{m,n}\mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1 V_{m,n} &= \mathbf{P}_1^{-1}\begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \end{bmatrix}\mathbf{Q}_1^{-1}\mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1\mathbf{P}_1^{-1}\begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \end{bmatrix}\mathbf{Q}_1^{-1} \\ &= \mathbf{P}_1^{-1}\begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \end{bmatrix}\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \end{bmatrix}\mathbf{Q}_1^{-1} \\ &= \mathbf{P}_1^{-1}\left(I_m + \mathbb{O}_m\right)\begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \end{bmatrix}\mathbf{Q}_1^{-1} \\ &= V_{m,n}, \end{aligned}$$

and the (2) holds since

$$\begin{aligned} \mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1 V_{m,n}\mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1 &= \mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1\mathbf{P}_1^{-1}\begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \end{bmatrix}\mathbf{Q}_1^{-1}\mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1 \\ &= \mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \end{bmatrix}\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1 \\ &= \mathbf{Q}_1\begin{bmatrix} I_m & \mathbb{O}_{m,n-m} \\ \mathbb{O}_{n-m,m} & \mathbb{O}_{n-m,n-m} \end{bmatrix}\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1 \\ &= \mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m} \end{bmatrix}\mathbf{P}_1, \end{aligned}$$

and finally, the (3) also holds since

$$\left(V_{m,n}\mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m}\end{bmatrix}\mathbf{P}_1\right)^* = \left(\mathbf{P}_1^{-1}\begin{bmatrix} I_m & \mathbb{O}_{m,n-m}\end{bmatrix}\mathbf{Q}_1^{-1}\mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m}\end{bmatrix}\mathbf{P}_1\right)^*$$

$$= \left(\mathbf{P}_1^{-1}\begin{bmatrix} I_m & \mathbb{O}_{m,n-m}\end{bmatrix}\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m}\end{bmatrix}\mathbf{P}_1\right)^*$$

$$= \left(\mathbf{P}_1^{-1}\left(I_m + \mathbb{O}_m\right)\mathbf{P}_1\right)^*$$

$$= I_m = V_{m,n}\mathbf{Q}_1\begin{bmatrix} I_m \\ \mathbb{O}_{n-m,m}\end{bmatrix}\mathbf{P}_1.$$

□

Theorem 3.2.2 *For the (II) case, the* $\{1,2,4\}$ *-inverse of the rectangular Vandermonde matrix is given by*

$$V_{n,m}^{\{1,2,4\}} = \mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 \in \mathcal{M}_{n,m}, \tag{3.2.15}$$

where the permutated matrices $\mathbf{P}_2$ *and* $\mathbf{Q}_2$ *are given by the expressions (3.2.9) and (3.2.10), respectively.*

Proof. Consider the expression (3.2.8), i.e.

$$\mathbf{P}_2 V_{m,n}\mathbf{Q}_2 = \begin{bmatrix} I_n \\ \mathbb{O}_{n-m,m}\end{bmatrix} \Leftrightarrow V_{m,n} = \mathbf{P}_2^{-1}\begin{bmatrix} I_n \\ \mathbb{O}_{n-m,m}\end{bmatrix}\mathbf{Q}_2^{-1}.$$

In order the matrix $\mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2$ to be the {1, 2, 4}-inverse of $V_{m,n}$, we have to prove the following three equalities,

$$(1)\ V_{m,n}\mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 V_{m,n} = V_{m,n},$$

$$(2)\ \mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 V_{m,n}\mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 = \mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2$$

and

$$(3)\ \left(\mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 V_{m,n}\right)^* = \mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 V_{m,n}.$$

Thus, the (1) holds since

$$
\begin{aligned}
V_{m,n}\mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 V_{m,n} &= \mathbf{P}_2^{-1}\begin{bmatrix} I_n \\ \mathbb{O}_{m-n,n}\end{bmatrix}\mathbf{Q}_2^{-1}\mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2\mathbf{P}_2^{-1}\begin{bmatrix} I_n \\ \mathbb{O}_{m-n,n}\end{bmatrix}\mathbf{Q}_2^{-1} \\
&= \mathbf{P}_2^{-1}\begin{bmatrix} I_n \\ \mathbb{O}_{m-n,n}\end{bmatrix}\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\begin{bmatrix} I_n \\ \mathbb{O}_{m-n,n}\end{bmatrix}\mathbf{Q}_2^{-1} \\
&= \mathbf{P}_2^{-1}\begin{bmatrix} I_n & \mathbb{O}_{n,m-n} \\ \mathbb{O}_{m-n,n} & \mathbb{O}_n\end{bmatrix}\begin{bmatrix} I_n \\ \mathbb{O}_{m-n,n}\end{bmatrix}\mathbf{Q}_2^{-1} \\
&= V_{m,n},
\end{aligned}
$$

and the (2) holds since

$$
\begin{aligned}
&\mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 V_{m,n}\mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 \\
&\qquad = \mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2\mathbf{P}_2^{-1}\begin{bmatrix} I_n \\ \mathbb{O}_{n-m,m}\end{bmatrix}\mathbf{Q}_2^{-1}\mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 \\
&\qquad = \mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\begin{bmatrix} I_n \\ \mathbb{O}_{n-m,m}\end{bmatrix}\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 \\
&\qquad = \mathbf{Q}_2\left(I_n + \mathbb{O}_n\right)\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 \\
&\qquad = \mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2,
\end{aligned}
$$

and finally, the (4) holds since

$$
\begin{aligned}
\left(\mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 V_{m,n}\right)^* &= \left(\mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2\mathbf{P}_2^{-1}\begin{bmatrix} I_n \\ \mathbb{O}_{n-m,m}\end{bmatrix}\mathbf{Q}_2^{-1}\right)^* \\
&= \left(\mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\begin{bmatrix} I_n \\ \mathbb{O}_{n-m,m}\end{bmatrix}\mathbf{Q}_2^{-1}\right)^* \\
&= \left(\mathbf{Q}_2\left(I_n + \mathbb{O}_n\right)\mathbf{Q}_2^{-1}\right)^* \\
&= I_n = \mathbf{Q}_2\begin{bmatrix} I_n & \mathbb{O}_{n,m-n}\end{bmatrix}\mathbf{P}_2 V_{m,n}.
\end{aligned}
$$

□

The (III) case has a very special interest. Here, the Moore-Penrose inverse (see also Remark 3.2.3) is derived. This inverse can be calculated easily, since an analytical formula derives.

Theorem 3.2.3 *For the (III) case, the Moore-Penrose inverse of the square Vandermonde matrix is given by*

$$V_n^{\dagger} = \mathbf{Q}_3\mathbf{P}_3 \in \mathcal{M}_n, \tag{3.2.16}$$

where the permutated matrices $\mathbf{P}_3$ *and* $\mathbf{Q}_3$ *are given by the expressions (3.2.12) and (3.2.13), respectively.*

Proof. Consider the expression (3.2.11), i.e.

$$\mathbf{P}_3 V_n \mathbf{Q}_3 = I_n \Leftrightarrow V_n = \mathbf{P}_3^{-1}\mathbf{Q}_3^{-1}.$$

In order the matrix $\mathbf{Q}_3\mathbf{P}_3$ to be the Moore-Penrose inverse of V_n, we have to prove the following four equalities,

$$(1)\ V_n\mathbf{Q}_3\mathbf{P}_3V_n = V_n,\ (2)\ \mathbf{Q}_3\mathbf{P}_3V_n\mathbf{Q}_3\mathbf{P}_3 = \mathbf{Q}_3\mathbf{P}_3,$$

$$(3)\ \left(\mathbf{Q}_3\mathbf{P}_3V_n\right)^* = \mathbf{Q}_3\mathbf{P}_3V_n \text{ and } (4)\ \left(V_n\mathbf{Q}_3\mathbf{P}_3\right)^* = V_n\mathbf{Q}_3\mathbf{P}_3.$$

Thus, the (1) holds since

$$V_n\mathbf{Q}_3\mathbf{P}_3V_n = \mathbf{P}_3^{-1}\mathbf{Q}_3^{-1}\mathbf{Q}_3\mathbf{P}_3\mathbf{P}_3^{-1}\mathbf{Q}_3^{-1} = \mathbf{P}_3^{-1}\mathbf{Q}_3^{-1} = V_n,$$

the (2) holds since

$$\mathbf{Q}_3\mathbf{P}_3V_n\mathbf{Q}_3\mathbf{P}_3 = \mathbf{Q}_3\mathbf{P}_3\mathbf{P}_3^{-1}\mathbf{Q}_3^{-1}\mathbf{Q}_3\mathbf{P}_3 = \mathbf{Q}_3\mathbf{P}_3,$$

the (3) holds since

$$\left(V_n\mathbf{Q}_3\mathbf{P}_3\right)^* = \left(\mathbf{P}_3^{-1}\mathbf{Q}_3^{-1}\mathbf{Q}_3\mathbf{P}_3\right)^* = I_n = V_n\mathbf{Q}_3\mathbf{P}_3,$$

and finally the (4) holds since

$$\left(\mathbf{Q}_3\mathbf{P}_3V_n\right)^* = \left(\mathbf{Q}_3\mathbf{P}_3\mathbf{P}_3^{-1}\mathbf{Q}_3^{-1}\right)^* = I_n = \mathbf{Q}_3\mathbf{P}_3V_n.$$

□

Remark 3.2.3 Obviously, the More-Penrose inverse (3.2.16) is also the Drazin inverse and the regular inverse V_n^{-1}, see also introduction. □

Example 3.2.2 Suppose that we have the 3×4 - Vandermonde matrix of Example 3.2.1, then the {1, 2, 3}-inverse of the Vandermonde matrix is given by

$$V_{4,3}^{\{1,2,3\}} = \mathbf{Q}_1 \begin{bmatrix} I_3 \\ \underline{0}_3^t \end{bmatrix} \mathbf{P}_1 = \begin{bmatrix} 1 & -3 & 30 & -3000 \\ 0 & 1 & -13 & 1330 \\ 0 & 0 & 1 & -113 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I_3 \\ \underline{0}_3^t \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1/7 & 1/7 & 0 \\ 1/679 & -1/630 & 1/8730 \end{bmatrix}$$

$$= \begin{bmatrix} 1.4728 & -0.4762 & 0.0034 \\ -0.1620 & 0.1635 & -0.0015 \\ 0.0015 & -0.0016 & 0.0001 \\ 0 & 0 & 0 \end{bmatrix}.$$

In the next section, a special matrix is discussed. The results that have been presented here are extended to that special case. This matrix has been recently appeared in an interesting control and system theory problem.

3.3 The Generalized Inverse of a Special Matrix

As we have already discussed extensively in the introduction, in an interesting recent applications of the control and system theory, see Karageorgos, Pantelous and Kalogeropoulos (2009), we need to calculate the generalized inverses of a very special matrix, like

$$\begin{bmatrix} 1 & \mu & \mu^2 & \mu^3 & * & \cdots & * & * & \cdots & \mu^{n-1} \\ 1 & \lambda & \lambda^2 & \lambda^3 & * & \cdots & * & * & \cdots & \lambda^{n-1} \\ 0 & 1 & 2\lambda & 3\lambda^2 & * & \cdots & * & * & \cdots & (n-1)\lambda^{n-2} \\ 0 & 0 & 1 & 3\lambda & * & \cdots & * & * & \cdots & (n-1)(n-2)\lambda^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & * & \cdots & \dfrac{1}{(m-1)}\dfrac{d^{m-1}}{d\lambda_j^{m-1}}\left(\lambda^{n-1}\right) \end{bmatrix} \in \mathcal{M}_{m+1,n},$$

where $\lambda \neq \mu \neq 0$.

In this section, we investigate the rectangular matrix, where $n > m$, using the first row of the Vandermonde matrix, see also introduction. The other two cases (where $n < m$ and $n = m$) can be straightforwardly derived using also the results of 2nd section. So, let assume that we want to investigate the following matrix.

$$S_{m,n} = \begin{bmatrix} 0 & \lambda-\mu & \lambda^2-\mu^2 & \lambda^3-\mu^3 & * & \cdots & * & * & \cdots & \lambda^{n-1}-\mu^{n-1} \\ 0 & 1 & 2\lambda & 3\lambda^2 & * & \cdots & * & * & \cdots & (n-1)\lambda^{n-2} \\ 0 & 0 & 1 & 3\lambda & * & \cdots & * & * & \cdots & (n-1)(n-2)\lambda^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & * & \cdots & \dfrac{1}{(m-1)}\dfrac{d^{m-1}}{d\lambda_j^{m-1}}\left(\lambda^{n-1}\right) \end{bmatrix} \in \mathcal{M}_{m,n}.$$

(3.3.1)

Consequently, as in the previous section, we transport the rectangular special matrix (3.3.1) into the following form,

$$\left[\underline{0}_m \quad I_m \quad \mathbb{O}_{m,n-m-1}\right].$$

Proposition 3.3.1 (*Special Matrix parameterization*)

There are invertible matrices $\mathbf{P}\in\mathcal{M}_m$ *and* $\mathbf{Q}\in\mathcal{M}_n$ *such that*

$$\mathbf{P}S_{m,n}\mathbf{Q}=\left[\underline{0}_m \quad I_m \quad \mathbb{O}_{m,n-m-1}\right], \tag{3.3.2}$$

where the permutated matrices are given analytically by the following expressions (3.3.3) and (3.3.4), i.e.

$$\mathbf{P}=\coprod_{m}^{s=2} P_{m-s+2}\left(\frac{1}{\lambda-\mu}\right)P_{m-s+2}\left(m-s+1,-1\right), \tag{3.3.3}$$

where $P_i\left(a\right)$ *and* $P_i\left(j,a\right)$ *are given by (3.2.1) and (3.2.2) respectively, and*

$$\mathbf{Q}=\prod_{s=2}^{m+1}\prod_{k=s+1}^{n} Q_k\left(s,-\frac{1}{(s-2)!}\frac{d^{s-2}}{d\lambda^{s-2}}\left(\sum_{\substack{k_1,k_2=0\\ k_1+k_2=k-2}}^{k-2}\mu^{k_1}\lambda^{k_2}\right)\right), \tag{3.3.4}$$

where $Q_i\left(j,a\right)$ *is given by (3.2.4).*

Proof. We start with the matrix (3.3.1) where $\lambda\neq\mu\neq 0$.

In this direction, we work as follows

$$P_2\left(1,-1\right)S_{m,n}\to P_2\left(\frac{1}{\lambda-\mu}\right)P_2\left(1,-1\right)S_{m,n}\to P_3\left(2,-1\right)P_2\left(\frac{1}{\lambda-\mu}\right)P_2\left(1,-1\right)S_{m,n}\to P_3\left(\frac{1}{\lambda-\mu}\right)P_3\left(2,-1\right)$$

$$P_2\left(\frac{1}{\lambda-\mu}\right)P_2\left(1,-1\right)S_{m,n}\to\cdots\to$$

$$P_m\left(\frac{1}{\lambda-\mu}\right)P_m\left(m-1,-1\right)P_{m-1}\left(\frac{1}{\lambda-\mu}\right)P_{m-1}\left(m-2,-1\right)\cdots P_3\left(\frac{1}{\lambda-\mu}\right)P_3\left(2,-1\right)P_2\left(\frac{1}{\lambda-\mu}\right)P_2\left(1,-1\right)S_{m,n}.$$

So, the matrix (3.3.1) is transformed to

$$\begin{bmatrix} 0 & 1 & \mu+\lambda & \mu^2+\mu\lambda+\lambda^2 & * & \cdots & * & * & \cdots & \sum\limits_{\substack{k_1,k_2=0\\ k_1+k_2=n-2}}^{n-2} \mu^{k_1}\lambda^{k_2} \\ 0 & 0 & 1 & \frac{d}{d\lambda}\left(\mu^2+\mu\lambda+\lambda^2\right) & * & \cdots & * & * & \cdots & \frac{d}{d\lambda}\sum\limits_{\substack{k_1,k_2=0\\ k_1+k_2=n-2}}^{n-2} \mu^{k_1}\lambda^{k_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & * & \cdots & \frac{1}{(m-1)!}\frac{d^{m-1}}{d\lambda^{m-1}}\sum\limits_{\substack{k_1,k_2=0\\ k_1+k_2=n-2}}^{n-2} \mu^{k_1}\lambda^{k_2} \end{bmatrix}$$

Thus, we conclude to the determination of matrix $\mathbf{P}$

$$\begin{aligned}\mathbf{P} &\triangleq P_m\left(\frac{1}{\lambda-\mu}\right)P_m(m-1,-1)\cdots P_2\left(\frac{1}{\lambda-\mu}\right)P_2(1,-1)\\ &=\prod_{m}^{s=2} P_{m-s+2}\left(\frac{1}{\lambda-\mu}\right)P_{m-s+2}(m-s+1,-1).\end{aligned}$$

Now we want to transfer the $\mathbf{P}S_{m,n}$ into the desired matrix $\left[\underline{0}_m \quad I_m \quad \mathbb{O}_{m,n-m-1}\right]$.

So, we act as follows

$$\mathbf{P}S_{m,n}Q_3\left(2,-\sum_{\substack{k_1,k_2=0\\ k_1+k_2=1}}^{1}\mu^{k_1}\lambda^{k_2}\right)Q_4\left(2,-\sum_{\substack{k_1,k_2=0\\ k_1+k_2=2}}^{2}\mu^{k_1}\lambda^{k_2}\right)\cdots Q_n\left(2,-\sum_{\substack{k_1,k_2=0\\ k_1+k_2=n-2}}^{n-2}\mu^{k_1}\lambda^{k_2}\right)\rightarrow$$

$$\mathbf{P}S_{m,n}Q_3\left(2,-\sum_{\substack{k_1,k_2=0\\ k_1+k_2=1}}^{1}\mu^{k_1}\lambda^{k_2}\right)Q_4\left(2,-\sum_{\substack{k_1,k_2=0\\ k_1+k_2=2}}^{2}\mu^{k_1}\lambda^{k_2}\right)\cdots Q_n\left(2,-\sum_{\substack{k_1,k_2=0\\ k_1+k_2=n-2}}^{n-2}\mu^{k_1}\lambda^{k_2}\right)\cdot$$

$$\cdot Q_3\left(3,-\frac{d}{d\lambda}\sum_{\substack{k_1,k_2=0\\ k_1+k_2=1}}^{2}\mu^{k_1}\lambda^{k_2}\right)\rightarrow\cdots\rightarrow$$

$$\mathbf{P}S_{m,n}Q_3\left(2,-\sum_{\substack{k_1,k_2=0\\ k_1+k_2=1}}^{1}\mu^{k_1}\lambda^{k_2}\right)\rightarrow\mathbf{P}S_{m,n}Q_3\left(2,-\sum_{\substack{k_1,k_2=0\\ k_1+k_2=1}}^{1}\mu^{k_1}\lambda^{k_2}\right)Q_4\left(2,-\sum_{\substack{k_1,k_2=0\\ k_1+k_2=2}}^{2}\mu^{k_1}\lambda^{k_2}\right)$$

$$\rightarrow\ldots\rightarrow$$

$$\mathbf{P}S_{m,n}Q_3\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{1}\mu^{k_1}\lambda^{k_2}\right)\cdots Q_n\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=n-2}}^{n-2}\mu^{k_1}\lambda^{k_2}\right)Q_4\left(3,-\frac{d}{d\lambda}\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{2}\mu^{k_1}\lambda^{k_2}\right)$$

$$\cdots Q_n\left(3,-\frac{d}{d\lambda}\sum_{\substack{k_1,k_2=0\\k_1+k_2=n-2}}^{n-2}\mu^{k_1}\lambda^{k_2}\right)$$

$$\to\ldots\to$$

$$\mathbf{P}S_{m,n}Q_3\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{1}\mu^{k_1}\lambda^{k_2}\right)\cdots Q_n\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=n-2}}^{n-2}\mu^{k_1}\lambda^{k_2}\right)Q_4\left(3,-\frac{d}{d\lambda}\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{2}\mu^{k_1}\lambda^{k_2}\right)$$

$$\cdots Q_n\left(3,-\frac{d}{d\lambda}\sum_{\substack{k_1,k_2=0\\k_1+k_2=n-2}}^{n-2}\mu^{k_1}\lambda^{k_2}\right)$$

$$\cdots Q_5\left(4,-\frac{1}{2!}\frac{d^2}{d\lambda^2}\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{3}\mu^{k_1}\lambda^{k_2}\right)\cdots Q_n\left(2,-\frac{1}{2!}\frac{d^2}{d\lambda^2}\sum_{\substack{k_1,k_2=0\\k_1+k_2=n-2}}^{n-2}\mu^{k_1}\lambda^{k_2}\right)\cdots$$

$$Q_{m+2}\left(m+1,-\frac{1}{(\rho_j-1)!}\frac{d^{m-1}}{d\lambda^{m-1}}\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{m}\mu^{k_1}\lambda^{k_2}\right)\cdots Q_n\left(m+1,-\frac{1}{(m-1)!}\frac{d^{m-1}}{d\lambda^{m-1}}\sum_{\substack{k_1,k_2=0\\k_1+k_2=n-2}}^{n-2}\mu^{k_1}\lambda^{k_2}\right)$$

We also define matrix

$$Q\triangleq Q_3\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{1}\mu^{k_1}\lambda^{k_2}\right)\cdots Q_n\left(2,-\sum_{\substack{k_1,k_2=0\\k_1+k_2=n-2}}^{n-2}\mu^{k_1}\lambda^{k_2}\right)Q_4\left(3,-\frac{d}{d\lambda}\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{2}\mu^{k_1}\lambda^{k_2}\right)\cdots$$

$$Q_n\left(3,-\frac{d}{d\lambda}\sum_{\substack{k_1,k_2=0\\k_1+k_2=n-2}}^{n-2}\mu^{k_1}\lambda^{k_2}\right)\cdots Q_5\left(4,-\frac{1}{2!}\frac{d^2}{d\lambda^2}\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{3}\mu^{k_1}\lambda^{k_2}\right)\cdots Q_n\left(2,-\frac{1}{2!}\frac{d^2}{d\lambda^2}\sum_{\substack{k_1,k_2=0\\k_1+k_2=n-2}}^{n-2}\mu^{k_1}\lambda^{k_2}\right)\cdots$$

$$Q_{m+2}\left(m+1,-\frac{1}{(\rho_j-1)!}\frac{d^{m-1}}{d\lambda^{m-1}}\sum_{\substack{k_1,k_2=0\\k_1+k_2=1}}^{m}\mu^{k_1}\lambda^{k_2}\right)\cdots Q_n\left(m+1,-\frac{1}{(m-1)!}\frac{d^{m-1}}{d\lambda^{m-1}}\sum_{\substack{k_1,k_2=0\\k_1+k_2=n-2}}^{n-2}\mu^{k_1}\lambda^{k_2}\right)$$

$$=\prod_{s=2}^{m+1}\prod_{k=s+1}^{n}P_k\left(s,-\frac{1}{(s-2)!}\frac{d^{s-2}}{d\lambda^{s-2}}\left(\sum_{\substack{k_1,k_2=0\\k_1+k_2=k-2}}^{k-2}\mu^{k_1}\lambda^{k_2}\right)\right)$$

Consequently, we have transposed the special matrix (3.3.1) into $\left[\underline{0}_m \quad I_m \quad \mathbb{O}_{m,n-m-1}\right]$. □

Example 3.3.1 Suppose that we have the 3×4 - special matrix,

$$S_{3,4}=\begin{bmatrix}0 & 10-3 & 10^2-3^2 & 10^3-3^3\\ 0 & 1 & 2\cdot 10 & 3\cdot 10^3\\ 0 & 0 & 1 & 3\cdot 10^3\end{bmatrix}\in \mathcal{M}_{3,4},$$

then by applying (3.3.3), we take

$$\mathbf{P}=\coprod_{3}^{s=2} P_{3-s+2}\left(\frac{1}{7}\right)P_{3-s+2}\left(3-s+1,-1\right)=P_3\left(\frac{1}{7}\right)P_3\left(2,-1\right)P_2\left(\frac{1}{7}\right)P_2\left(1,-1\right)$$
$$=\begin{bmatrix}1 & 0 & 0\\ -1/7 & 1/7 & 0\\ 1/49 & -1/49 & 1/7\end{bmatrix}.$$

The matrix $S_{3,4}$ is being transformed into the following

$$\mathbf{P}S_{3,4}=\begin{bmatrix}0 & 1 & \mu+\lambda & \mu^2+\mu\lambda+\lambda^2\\ 0 & 0 & 1 & \mu+2\lambda\\ 0 & 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}0 & 1 & 3+10 & 3^2+3\cdot 10+10^2\\ 0 & 0 & 1 & 3+2\cdot 10\\ 0 & 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}0 & 1 & 13 & 139\\ 0 & 0 & 1 & 23\\ 0 & 0 & 0 & 1\end{bmatrix}.$$

Now, we want to transfer the $\mathbf{P}S_{3,4}$ into the matrix $\begin{bmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$, so we apply (3.3.4),

i.e.

$$\mathbf{Q}=\prod_{s=2}^{4}\prod_{k=s+1}^{4} Q_k\left(s,-\frac{1}{(s-2)!}\frac{d^{s-2}}{d\lambda^{s-2}}\left(\sum_{\substack{k_1,k_2=0\\ k_1+k_2=k-2}}^{k-2}\mu^{k_1}\lambda^{k_2}\right)\right)=Q_3\left(2,-13\right)Q_3\left(2,-139\right)Q_4\left(3,-23\right)$$
$$=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & -13 & 160\\ 0 & 0 & 1 & -23\\ 0 & 0 & 0 & 1\end{bmatrix}.$$

So, we take the parameterization

$$\mathbf{P}S_{3,4}\mathbf{Q}=\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix},$$

where

$$\mathbf{P}=\begin{bmatrix}1&0&0\\-1/7&1/7&0\\1/49&-1/49&1/7\end{bmatrix}\quad\text{and}\quad\mathbf{Q}=\begin{bmatrix}1&0&0&0\\0&1&-13&160\\0&0&1&-23\\0&0&0&1\end{bmatrix}.$$

In the next lines, the generalized inverse of the rectangular special matrix (3.3.1) is derived.

Theorem 3.3.1 *The* $\{1,2,3\}$*-inverse of the rectangular special matrix (3.3.1) is given by*

$$S_{n,m}^{\{1,2,3\}}=\mathbf{Q}\begin{bmatrix}\underline{0}_m^t\\I_m\\\mathbb{O}_{n-m-1,m}\end{bmatrix}\mathbf{P},\tag{3.3.5}$$

where the permutated matrices $\mathbf{P}$ *and* $\mathbf{Q}$ *are given by the expressions (3.3.3) and (3.3.4), respectively.*

Proof. Consider the expression (3.3.2), i.e.

$$\mathbf{P}S_{m,n}\mathbf{Q}=\begin{bmatrix}\underline{0}_m&I_m&\mathbb{O}_{m,n-m-1}\end{bmatrix}\Leftrightarrow S_{m,n}=\mathbf{P}^{-1}\begin{bmatrix}\underline{0}_m&I_m&\mathbb{O}_{m,n-m-1}\end{bmatrix}\mathbf{Q}^{-1}.$$

In order the matrix $\mathbf{Q}\begin{bmatrix}\underline{0}_m^t\\I_m\\\mathbb{O}_{n-m-1,m}\end{bmatrix}\mathbf{P}$ to be the {1, 2, 3}-inverse of $S_{m,n}$, we have to prove the following three equalities,

$$(1)\quad S_{m,n}\mathbf{Q}\begin{bmatrix}\underline{0}_m^t\\I_m\\\mathbb{O}_{n-m-1,m}\end{bmatrix}\mathbf{P}S_{m,n}=S_{m,n},$$

$$(2)\quad \mathbf{Q}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}S_{m,n}\mathbf{Q}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P} = \mathbf{Q}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}$$

and

$$(3)\quad \left(S_{m,n}\mathbf{Q}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}\right)^* = S_{m,n}\mathbf{Q}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}\,.$$

Thus, the (1) holds since

$$\begin{aligned} S_{m,n}\mathbf{Q}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}S_{m,n} &= \mathbf{P}^{-1}\begin{bmatrix}\underline{0}_m & I_m & \mathbb{O}_{m,n-m-1}\end{bmatrix}\mathbf{Q}^{-1}\mathbf{Q}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}\mathbf{P}^{-1}\begin{bmatrix}\underline{0}_m & I_m & \mathbb{O}_{m,n-m-1}\end{bmatrix}\mathbf{Q}^{-1} \\ &= \mathbf{P}^{-1}\begin{bmatrix}\underline{0}_m & I_m & \mathbb{O}_{m,n-m-1}\end{bmatrix}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\begin{bmatrix}\underline{0}_m & I_m & \mathbb{O}_{m,n-m-1}\end{bmatrix}\mathbf{Q}^{-1} \\ &= \mathbf{P}^{-1}\left(\mathbb{O}_m + I_m + \mathbb{O}_m\right)\begin{bmatrix}\underline{0}_m & I_m & \mathbb{O}_{m,n-m-1}\end{bmatrix}\mathbf{Q}^{-1} \\ &= S_{m,n}, \end{aligned}$$

and the (2) holds since

$$\begin{aligned} \mathbf{Q}\begin{bmatrix} \underline{0}^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}S_{m,n}\mathbf{Q}\begin{bmatrix} \underline{0}^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P} &= \mathbf{Q}\begin{bmatrix} \underline{0}^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}\mathbf{P}^{-1}\begin{bmatrix}\underline{0}_m & I_m & \mathbb{O}_{m,n-m-1}\end{bmatrix}\mathbf{Q}^{-1}\mathbf{Q}\begin{bmatrix} \underline{0}^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P} \\ &= \mathbf{Q}\begin{bmatrix} \underline{0}^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\begin{bmatrix}\underline{0}_m & I_m & \mathbb{O}_{m,n-m-1}\end{bmatrix}\begin{bmatrix} \underline{0}^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P} \end{aligned}$$

$$= \mathbf{Q}\begin{bmatrix} 0 & \underline{0}_m^t & \underline{0}_{n-m-1}^t \\ \underline{0}_m & I_m & \mathbb{O}_{m,n-m-1} \\ \underline{0}_{n-m-1} & \mathbb{O}_{n-m-1,m} & \mathbb{O}_{n-m-1} \end{bmatrix}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}$$

$$= \mathbf{Q}_1\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}_1,$$

and finally, the (3) also holds since

$$\left(S_{m,n}\mathbf{Q}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}\right)^* = \left(\mathbf{P}^{-1}\begin{bmatrix} \underline{0}_m & I_m & \mathbb{O}_{m,n-m-1} \end{bmatrix}\mathbf{Q}^{-1}\mathbf{Q}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}\right)^*$$

$$= \left(\mathbf{P}^{-1}\begin{bmatrix} \underline{0}_m & I_m & \mathbb{O}_{m,n-m-1} \end{bmatrix}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}\right)^*$$

$$= \left(\mathbf{P}^{-1}\left(\mathbb{O}_m + I_m + \mathbb{O}_m\right)\mathbf{P}\right)^*$$

$$= I_m = S_{m,n}\mathbf{Q}\begin{bmatrix} \underline{0}_m^t \\ I_m \\ \mathbb{O}_{n-m-1,m} \end{bmatrix}\mathbf{P}.$$

□

Example 3.3.2 Suppose that we have the 3×4- special matrix of Example 3.3.1, then the {1, 2, 3}-inverse of the special matrix is given by

$$S_{4,3}^{\{1,2,3\}} = \mathbf{Q}\begin{bmatrix} \underline{0}_3^t \\ I_3 \end{bmatrix}\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -13 & 160 \\ 0 & 0 & 1 & -23 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \underline{0}_3^t \\ I_3 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -1/7 & 1/7 & 0 \\ 1/49 & -1/49 & 1/7 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 6.1224 & -5.1224 & 22.8571 \\ -0.6122 & 0.6122 & -3.2857 \\ 0.0204 & -0.0204 & 0.1429 \end{bmatrix}.$$

3.4 Conclusions – Further Research

In the present section, three main results have been proposed: First, we have provided a (quasi) LU factorization, and secondly we have calculated analytically the generalized inverses of the rectangular (and square) Vandermonde matrix, which is defined in terms of scalars $\lambda_1, \lambda_2, ..., \lambda_m \in \mathbb{R}$ (where $m \neq n$) by the following expression:

$$V_{m,n} \equiv V_n(\lambda_1, \lambda_2, ..., \lambda_m) \triangleq \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_m & \cdots & \lambda_m^{n-1} \end{bmatrix}.$$

Finally, similar results with the Vandermonde matrix have been presented for a special structure matrix. Both matrices have been appeared recently in the control and system theory's literature, where the change of the initial state of a linear system in zero time is required.

As a further extension of this chapter,

- we are interested in extending the presenting results to the complex case, where $\lambda_1, \lambda_2, ..., \lambda_m \in \mathbb{C}$.
- Moreover, based on our approach, we want to extend Martinez and Peña (1998b) and Eisinberg, Franzé and Salerno (2001) research works. In the first case, i.e. Eisinberg, Franzé and Salerno (2001), we have a special type of $\lambda_i = \cos\left[\frac{2i-1}{2n}\pi\right]$ for $i = 1, 2, ..., n$ (Chebychev nodes) and
- in the next case, i.e. Martinez and Peña (1998b), we want to calculate the appropriate complete symmetric function, in order to determine the LU factorization of the rectangular Vandermonde matrix.

Chapter 4

Generalized Regular Differential Systems with Distributed Delay

4.1 Introduction

Now days, it is assumed that *Generalized Differential Delay Systems* (GDDSs) provide an excellent mathematical modelling framework for many applications in economical, physical and biological aspects, as well. In many differential models, for instance models for biological population composed of adult and juvenile individuals; it is sometimes meaningless not only have time dependence on the past but also some weighted (distributed) average of previous values on the growth at time t. This has been known for some time, but the theory of such systems with piecewise constant or continuous lagging arguments has been extensively developed only recently.

Our long-term purpose is to study GDDSs within the mainstream of matrix pencil theory. This approach has been extensively used in control theory for the study of generalized linear time invariant dynamical systems without delay, see Gantmacher (1959), Campbell (1980, 1982), Karcanias (1979), Karcanias and Hayton (1981), Van Dooren (1983) and Kalogeropoulos (1985). However, quite recently, in Kalogeropoulos and Stratis (1999) and Wei (2004) a first discussion of generalized differential systems with delay is offered by the matrix pencil and the Drazin inverse matrix theory approach, respectively.

This section is organized as follows: In sub-section 4.2 the necessary preliminary concepts from matrix pencil theory are presented. Sub-section 4.3 contains a brief ac-

count of the required elements of the theory of *Differential Systems with Distributed Delay* (DDDSs). In Sub-section 4.4 the main results of this work are developed. Thus, we investigate the solution of GDDSs with constant coefficients, that means

$$E\underline{x}'(t) = A\int_{t_o}^{t_o+\tau} \underline{x}(t-s)d\mu(s) + B\underline{u}(t), \tag{4.1.1}$$

where $E, A \in \mathbb{C}^{n\times n}$, where $\det E = 0$ and $B \in \mathbb{C}^{n\times l}$ are constant matrices, $\underline{u} \in C\left([t_o,\infty),\mathbb{C}^l\right)$ is a control (column vector function of dimension l), and $t \geq t_o$, where $\tau > 0$ is constant. Furthermore, there exists a unique *normalized bounded variation (**NBV**) function (or distribution)* $\mu:[t_o,t_o+\tau]\to\mathbb{C}$. Moreover, the system (4.1.1) may be reduced to studying a GDDS of the form:

$$F\underline{x}'(t) = G\int_{t_o}^{t_o+\tau} \underline{x}(t-s)\,d\mu(s), \tag{4.1.2}$$

under the common control theory assumption that a state-derivative and continuous delay controller of the following form is obtained:

$$\underline{u}(t) = \tilde{E}\underline{x}'(t) - \tilde{A}\int_{t_o}^{t_o+\tau} \underline{x}(t-s)d\mu(s),$$

when $sF-G$ is a regular pencil the system (4.1.2) is transformed using the Weierstrass canonical decomposition form of the pencil $sF-G$, in two subsystems. One of them is in standard DDS form, while the other is a nilpotent system. This procedure also suggests the form that the initial function should have, such that the corresponding (4.1.1) initial value problem admits a unique solution. Finally, in sub-section 4.5 an illustrative application is presented using MatLab DDE initial value problem solver. Sub-section 4.6 concludes this chapter.

4.2 Mathematical Background from Matrix Pencil Theory

We begin this section by introducing some preliminary concepts and definitions from matrix pencil theory which are used throughout the chapter. Firstly, let there be given the constant matrices F and $G \in \mathbb{C}^{m\times n}$, which uniquely determine the underlying matrix pencil $sF-G$ of system (4.1.2).

Definition 4.2.1 *Given $F,G \in \mathbb{C}^{m\times n}$ and an indeterminate s, the matrix pencil $sF-G$ is called regular when $m=n$ and $\det(sF-G) \neq \mathbb{O}$, where $\mathbb{O}$ is the zero polynomial. In any other case, the pencil will be called singular.* □

In the present section, we focus on regular pencils. Now, let $\mathcal{L}_{n,n}^{r}$ be the set on $n\times n$ regular pencils, i.e.

$$\mathcal{L}_{n,n}^{r} \triangleq \left\{ sF-G : F,G \in \mathbb{C}^{n\times n} \text{ and } sF-G \text{ regular} \right\} \tag{4.2.1}$$

Definition 4.2.2 *The pencil $sF-G \in \mathcal{L}_{n,n}^{r}$ is said to be strictly equivalent to the pencil $sF_1-G_1 \in \mathcal{L}_{n,n}^{r}$ if and only if $P(sF-G)Q = sF_1-G_1$, where $P,\ Q \in \mathbb{C}^{n\times n}$, and $\det P$, $\det Q \neq 0$.* □

The strict equivalence relation can be defined rigorously on the set of regular pencils as follows: Consider the set

$$g \triangleq \left\{ (P,Q) : P,\ Q \in \mathbb{C}^{n\times n},\ \det P,\ \det Q \neq 0 \right\}$$

and a composition rule $*$ defined on g as follows:

$$*: g\times g \to g \text{ such that } (P_1,Q_1)*(P_2,Q_2) \triangleq (P_1\cdot P_2, Q_2\cdot Q_1)$$

It may readily be verified that $(g,*)$ forms a non-abelian group.

Furthermore, an action $\circ$ of the group $(g,*)$ on the set of regular matrix pencils (4.2.1) is defined by

$$\circ : g \times \mathcal{L}_{n,n}^{r} \to \mathcal{L}_{n,n}^{r}$$

such that $\big((P,Q), sF-G\big) \to (P,Q)\circ(sF-G) \triangleq P(sF-G)Q$

This group has the following properties:

(a) $(P_1,Q_1)\circ\big[(P_2,Q_2)\circ(sF-G)\big] = (P_1,Q_1)*(P_2,Q_2)\circ(sF-G)$

for every $P_1, P_2 \in \mathbb{C}^{n\times n}$, $Q_1, Q_2 \in \mathbb{C}^{n\times n}$, $\det P_1$, $\det P_2$, $\det Q_1$, and $\det Q_2 \neq 0$.

(b) $e_g \circ (sF-G) = sF-G$, $sF-G \in \mathcal{L}_{n,n}^{r}$, where $e_g = (I_n, I_n)$ is the identity element of the group $(g,*)$ on the set of $\mathcal{L}_{n,n}^{r}$ defines a transformation group denoted by $\mathcal{M}$, see Kalogeropoulos (1985).

Definition 4.2.3 *For* $sF-G \in \mathcal{L}_{n,n}^{r}$, *the subset*

$$g\circ(sF-G) \triangleq \big\{(P,Q)\circ(sF-G) : (P,Q)\in g\big\} \subseteq \mathcal{L}_{n,n}^{r}$$

will be called the orbit of $sF-G$ *at* g. □

Also $\mathcal{M}$ defines an equivalence relation on $\mathcal{L}_{n,n}^{r}$ which is called a strict equivalence relation and is denoted by $\mathcal{E}_s$. So, $(sF-G)\mathcal{E}_s(sF_1-G_1)$ if and only if $P(sF-G)Q = sF_1-G_1$, where P, $Q \in \mathbb{C}^{n\times n}$ and $\det P$, $\det Q \neq 0$.

The class of $\mathcal{E}_s(sF-G)$ is characterized by a uniquely defined element, known as a *complex* Weierstrass canonical form, $sF_w - Q_w$; see Kalogeropoulos (1985), specified by the complete set of invariants of $\mathcal{E}_s(sF-G)$.

This is the set of (e.d.) obtained by factorizing the invariant polynomials $f_i(s,\hat{s})$ into powers of homogeneous polynomials irreducible over $\mathbb{C}$.

In the case where $sF-G \in \mathcal{L}_{n,n}^{r}$ and $\det F = 0$, we have elementary divisors of the following type:

- *zero elementary divisors* (**z.e.d.**) are those of type s^p;
- *nonzero finite elementary divisors* (**nz. f.e.d.**) are those of type $(s-a)^\pi$, with $a \neq 0$;
- *infinite elementary divisors* (**i.e.d**) are those of type $\hat{s}^q$.

Then, the complex Weierstrass form $sF_w - Q_w$ of the regular pencil $sF-G$, $\det F = 0$ is defined by $sF_w - Q_w \triangleq block\ diag\{sI_p - J_p, sH_q - I_q\}$, where the first normal Jordan type block $sI_p - J_p$ is uniquely defined by the set of f.e.d.

$$(s-a_1)^{p_1},\ldots,(s-a_\nu)^{p_\nu},\ \sum_{j=1}^{\nu} p_j = p$$

of $sF-G$ and has the form

$$sI_p - J_p \triangleq block\ diag\left\{sI_{p_1} - J_{p_1}(a_1),\ldots,sI_{p_\nu} - J_{p_\nu}(a_\nu)\right\},$$

and also the q blocks of the second uniquely defined block $sH_q - I_q$ correspond to the i.e.d.

$$(\hat{s})^{q_1},\ldots,(\hat{s})^{q_\sigma},\ \sum_{j=1}^{\sigma} q_j = q$$

of $sF-G$ and it has the form

$$sH_q - I_q \triangleq block\ diag\left\{sH_{q_1} - I_{q_1},\ldots,sH_{q_\sigma} - I_{q_\sigma}\right\}.$$

Thus the H_q is a nilpotent matrix of index $q^* = \max\left\{q_j : j = 1,2,\ldots,\sigma\right\}$,

where $H_q^{q^*} = \mathbb{O}$

$I_{p_j}, J_{p_j}(a), H_{q_j}$ are the matrices:

$$I_{p_j} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{p_j \times p_j}, \qquad J_{p_j}(a) = \begin{bmatrix} a & 1 & 0 & \cdots & 0 \\ 0 & a & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & a & 1 \\ 0 & 0 & 0 & 0 & a \end{bmatrix} \in \mathbb{C}^{p_j \times p_j}$$

$$H_{q_j} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{q_j \times q_j} \tag{4.2.2}$$

4.3 Delay Differential Equations and Renewal Equations

In this section, we describe briefly the necessary theory of Delay Differential Equations (DDEs). For DDEs we must provide not just the value of the solution at the initial point, but also the "history", i.e. the solution at time prior to the initial point. Thus, the main result is

Theorem 4.3.1 *Consider the system*

$$\underline{x}'(t) = A\int_{t_o}^{t_o+\tau} d\mu(s)\,\underline{x}(t-s)\ t \geq t_o, \tau > 0 \tag{4.3.1}$$

and the initial condition

$$\underline{x}(t) = \underline{\varphi}(t), \tag{4.3.2}$$

for $A \in \mathbb{C}^{n\times n}$ *constant matrix, with NBV function* $\mu:[t_o, t_o+\tau] \to \mathbb{C}^{n\times n}$ *be given, and* $\underline{\varphi} \in C\left([[t_o-\tau, t_o)], \mathbb{C}^n\right)$. *Then there exists a unique function*

$$\underline{x} \in C[t_o-\tau, \infty) \bigcap C^1(t_o, \infty)$$

that satisfies (4.3.1) and (4.3.2). □

The existence and uniqueness may be found in Bellmann and Cooke (1963), Hale (1977), Hale and Verduyn Lunel (1993) and Diekmann et al. (1995).

Definition 4.3.1 *Denote the convolution product by* $*$, *such as* $f * g \in L^1$ *is defined by*

$$f * g \triangleq \int_{t_o}^{t} f(t-\tau)\, g(\tau)\, d\tau$$

where f *is a (possibly* $n \times n$ *matrix - valued)* L^1*-function.* □

Definition 4.3.2 *Equations of the form*

$$y = A(\mu * y) + h,$$

where the kernel μ, the forcing function h are given and y is the unknown parameter, are called (linear) renewal matrix-valued equations or, alternatively, Volterra convolution integral matrix-valued equations (of second kind). □

Now, we reformulate Theorem 4.3.1 taking into consideration the two definitions above. Our strategy is to rewrite the initial-value problem for a linear autonomous delay differential equation as a renewal equation and then to use the resolvent to obtain a representation of the solution.

Theorem 4.3.2 *Let is μ NBV function and $\underline{\varphi} \in C\left([t_o - \tau, t_o)], \mathbb{C}^n\right)$ be given. Define $\underline{g}$ and $\underline{f}$ in terms of μ and $\underline{\varphi}$ by, respectively*

$$\underline{g} = A\left[\mu(t)\underline{\varphi}(t_o) + \int_t^{t_o+\tau} d\mu(s)\underline{\varphi}(t-s)\right]$$

and

$$\underline{f} = \underline{\varphi}(t_o) + A\int_{t_o}^{t_o+\tau}\left[\mu(t+s) - \mu(s)\right]\underline{\varphi}(t_o - s)ds$$

the delay differential system (4.3.1) provided with the initial condition (4.3.2) admits a unique solution on $[t_o - \tau, \infty)$. For $t \geq t_o$ this solution coincides with

$$\underline{x} = A(\mu * \underline{x}) + \underline{f} \tag{4.3.3}$$

whereas the derivative $\underline{x}'$ coincides with the unique solution of the renewal equation

$$\underline{x}' = A(\mu * \underline{x}') + \underline{g}\,.$$ □

The proof may be found in Diekmann et al. (1995).

Definition 4.3.3 *The characteristic matrix $\Delta(z)$ is defined by the expression*

$$\Delta(z) \triangleq zI - A\int_{t_o}^{t_o+\tau} e^{-zt} d\mu(t)\,.$$ □

In the next theorem, we obtain a representation formula using Laplace transformation.

Theorem 4.3.3 *For* μ *NBV function and* $\underline{f} \in C\left(\left[[t_o, \infty)\right], \mathbb{C}^n\right)$ *of bounded variation and constant for* $t \geq \tau$ *the solution* $\underline{x}$ *of the renewal equation (4.3.3) admits for* $t > t_o$ *the representation*

$$\underline{x} = \frac{1}{2\pi i} \int_{L(\gamma)} e^{zt} \Delta(z)^{-1} \left(\underline{f}(t_o) + \int_{t_o}^{\tau} e^{-zs} d\underline{f}(s) \right) dz \tag{4.3.4}$$

for $\gamma > \sup\left\{\operatorname{Re} z : \det \Delta(z) = 0\right\}$, *and* $L(\gamma)$ *denotes the line* $\left\{z \mid \operatorname{Re} z = \gamma\right\}$ *parallel to the imaginary axis in the complex plane. Moreover, by* $\int_{L(\gamma)} \ldots dz$ *we denote the so-called principal value integral* $\lim_{\omega \to \infty} \int_{\gamma - i\omega}^{\gamma + i\omega} \ldots dz$. □

The proof may be also found in Diekmann et al. (1995).

Remark 4.3.1

4.3.1.1 There are several numerical computation methods for the characteristic roots, $\det \Delta(z) = 0$, of linear delay differential equations (4.3.1); see Shampine and Thompson (2001).

4.3.1.2 A detailed study of the asymptotic behaviour of the solution (4.3.4) is also available on Diekmann et al. (1995). □

4.4 Systems of Generalized Linear Differential Equations with Distributed Delay

In this section, we deal with the initial value problem for generalized regular DDSs. These systems of the form

$$E\underline{x}'(t) = A\int_{t_o}^{t_o+\tau} \underline{x}(t-s)\,d\mu(s) + B\underline{u}(t) \quad t \geq t_o,\ \tau > 0 \tag{4.4.1}$$

$$\underline{x}(t) = \underline{\varphi}(t), \qquad t_o - \tau \leq t \leq t_o \tag{4.4.2}$$

where E, $A \in \mathbb{C}^{n\times n}$, with $\det E = 0$, and $B \in \mathbb{C}^{n\times l}$ are constant matrices; the matrix pencil $sE - A$ is supposed to be regular, $\underline{u} \in C\left([t_o,\infty),\mathbb{C}^l\right)$ is a control (column vector function of dimension l), and $t \geq t_o$, where $\tau > 0$ is constant. Furthermore, there exists a unique normalized bounded variation (NBV) function (distribution) $\mu:[t_o, t_o+\tau] \to \mathbb{C}$.

Additionally, let suppose that $\underline{\varphi} \in C^1[t_o - \tau_r, t_o)$ is C^1-differentiable.

Lemma 4.4.1 *The system (4.4.1) may be reduced to studying a generalized linear DDS of the form*

$$F\underline{x}'(t) = G\int_{t_o}^{t_o+\tau} \underline{x}(t-s)\,d\mu(s). \tag{4.4.3}$$

Proof. Assume that the state-derivative and delay feedback controller has the following form

$$\underline{u}(t) = \tilde{E}\underline{x}'(t) - \tilde{A}\int_{t_o}^{t_o+\tau} \underline{x}(t-s)\,d\mu(s), \tag{4.4.4}$$

where $\tilde{E}$, $\tilde{A} \in \mathbb{C}^{l\times n}$ are constant matrices. Then by substituting the above expression into (4.4.1), it is obtained (4.4.3) where

$$F = E - B\tilde{E} \in \mathbb{C}^{n\times n} \text{ and } G = A - B\tilde{A} \in \mathbb{C}^{n\times n}. \qquad \square$$

Definition 4.4.1 *System (4.4.1) is being called normalized if a feedback controller (4.4.4) may be chosen such that its closed-loop (4.4.3) is normal, i.e.*

$$\det F = \det\left(E - B\tilde{E}\right) \neq 0. \qquad \square$$

Moreover, as long as the normalized condition is satisfied, the closed-loop system (4.4.3) would become

$$\underline{x}'(t) = F^{-1}G \int_{t_o}^{t_o+\tau} \underline{x}(t-s)\, d\mu(s) \tag{4.4.5}$$

and its plain feature is its finite poles, i.e. there is not any infinite pole. The solution of the above equation is discussed in Section 4.3 (see also remark 4.4.1).

Lemma 4.4.2 *The system (4.4.1) may be reduced to studying a normalizable linear DDS (4.4.5) if and only if the compound matrix* $C_n\left([E\vdots B]\right) \neq \underline{0}^t$, *where* $C_n\left([E\vdots B]\right) \in \mathbb{C}^{1\times\binom{2n}{n}}$ *is the* n *-order compound matrix of* $[E\vdots B]$. $\square$

The proof may be found in Kytagias (1993), Dai(1989), Kalogeropoulos, Pantelous and Papachristopoulos (2008) research work.

Lemma 4.4.3 *If* $\det E = r < n$, *there exists* $P, Q \in \mathbb{C}^{n\times n}$, $\det P$, $\det Q \neq 0$ *such as*

$$PEQ = \begin{bmatrix} I_r & \mathbb{O}_{r,n-r} \\ \mathbb{O}_{n-r,r} & \mathbb{O}_{n-r,n-r} \end{bmatrix} \qquad \textit{and} \qquad \begin{bmatrix} \tilde{B}_{r,l} \\ \tilde{B}_{n-r,l} \end{bmatrix} \triangleq PB$$

with $\det \tilde{B}_{n-r,l} = n - r$, *then the matrix* $\tilde{E} = -\left[\mathbb{O}_{l,r} \vdots \tilde{B}^t_{l,n-r}\right] Q^{-1}$ *normalizes system (4.4.1).*

Proof. The proof is a direct consequence of lemma 4.4.2. Thus,

suppose that $C_n\left([E\vdots B]\right) \neq \underline{0}^t$, then it should be proved that there exists a matrix $\tilde{E} = -\left[\mathbb{O}_{l,r} \vdots \tilde{B}^t_{l,n-r}\right] Q^{-1}$ such that $\det\left(E - B\tilde{E}\right) \neq 0$.

Analytically, since the $C_n\left(\left[E\vdots B\right]\right)\neq \underline{0}^t$ it is derived that $rank\left[E\vdots B\right]=n$.

Now, since also the $\det E=r<n$ then there exists $P,Q\in\mathbb{C}^{n\times n}$, $\det P$, $\det Q\neq 0$ such as

$$PEQ=\begin{bmatrix} I_r & \mathbb{O}_{r,n-r} \\ \mathbb{O}_{n-r,r} & \mathbb{O}_{n-r,n-r}\end{bmatrix},$$

and $n=rank\left[E\vdots B\right]=rank\left[PEQ\vdots PB\right]=rank\begin{bmatrix} I_r & \mathbb{O}_{r,n-r} & \tilde{B}_{r,l} \\ \mathbb{O}_{n-r,r} & \mathbb{O}_{n-r,n-r} & \tilde{B}_{n-r,l}\end{bmatrix}=r+rank\ \tilde{B}_{n-r,l}$,

Moreover, it is defined $PB\triangleq\begin{bmatrix}\tilde{B}_{r,l}\\ \tilde{B}_{n-r,l}\end{bmatrix}$ and it is clear that the $rank\ \tilde{B}_{n-r,l}=n-r$.

Consequently, the matrix $\tilde{B}_{n-r,l}$ is full row rank.

Now, assume that the matrix $\tilde{B}_{l,n-r}^t\in\mathbb{C}^{l\times(n-r)}$ and define also $X=\tilde{B}_{n-r,l}\tilde{B}_{l,n-r}^t\in\mathbb{C}^{(n-r)\times(n-r)}$ such as

$$\det X=C_{n-r}\left(\tilde{B}_{n-r,l}\right)C_{n-r}\left(\tilde{B}_{l,n-r}^t\right)=C_{n-r}\left(\tilde{B}_{n-r,l}\right)C_{n-r}^t\left(\tilde{B}_{n-r,l}\right)=\left\|C_{n-r}\left(\tilde{B}_{n-r,l}\right)\right\|_2^2>0.$$

Denote the matrix $\tilde{E}=-\left[\mathbb{O}_{l,r}\vdots\tilde{B}_{l,n-r}^t\right]Q^{-1}$, then

$$\begin{aligned}E-B\tilde{E}&=P^{-1}\begin{bmatrix} I_r & \mathbb{O}_{r,n-r} \\ \mathbb{O}_{n-r,r} & \mathbb{O}_{n-r,n-r}\end{bmatrix}Q^{-1}+P^{-1}\begin{bmatrix}\tilde{B}_{r,l}\\ \tilde{B}_{n-r,l}\end{bmatrix}\left[\mathbb{O}_{l,r}\vdots\tilde{B}_{l,n-r}^t\right]Q^{-1}\\ &=P^{-1}\begin{bmatrix} I_r & \mathbb{O}_{r,n-r} \\ \mathbb{O}_{n-r,r} & \mathbb{O}_{n-r,n-r}\end{bmatrix}Q^{-1}+P^{-1}\begin{bmatrix} I_r & \tilde{B}_{r,l}\tilde{B}_{l,n-r}^t \\ \mathbb{O}_{n-r,r} & \tilde{B}_{n-r,l}\tilde{B}_{l,n-r}^t\end{bmatrix}Q^{-1}\\ &=P^{-1}\begin{bmatrix} I_r & \tilde{B}_{r,l}\tilde{B}_{l,n-r}^t \\ \mathbb{O}_{n-r,r} & \tilde{B}_{n-r,l}\tilde{B}_{l,n-r}^t\end{bmatrix}Q^{-1}.\end{aligned}$$

Finally, it is easily derived that

$$\det\left(E-B\tilde{E}\right)=\det P^{-1}\det Q^{-1}\det\begin{bmatrix} I_r & \tilde{B}_{r,l}\tilde{B}_{l,n-r}^t \\ \mathbb{O}_{n-r,r} & \tilde{B}_{n-r,l}\tilde{B}_{l,n-r}^t\end{bmatrix}\neq 0\,.$$ □

In view of Lemmas 4.4.2 and 4.4.3 it is considered, in what follows, systems of the form (4.4.3) where the corresponding matrix pencil $sF-G$ is regular, i.e. $C_n\left(\left[E\vdots B\right]\right)=\underline{0}^t$ and $\det F=\det\left(E-B\tilde{E}\right)=0$.

From the regularity of $sF-G$, there exist non-singular $\mathbb{C}^{n\times n}$ matrices P and Q such that (see also section 4.2)

$$PFQ=F_w=\begin{bmatrix} I_p & \mathbb{O}_{p,q} \\ \mathbb{O}_{q,p} & H_q \end{bmatrix}, \tag{4.4.6}$$

$$PGQ=G_w=\begin{bmatrix} J_p & \mathbb{O}_{p,q} \\ \mathbb{O}_{q,p} & I_q \end{bmatrix}, \tag{4.4.7}$$

where I_p, J_p, H_q are given by (4.2.2).

Proposition 4.4.1 *The system (4.4.3) may be written in the form*

$$\underline{y}_p'(t)=J_p\int_{t_o}^{t_o+\tau}\underline{y}_p(t-s)\,d\mu(s), \tag{4.4.8}$$

$$H_q\underline{y}_q'(t)=\int_{t_o}^{t_o+\tau}\underline{y}_q(t-s)\,d\mu(s). \tag{4.4.9}$$

Proof. Consider the transformation

$$\underline{x}(t)=Q\underline{y}(t). \tag{4.4.10}$$

Substituting the previous expression into (4.4.3), we obtain

$$FQ\underline{y}'(t)=GQ\int_{t_o}^{t_o+\tau}\underline{y}(t-s)\,d\mu(s).$$

Whereby, multiplying by P, we arrive at

$$F_w\underline{y}'(t)=G_w\int_{t_o}^{t_o+\tau}\underline{y}(t-s)\,d\mu(s)$$

Moreover, we can write $\underline{y}(t)$ as $\underline{y}(t)=\begin{bmatrix}\underline{y}_p\\ \underline{y}_q\end{bmatrix}(t)$.

And taking into account the above expressions, we arrive easily at (4.4.7) and (4.4.8).

□

Remark 4.4.1 The system (4.4.7) is in the standard form of systems of linear delay differential equations, and the corresponding initial value problem

$$\underline{y}_p'(t)=J_p\int_{t_o}^{t_o+\tau}\underline{y}_p(t-s)\,d\mu(s),\qquad t\geq t_o,\ \tau>0 \tag{4.4.11}$$

$$\underline{y}_p(t)=\underline{\tilde{\varphi}}_p(t),\qquad t_o-\tau\leq t\leq t_o \tag{4.4.12}$$

may be treated by classical methods (see, section 4.2). Additionally, as it is derived from expression (4.4.12), the initial state function (4.4.2) obtains the following general format:

$$\underline{\tilde{\varphi}}(t)=\begin{bmatrix}\tilde{\varphi}_p(t)\\ \tilde{\varphi}_q(t)\end{bmatrix}=Q^{-1}\underline{\varphi}(t)=Q^{-1}\begin{bmatrix}\varphi_p(t)\\ \varphi_q(t)\end{bmatrix},\ \text{as}\ \det Q\neq 0. \qquad \square$$

Proposition 4.4.2 *The system (4.4.9) has the following solution*

$$\underline{y}_q(t)=\underline{0}_q,\ \text{where,}\ \underline{y}_q(t)=\underline{\tilde{\varphi}}_q(t),\ \text{for}\ t_o-\tau\leq t\leq t_o.$$

Proof. We start by observing that - as it is well known - there exists a $q^*\in\mathbb{N}$ such that

$$H_q^{q^*}=\mathbb{O} \tag{4.4.13}$$

i.e. the q^* is the annihilation index of H_q. We obtain

$$H_q\underline{y}_q'(t)=\int_{t_o}^{t_o+\tau}\underline{y}_q(t-s_1)\,d\mu(s_1) \tag{4.4.14}$$

whereby differentiating, and multiplying by H_q, we get

$$H_q^2 \underline{y}_q''(t) = \int_{t_o}^{t_o+\tau} H_q \underline{y}_q'(t-s_1)\, d\mu(s_1) \tag{4.4.15}$$

and substituting (4.4.14) into (4.4.15)

$$H_q^2 \underline{y}_q''(t) = \int_{t_o}^{t_o+\tau} \int_{t_o}^{t_o+\tau} \underline{y}_q(t-s_1-s_2)\, d\mu(s_2) d\mu(s_1). \tag{4.4.16}$$

By differentiating and multiplying again by H_q the expression (4.4.16) it is obtained

$$H_q^3 \underline{y}_q^{(3)}(t) = \int_{t_o}^{t_o+\tau} \int_{t_o}^{t_o+\tau} \int_{t_o}^{t_o+\tau} \underline{y}_q(t-s_1-s_2-s_3)\, d\mu(s_3) d\mu(s_2) d\mu(s_1)$$

Repeating this argument a sufficient number of times, we end up with

$$H_q^{q^*} \underline{y}_q^{(q^*)}(t) = \int_{t_o}^{t_o+\tau} \int_{t_o}^{t_o+\tau} \cdots \int_{t_o}^{t_o+\tau} \underline{y}_q\left(t-s_1-s_2-s_3 \ldots - s_{q^*}\right) d\mu\left(s_{q^*}\right) \ldots d\mu(s_3) d\mu(s_2) d\mu(s_1).$$

(4.4.17)

Taking into consideration the expression (4.4.13), and all the other similar relations up to and including (4.4.17) as well, we arrive at

$$\int_{t_o}^{t_o+\tau} \int_{t_o}^{t_o+\tau} \cdots \int_{t_o}^{t_o+\tau} \underline{y}_q\left(t-s_1-s_2-s_3-\ldots-s_{q^*}\right) d\mu\left(s_{q^*}\right) \ldots d\mu(s_3) d\mu(s_2) d\mu(s_1) = \underline{0}_q$$

which gives $\underline{y}_q(t) = \underline{0}_q$ with history, $\underline{y}_q(t) = \underline{\tilde{\varphi}}_q(t) \in C[t_o - \tau, t_o)$. □

We conclude this section with the following theorem; its proof follows the proceeding discussion.

Theorem 4.4.1 *The initial value problem for the homogeneous generalized linear regular DDS of the form:*

$$F\underline{x}'(t) = G \int_{t_o}^{t_o+\tau} \underline{x}(t-s)\, d\mu(s) \quad t \geq t_o,\ \tau > 0,$$

and the initial condition $\underline{x}(t) = \underline{\varphi}(t)$, $t_o - \tau \le t \le t_o$

has a unique solution that $\underline{\varphi}(t) = \begin{bmatrix} \underline{\varphi}_p(t) \\ \underline{\mathbb{O}}_q \end{bmatrix} \in C[t_o - \tau, t_o)$ *, and* $p = \sum_{j=1}^{\nu} p_j$ *(i.e. the sum of the degrees of the f.e.d.* $p + q = n$*).* □

4.5 A Numerical Application

In this section, we illustrate the straightforward solution of a GDDS (4.5.1) by computing and plotting the solution. A detailed discussion of the method used by **dde23** (Matlab; m -file) can be found in Shampine and Thompson (2001). Now, suppose that we obtain the following GDDS

$$E\underline{x}'(t) = A\int_0^1 \underline{x}(t-s)e^{-s}ds, \tag{4.5.1}$$

where $E, A \in \mathbb{C}^{4\times 4}$, are constant matrices with $\det E = 0$, and delay period of $-1 < s < 0$. By the associated matrix pencil, $sE - A$, and the results of the 4th subsection, it is supposed that the following two subsystems are obtained

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \int_0^1 y_1(t-s)e^{-s}ds \\ \int_0^1 y_2(t-s)e^{-s}ds \end{bmatrix}, \tag{4.5.2}$$

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_3'(t) \\ y_4'(t) \end{bmatrix} = \begin{bmatrix} \int_0^1 y_3(t-s)e^{-s}ds \\ \int_0^1 y_4(t-s)e^{-s}ds \end{bmatrix}, \tag{4.5.3}$$

which are solved on $t \in [0,10]$ with history; $y_i(s) = 0.1$ for $i = 1,2$ and $-1 \le s \le 0$, respectively.

Firstly, according to the Proposition 4.4.2 the subsystems (4.5.3) has the zero solution

$$\begin{bmatrix} y_3(t) \\ y_4(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Additionally, the system (4.5.2) is reformed to (4.5.4)

$$\begin{cases} y_1'(t) = \int_0^1 e^{-s} y_1(t-s)\,ds + \int_0^1 e^{-s} y_2(t-s)\,ds \\ y_2'(t) = \int_0^1 e^{-s} y_2(t-s)\,ds. \end{cases} \tag{4.5.4}$$

So, we take advantage of the specified history of $y_i(t)$, $i=1,2$ for computing and plotting (figure 4.5.1) the solution of the (4.5.4) system.

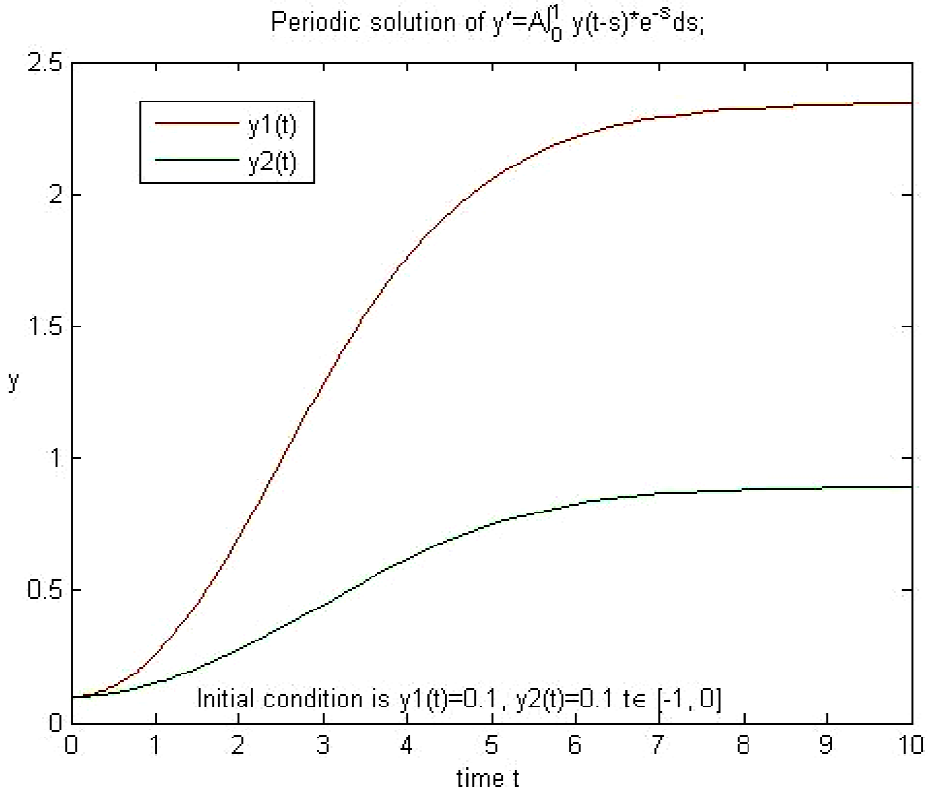

Figure 4.5.1: The plot of the solution of (4.5.4) system into $[0,10]$

4.6 Conclusions - Further Research

In this section, a special class of generalized regular differential delay systems with constant coefficients is extensively studied. In practice, these kinds of systems can model the size of a population or the value of an investment. By considering the regular Matrix Pencil approach, we finally decompose it into two subsystems, whose solutions are obtained. Moreover, since the initial function is given, the corresponding initial value problem is uniquely solvable. Finally, an illustrative application is presented using **dde23** MatLab (m –) file based on the explicit Runge - Kutta method.

As a further extension of the present chapter,

- we want to investigate the special properties of the control input. Thus, several known controllability and stability criteria (see Wei, 2004) can be further extended.
- The introduction of special *normalized bounded variation (**NBV**) functions (or distributions)* $\mu:[t_o,t_o+\tau]\to\mathbb{C}$ is also of a great mathematical interest and importance. In order to be able to investigate the existence and the uniqueness of the solution, some elements of the Functional Analysis are required, see for instance Yosida (1966), Hirch and Lacombe (1999), and Pedersen (2000).
- Additionally, the results of the 2nd chapter can be applied and further extended into such kind of differential systems. Thus, the change of the state in zero time, and the related impulsive behaviour can be combined with the special *normalized bounded variation (**NBV**) functions (or distributions)* $\mu:[t_o,t_o+\tau]\to\mathbb{C}$.

Moreover,

- we want to consider a more general system, see (4.1.2), i.e.

$$E\underline{x}'(t)=A\int_{t_o}^{t_o+\tau}\underline{x}(t-s)d\mu(s)+B\underline{u}(t),$$

where the matrices E and A are rectangular time-invariant coefficients or with a special structure (symmetric, skew symmetric, Toeplitz, non-negative etc). Then some

more special canonical forms, like Kronecker or Tompson etc should be applied. In several applications, see Kalogeropoulos, Karageorgos and Pantelous (2009) and references therein, analytical solutions are also required, where some elements of ODEs and Operator Theory have to be applied.

- secondly, we want to consider the stochastic version of the system (4.1.2), see also 6^{th} section of the present thesis. Under the introduction of irregular inputs, several other aspects can be further discussed.

Chapter 5

On Linear Generalized Neutral Differential Delay Systems

5.1 Introduction

Nowadays, it has been assumed that the class of generalized **n**eutral **d**ifferential **d**elay **s**ystems (**neutral dds**) provides an excellent mathematical modelling framework for numerous applications in natural science and technology. For instance, they are frequently used for the study of distributed networks containing lossless transmission lines; see for instance the comments in Hale (1977). This has been known for some time, but the theory of such systems, with piecewise constant or continuous lagging arguments, has been extensively developed only recently.

In this section, our long-term purpose is to bring the solution's properties for linear generalized neutral dds into the mainstream of matrix pencil theory. This approach has been extensively used in control theory for the study of linear generalized time-invariant dynamical systems without delay; see for more details previous chapter.

The present section is organized as follows: In sub-section 5.2, the necessary preliminary concepts from matrix pencil theory are presented. In sub-section 5.3, the main results of this work are developed. Thus, we investigate the solution of linear generalized neutral dds with constant coefficients, that means

$$E\underline{x}'(t) = A\underline{x}(t) - \sum_{i=1}^{\rho} B_i \underline{x}'(t-\tau_i) + \sum_{i=1}^{\rho} C_i \underline{x}(t-\tau_i) + D\underline{u}(t) \qquad (5.1.1)$$

where, E, A and $B_i, C_i \in \mathbb{C}^{n\times n}$ for $i = 1,2,\ldots,\rho$ are constant matrices, with $\det E = 0$, and the input function $\underline{u} \in C^1[t_o, \infty)$ (column vector function of dimension l) is assumed to consist of all differentiable functions whose derivative is continuous (*continuously differentiable*), and $t \geq t_o$, $0 < \tau_1 < \tau_2 < \ldots < \tau_\rho$ are constants. This kind of delay system is very comon, as Baker, Paul and Willé (1998) claim, due to the fact that constant delay (lag) functions arise frequently in the literature of applications. Furthermore, the system (5.1.1) may be reduced to studying a generalized singular differential system of the form:

$$F\underline{z}'\left(t;\tau_i, i=1,2,\ldots,\rho\right) = G\underline{z}\left(t;\tau_i, i=1,2,\ldots,\rho\right), \tag{5.1.2}$$

where $F, G \in \mathbb{C}^{n\times\rho n}$, $F \triangleq \left[M_\rho \vdots \cdots \vdots M_1 \vdots \tilde{F}\right] \in \mathbb{C}^{n\times\rho n}$, $G \triangleq \left[N_\rho \vdots \cdots \vdots N_1 \vdots \tilde{G}\right] \in \mathbb{C}^{n\times\rho n}$,

$$\underline{z}\left(t;\tau_i, i=1,2,\ldots,\rho\right) \triangleq \begin{bmatrix} \underline{x}\left(t-\tau_\rho\right) \\ \underline{x}\left(t-\tau_{\rho-1}\right) \\ \vdots \\ \underline{x}\left(t-\tau_1\right) \\ \underline{x}\left(t\right) \end{bmatrix}.$$

Under the, usual in control theory, assumption that

$$\underline{u}\left(t\right) = \tilde{E}\underline{x}'\left(t\right) - \tilde{A}\underline{x}\left(t\right) + \sum_{i=1}^{\rho} \tilde{B}_i \underline{x}'\left(t-\tau_i\right) - \sum_{i=1}^{\rho} \tilde{C}_i \underline{x}\left(t-\tau_i\right), \tag{5.1.3}$$

when $sF - G$ is a singular pencil the system (5.1.2) is transformed using the complex Kronecker canonical decomposition form of the pencil $sF - G$ into five subsystems. Finally, in sub-section 5.4 an illustrative application is presented.

5.2 Matrix Pencil Theory Background

This section introduces some preliminary concepts and definitions from Matrix Pencil theory that are been used throughout the chapter. Our result will be restricted to linear generalized neutral dds of the form (5.1.1) where $E, A, D \in \mathbb{C}^{n\times n}$ and $B_i, C_i \in \mathbb{C}^{n\times n}$ for $i = 1, 2, \ldots, \rho$ are time-invariant matrices, with $\det E = 0$, and the input function $\underline{u}(t)$ is assumed to be continuously differentiable. Through this paper we shall adopt the following notation: $\mathbb{R}, \mathbb{C}$ denote the field of real numbers and complex numbers, respectively. $\mathbb{N}$ is the set of natural numbers. If $\mathbb{F}$ is a field, $\mathbb{F}^{m\times n}$ denotes the set of $m \times n$ matrices with elements from $\mathbb{F}$.

Now, if we have $sF - G$, $sF_1 - G_1 \in \mathcal{L}^r_{m,n}$, then $(sF - G)\mathcal{E}_s(sF_1 - G_1)$ if and only if $P(sF - G)Q = sF_1 - G_1$, where $P \in \mathbb{F}^{m\times m}$, $Q \in \mathbb{F}^{n\times n}$ and $\det P$, $\det Q \neq 0$.

The class of $\mathcal{E}_s(sF - G)$ is characterized by a uniquely defined element, known as a complex Kronecker canonical form, $sF_k - Q_k$, see Gantmacher (1959), specified by the complete set of invariants of $\mathcal{E}_s(sF - G)$. Unlike the case of regular pencils, however, the characterization of the $\mathcal{E}_s(sF - G)$, $sF - G \in \mathcal{L}^r_{m,n}$ apart from the set of elementary divisors requires the definition of additional sets of invariants, the minimal indices.

The sets of the minimal degrees $\{v_i,\ 1 \le i \le n - r\}$ and $\{u_j,\ 1 \le j \le m - r\}$ are known by Gantmacher (1959) as column minimal indices (c.m.i.) and row minimal indices (r.m.i.) of $sF - G$, respectively. Furthermore, If $r = rank(sF - G) < \min\{m, n\}$ it is evident that

$$r = \sum_{i=g+1}^{n-r} v_i + \sum_{j=h+1}^{m-r} u_j + rank(sF_w - G_w), \tag{5.2.1}$$

where $sF_w - G_w$ is the complex Weierstrass canonical form specified by the set of elementary divisors (e.d.) obtained by factorizing the invariant polynomials $f(s,\hat{s})$ over $\mathbb{F}[s,\hat{s}]$ (the ring of polynomials in s and $\hat{s}=1/s$ with coefficients on $\mathbb{F}$), which are the nonzero elements of the diagonal of Smith canonical form of the homogeneous pencil $sF-\hat{s}G$, into powers of homogeneous polynomials irreducible over $\mathbb{F}$. In the case where $sF-G$ is a singular pencil, we have elementary divisors of the following type:

- **e.d.** of the type s^d, $d\in\mathbb{N}$, are called zero elementary divisors (**z.e.d.**).
- **e.d.** of the type $(s-a)^c$, $a\neq 0$, $c\in\mathbb{N}$ are called non-zero finite elementary divisors (**nz. f.e.d.**).
- **e.d.** of the type $\hat{s}^q$ are called infinite elementary divisors (**i.e.d**).
- **c.m.i** of the type $\nu\in\mathbb{N}\cup\{0\}$ are called column minimal indices (**c.m.i.**) deduced from the column degrees of minimal polynomial bases of the maximal sub module $\mathcal{M}_N$ embedded in $\mathcal{N}_{right}(s)=\{\underline{x}(s)\in\mathbb{F}^n(s):(sF-G)\underline{x}(s)=\underline{0}\}$ with a free $\mathbb{F}(s)$-module structure.
- **r.m.i** of the type $u\in\mathbb{N}\cup\{0\}$ are called row minimal indices (**r.m.i**.) deduced from the row degrees of minimal polynomial bases of the maximal sub module $\mathcal{M}_N$ embedded in $\mathcal{N}_{left}(s)=\{\underline{\psi}(s)\in\mathbb{F}^m(s):\underline{\psi}^t(sF-G)=\underline{0}^t\}$ with a free $\mathbb{F}(s)$-module structure.

See for more details Gantmacher (1959), Forney (1975), Karcanias (1979), Karcanias and Hayton (1981), Kalogeropoulos (1985) et al. The complex Kronecker form sF_k-Q_k of the singular pencil $sF-G$ is defined.

$$sF_k - Q_k \triangleq block\ diag\{\mathbb{O}_{h,g}, s\Lambda_\nu-\lambda_\nu, s\Lambda_u^t-\lambda_u^t, sI_p-J_p, sH_q-I_q\} \qquad (5.2.2)$$

Analytically, we present the following definition.

Definition 5.2.1 *a) The* $\mathbb{O}_{h,g}$ *is uniquely defined by the sets* $\underbrace{\{0,0,\ldots,0\}}_{g}$ *and* $\underbrace{\{0,0,\ldots,0\}}_{h}$ *of zero, column, and row minimal indices, respectively.*

b) The second normal block $s\Lambda_v - \lambda_v$ *of (5.2.2) is uniquely defined by the set of non-zero column minimal indices (a new arrangement of the indices of* v *must be noted in order to simplify the notation)* $\{v_1 \le v_2 \le \cdots \le v_{n-r-g}\}$ *of* $sF - Q$ *and has the form*

$$s\Lambda_v - \lambda_v \triangleq block\ diag\left\{s\Lambda_{v_1} - \lambda_{v_1}, s\Lambda_{v_2} - \lambda_{v_2}, \ldots, s\Lambda_{v_i} - \lambda_{v_i}, \ldots, s\Lambda_{v_{n-r-g}} - \lambda_{v_{n-r-g}}\right\}, \quad (5.2.3)$$

where $\Lambda_{v_i} = \left[I_{v_i} \vdots \underline{0}\right]$, $\lambda_{v_i} = \left[\underline{0} \vdots I_{v_i}\right]$ *for every* $i = 1,2,\ldots,n-r-g$, *and* I_{v_i} *and* $\underline{0}$ *denote the* $v_i \times v_i$ *identity matrix and the zero column vector, respectively.*

c) The third also normal block $s\Lambda_u^t - \lambda_u^t$ *of (5.2.2) is uniquely determined by the set of non-zero row minimal indices (a new arrangement of the indices of* u *must be noted in order to simplify the notation)* $\{u_1 \le u_2 \le \cdots \le u_{m-r-h}\}$ *of* $sF - G$ *and has the form*

$$s\Lambda_u^t - \lambda_u^t \triangleq block\ diag\left\{s\Lambda_{u_1}^t - \lambda_{u_1}^t, s\Lambda_{u_2}^t - \lambda_{u_2}^t, \ldots, s\Lambda_{u_j}^t - \lambda_{u_j}^t, \ldots, s\Lambda_{u_{m-r-h}}^t - \lambda_{u_{m-r-h}}^t\right\}, \quad (5.2.4)$$

where $\Lambda_{u_j}^t = \begin{bmatrix} I_{u_j} \\ \cdots \\ \underline{0}^t \end{bmatrix}$, $\lambda_{u_j}^t = \begin{bmatrix} \underline{0}^t \\ \cdots \\ I_{u_j} \end{bmatrix}$ *for every* $j = 1,2,\ldots,m-r-h$, *and* I_{u_j} *and* $\underline{0}$ *denote the* $u_j \times u_j$ *identity matrix and the zero column matrix, respectively.*

d) The forth and the fifth normal matrix block of expression (5.2.2) is the complex Weierstrass form $sF_w - Q_w$ *of the regular pencil* $sF - G$, *(i.e.* $\det F = 0$, *if* $n = m$*) is defined by*

$$sF_w - Q_w \triangleq block\ diag\left\{sI_p - J_p, sH_q - I_q\right\}, \quad (5.2.4)$$

where the first normal Jordan type block $sI_p - J_p$ *is uniquely defined by the set of f.e.d.*

$$(s-a_1)^{p_1},\ldots,(s-a_\nu)^{p_\nu},\ \sum_{j=1}^{\nu} p_j = p$$

of $sF-G$ *and has the form*

$$sI_p - J_p \triangleq \text{block diag}\left\{sI_{p_1} - J_{p_1}(a_1),\ldots,sI_{p_\nu} - J_{p_\nu}(a_\nu)\right\}. \tag{5.2.5}$$

And also the q *blocks of the second uniquely defined block* $sH_q - I_q$ *correspond to the i.e.d.*

$$(\hat{s})^{q_1},\ldots,(\hat{s})^{q_\sigma},\ \sum_{j=1}^{\sigma} q_j = q$$

of $sF-G$ *and has the form*

$$sH_q - I_q \triangleq \text{block diag}\left\{sH_{q_1} - I_{q_1},\ldots,sH_{q_\sigma} - I_{q_\sigma}\right\}. \tag{5.2.6}$$

Thus the H_q *is a nilpotent matrix of index* $q^* = \max\left\{q_j : j = 1,2,\ldots,\sigma\right\}$,

where

$$H_q^{q^*} = \mathbb{O}_q, \tag{5.2.7}$$

I_p, $J_p(a)$, H_q *are the matrices:*

$$I_p = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{p\times p},\ J_p(a) = \begin{bmatrix} a & 1 & 0 & \cdots & 0 \\ 0 & a & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & a & 1 \\ 0 & 0 & 0 & 0 & a \end{bmatrix} \in \mathbb{F}^{p\times p},$$

and

$$H_q = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{q\times q}. \tag{5.2.8}$$

5.3 Systems of Linear Generalized Neutral Differential Delay Equations

In this sub-section, we deal with the initial value problem for linear generalized neutral dds. These systems of the form

$$E\underline{x}'(t) = A\underline{x}(t) - \sum_{i=1}^{\rho} B_i \underline{x}'(t-\tau_i) + \sum_{i=1}^{\rho} C_i \underline{x}(t-\tau_i) + D\underline{u}(t), \qquad t > t_o,$$

$$0 < \tau_1 < \tau_2 < \ldots < \tau_\rho \quad (5.3.1)$$

$$\underline{x}(t) = \underline{f}(t), \qquad t_o - \tau_\rho \le t \le t_o \quad (5.3.2)$$

where $E, A \in \mathbb{C}^{n\times n}$, $B_i, C_i \in \mathbb{C}^{n\times n}$ for $i = 1, 2, \ldots, \rho$ and $D \in \mathbb{C}^{l\times n}$ are constant matrices, $\underline{u} \in C[t_o, \infty)$ is a control (column vector function of dimension l), and $t > t_o$, $0 < \tau_1 < \tau_2 < \ldots < \tau_\rho$. Additionally, let suppose that $\underline{f} \in C^1[t_o - \tau_\rho, t_o)$ is continuously differentiable.

Lemma 5.3.1 *The system (5.3.1) may be reduced to studying a linear generalized neutral dds of the form:*

$$\tilde{F}\underline{x}'(t) = \tilde{G}\underline{x}(t) - \sum_{i=1}^{\rho} M_i \underline{x}'(t-\tau_i) + \sum_{i=1}^{\rho} N_i \underline{x}(t-\tau_i). \quad (5.3.3)$$

Proof. Assume that the state-derivative and delay feedback controller has the following type

$$\underline{u}(t) = \tilde{E}\underline{x}'(t) - \tilde{A}\underline{x}(t) + \sum_{i=1}^{\rho} \tilde{B}_i \underline{x}'(t-\tau_i) - \sum_{i=1}^{\rho} \tilde{C}_i \underline{x}(t-\tau_i). \quad (5.3.4)$$

where $\tilde{E}, \tilde{A} \in \mathbb{C}^{l\times n}$, and $\tilde{B}_i, \tilde{C}_i \in \mathbb{C}^{l\times n}$ for $i = 1, 2, \ldots, \rho$ are time invariant matrices. Then by substituting the above expression into (5.3.1), we obtain (5.3.3), where

$$\tilde{F} = E - D\tilde{E} \in \mathbb{C}^{n\times n}, \quad \tilde{G} = A - D\tilde{A} \in \mathbb{C}^{n\times n}, \quad M_i = B_i - D\tilde{B}_i \in \mathbb{C}^{n\times n}$$

$$\text{and } N_i = C_i - D\tilde{C}_i \in \mathbb{C}^{n\times n}, \; i = 1, 2, \ldots, \rho. \qquad \square$$

In the theory of control science, see analogously Dai (1989), the proportional and the derivative feedback may be viewed as the “speed” feedback. In our case, see eq. (5.3.4), we provide to our accelerate feedback controller a memory. In the literature, little is known about (5.3.4). Thus, it should be noticed here that it is planned to investigate this kind of controller more in the near future, since the applications are numerous and very interesting.

Definition 5.3.1 (Kytagias, 1993, Kalogeropoulos, Pantelous and Papachristopoulos, 2008) *System (5.3.1) is called normalized if a feedback controller (5.3.4) may be chosen such that its closed-loop (5.3.3) is normal, i.e.*

$$\det \tilde{F} = \det\left(E - D\tilde{E}\right) \neq 0 \tag{5.3.5}$$

□

Moreover, as long as the normalized condition is satisfied, the closed-loop system (5.3.3) would become

$$\underline{x}'(t) = \tilde{F}^{-1}\tilde{G}\underline{x}(t) - \sum_{i=1}^{\rho} \tilde{F}^{-1}M_i\underline{x}'(t-\tau_i) + \sum_{i=1}^{\rho} \tilde{F}^{-1}N_i\underline{x}(t-\tau_i), \tag{5.3.6}$$

and its plain feature is its finite poles, i.e. it is transformed to a delay system without any infinite pole. The following results present the necessary and sufficient conditions to succeed in transferring our system (3.1) into a normalizable neutral DDS.

Now, we present an alternative approach for treating (3.1). In order to find a solution of the corresponding neutral DDS, some additional results are required.

Lemma 5.3.2 *The system (5.3.3) may be reduced to studying a linear generalized system of the type*

$$F\underline{z}'(t;\tau_i, i=1,2,\ldots,\rho) = G\underline{z}(t;\tau_i, i=1,2,\ldots,\rho), \tag{5.3.7}$$

where $F, G \in \mathbb{C}^{n\times\rho n}$.

Proof. The expression (5.3.3) can be transposed into

$$\tilde{F}\underline{x}'(t)+\sum_{i=1}^{\rho}M_i\underline{x}'(t-\tau_i)=\tilde{G}\underline{x}(t)+\sum_{i=1}^{\rho}N_i\underline{x}(t-\tau_i),$$

or equivalently into

$$\left[M_\rho\vdots\cdots\vdots M_1\vdots\tilde{F}\right]\begin{bmatrix}\underline{x}'(t-\tau_\rho)\\ \underline{x}'(t-\tau_{\rho-1})\\ \vdots\\ \underline{x}'(t-\tau_1)\\ \underline{x}'(t)\end{bmatrix}=\left[N_\rho\vdots\cdots\vdots N_1\vdots\tilde{G}\right]\begin{bmatrix}\underline{x}(t-\tau_\rho)\\ \underline{x}(t-\tau_{\rho-1})\\ \vdots\\ \underline{x}(t-\tau_1)\\ \underline{x}(t)\end{bmatrix}. \tag{5.3.8}$$

Then, if we set

$$F\triangleq\left[M_\rho\vdots\cdots\vdots M_1\vdots\tilde{F}\right]\in\mathbb{C}^{n\times\rho n},\ G\triangleq\left[N_\rho\vdots\cdots\vdots N_1\vdots\tilde{G}\right]\in\mathbb{C}^{n\times\rho n},$$

and

$$\underline{z}(t;\tau_i,i=1,2,\ldots,\rho)\triangleq\begin{bmatrix}\underline{x}(t-\tau_\rho)\\ \underline{x}(t-\tau_{\rho-1})\\ \vdots\\ \underline{x}(t-\tau_1)\\ \underline{x}(t)\end{bmatrix},$$

and substitute the above expressions into (5.3.8), we obtain (5.3.7). □

In view of Lemma 5.3.2, we consider, in what follows, systems of the form (5.3.7) where the corresponding matrix pencil $sF-G$ is singular, as $n\neq\rho n$ (while $\rho>1$).

From the singularity of $sF-G$, there exist non singular matrices $P\in\mathbb{C}^{n\times n}$ and $Q\in\mathbb{C}^{\rho n\times\rho n}$ such that:

$$PFQ=F_k\triangleq block\ diag\left\{\mathbb{O}_{h,g},\Lambda_v,\Lambda_u^t,I_p,H_q\right\} \tag{5.3.10}$$

$$PGQ=G_k\triangleq block\ diag\left\{\mathbb{O}_{h,g},\lambda_v,\lambda_u^t,J_p,I_q\right\}, \tag{5.3.11}$$

where the elements of (5.3.10) and (5.3.11) are well determined in Section 5.2. Additionally, it is profound that $r = rank\left(sF - G\right) < n$, and the expression (5.2.1) can be transposed to

$$rank\left(sF_w - G_w\right) < n - \sum_{i=g+1}^{\rho n-r} v_i - \sum_{j=h+1}^{n-r} u_j < n. \tag{5.3.12}$$

Theorem 5.3.1 *The system (5.3.7) may be decomposed into the equivalent set of subsystems*

$$\mathbb{O}_{h,g}\underline{y}_1'\left(t;\tau\right) = \mathbb{O}_{h,g}\underline{y}_1\left(t;\tau\right), \tag{5.3.13}$$

$$\Lambda_v\underline{y}_v'\left(t\right) = \lambda_v\underline{y}_v\left(t\right), \tag{5.3.14}$$

$$\Lambda_u^t\underline{y}_u'\left(t\right) = \lambda_u^t\underline{y}_u\left(t\right), \tag{5.3.15}$$

$$\underline{y}_p'\left(t\right) = J_p\underline{y}_p\left(t\right), \tag{5.3.16}$$

$$H_q\underline{y}_q'\left(t\right) = \underline{y}_q\left(t\right). \tag{5.3.17}$$

where

$\underline{y}_1\left(t;\tau\right) = \begin{bmatrix} \underline{y}\left(t-\tau_\rho\right) \\ \underline{y}\left(t-\tau_{\rho-1}\right) \\ \vdots \\ \underline{y}_g\left(t\right) \end{bmatrix} \in C^1\left(\left[t_o-\tau_\rho, t_o\right]\right)$, *while we define* $\underline{y}_g\left(t\right)$ *as the* g^{th} *first rows of the* $\underline{y}\left(t\right)$ *vector, under a suitable transformation of* $\underline{z}\left(t\right)$.

Proof. Consider the transformation

$$\underline{x}\left(t;\tau\right) = Q\underline{y}\left(t;\tau\right). \tag{5.3.18}$$

Substituting the previous expression into (5.3.7), we obtain

$$FQ\underline{y}'\left(t;\tau\right) = GQ\underline{y}\left(t;\tau\right).$$

Whereby, multiplying by P, we arrive at

$$F_k \underline{y}'(t;\tau) = G_k \underline{y}(t;\tau).$$

By writing $\underline{y}(t;\tau) = \left[\underline{y}_1^t(t;\tau) \vdots \underline{y}_v^t(t) \vdots \underline{y}_u^t(t) \vdots \underline{y}_p^t(t) \vdots \underline{y}_q^t(t) \right]$ and taking into account the above expressions, we arrive easily at (5.3.13) - (5.3.17). □

In the sequel, the initial value problem is studied corresponding to the subsystems (5.3.13) - (5.3.17) taking into account that

$$\underline{\varphi}(t) = \left[\underline{\varphi}_g^t(t) \ \vdots \ \underline{\varphi}_v^t(t) \ \vdots \ \underline{\varphi}_u^t(t) \ \vdots \ \underline{\varphi}_p^t(t) \ \vdots \ \underline{\varphi}_q^t(t) \ \right]^t \in C\left(\left[t_o - \tau_\rho, t_o \right] \right). \qquad (5.3.19)$$

Additionally, according to expression (3.12), we should always keep in mind that

$$v + u + p + q < n.$$

Proposition 5.3.1 *The initial value problem*

$$\mathbb{O}_{h,g} \underline{y}_1'(t;\tau) = \mathbb{O}_{h,g} \underline{y}_1(t;\tau), \ t > t_o, \ \tau_\rho > \tau_{\rho-1} > \ldots > \tau_{\rho-k} > 0 \qquad (5.3.20)$$

$$\underline{y}_1(t;\tau) = \underline{\varphi}_g(t), \qquad t_o - \tau_\rho < t < t_o \qquad (5.3.21)$$

where $\underline{y}_1(t;\tau) = \begin{bmatrix} \underline{y}(t-\tau_\rho) \\ \underline{y}(t-\tau_{\rho-1}) \\ \vdots \\ \underline{y}_g(t) \end{bmatrix} \in C^1\left(\left[t_o - \tau_\rho, t_o \right] \right)$, *is satisfied for any initial column vector function* $\underline{\varphi}_g(t) \in C^1\left(\left[t_o - \tau_\rho, \infty \right] \right)$ *of* $n(\rho - 1) + g$ *coordinates.*

Proof. The proof is obvious from the fact that the left factors of $\underline{y}_1'(t;\tau)$ and $\underline{y}_1(t;\tau)$ are the $h \times \left[(\rho - 1)n + g \right]$ zero matrices. □

Proposition 5.3.2 *Let* $v_i \in \mathbb{N}$ *be a non-zero column minimal index of the pencil* $sF - G \in \mathcal{L}_{n,\rho n}^s$. *Moreover, let the corresponding typical initial value problem from (5.3.14) be*

$$\Lambda_{v_i}\underline{y}'_{v_i+1}(t)=\lambda_{v_i}\underline{y}_{v_i+1}(t),\ t>t_o \tag{5.3.22}$$

$$\underline{y}_{v_i+1}(t)=\underline{\varphi}_{v_i+1}(t), \tag{5.3.23}$$

with index i *taking values between 1 and* $\rho n-r-(\rho-1)n-g$.

By taking an initial function $y_{v_i+1}:[t_o,\infty)\to\mathbb{R}$ *to be an arbitrary* v_i*-times integrable function over* $[t_o,\infty)$*, the solution is given by*

$$\underline{y}_{v_i+1}(t)=\left[\underbrace{\int_{t_o}^{t}\cdots\int_{t_o}^{t}\int_{t_o}^{t}y_{v_i+1}(s)\,ds}_{v_i}\quad\cdots\quad\int_{t_o}^{t}y_{v_i+1}(s)\,ds\quad y_{v_i+1}(t)\right]^{t}.$$

Proof. By the definition of Λ_{v_i} and λ_{v_i} it follows that the first of the system (5.3.22) - (5.3.23) can be written as

$$\left[I_{v_i}\ \vdots\ \underline{0}\right]\begin{bmatrix}\dot{y}_1(t)\\ \dot{y}_2(t)\\ \vdots\\ \dot{y}_{v_i+1}(t)\end{bmatrix}=\left[\underline{0}\ \vdots\ I_{v_i}\right]\begin{bmatrix}y_1(t)\\ y_2(t)\\ \vdots\\ y_{v_i+1}(t)\end{bmatrix},$$

which is equivalent to

$$\begin{bmatrix}\dot{y}_1(t)\\ \dot{y}_2(t)\\ \vdots\\ \dot{y}_{v_i}(t)\end{bmatrix}=\begin{bmatrix}y_2(t)\\ y_3(t)\\ \vdots\\ y_{v_i+1}(t)\end{bmatrix}. \tag{5.3.24}$$

Such a system is always consistent. If we take $y_{v_i+1}(t)$ to be an arbitrary v_i-times integrable function, then all y_j, $j=1,2,\ldots,v_i$ may be determined by successive integrations of $y_{v_i+1}(t)$ from (5.3.24). It is also clear that $\underline{y}_{v_i+1}(t)$ satisfies the initial value problem. □

Remark 5.3.1 The system (5.3.24) can be written as below

$$\begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \\ \vdots \\ \dot{y}_{v_i}(t) \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}}_{H_{v_i}} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{v_i}(t) \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}}_{e_{v_i}} y_{v_i+1}(t), \tag{5.3.25}$$

where H_{v_i} is a nilpotent matrix of index $v_i \le v_{n-r-g}$, while $v_1 \le v_2 \le \ldots < v_{n-r-g}$.

Furthermore, the unique solution of (5.3.27) is given by the expressions

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{v_i}(t) \end{bmatrix} = e^{H_{v_i}(t-t_o)} \begin{bmatrix} y_1(t_o) \\ y_2(t_o) \\ \vdots \\ y_{v_i}(t_o) \end{bmatrix} + \int_{t_o}^{t} e^{H_{v_i}(t-s)} \underline{e}_{v_i} y_{v_i+1}(s)\, ds$$

while $$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{v_i+1}(t) \end{bmatrix} = \begin{bmatrix} e^{H_{v_i}(t-t_o)} \underline{\varphi}_{v_i}(t_o) + \int_{t_o}^{t} e^{H_{v_i}(t-s)} \underline{e}_{v_i} \varphi_{v_i+1}(s)\, ds \\ \varphi_{v_i+1}(t) \end{bmatrix},$$

taking into consideration that $y_{v_i+1}(t)$ to be an arbitrary v_i-times integrable function, $\forall t \ge t_o$, which satisfies the initial value problem (5.3.22) - (5.3.23). □

Proposition 5.3.3 *Let $u_i \in \mathbb{N}$ be a non-zero column minimal index of the pencil $sF - G \in \mathcal{L}^{s}_{n,\rho n}$. Moreover, let the corresponding typical initial value problem from (5.3.15) be*

$$\Lambda^{t}_{u_i} \underline{\dot{y}}_{u_i}(t) = \lambda^{t}_{u_i} \underline{\dot{y}}_{u_i}(t)\ ,\ t \ge t_o \tag{5.3.26}$$

$$\underline{y}_{u_i}(t) = \underline{\varphi}_{u_i}(t), \tag{5.3.27}$$

with index i taking values between 1 and $n-r-g$, has only the trivial solution.

Proof. By the definition of $\Lambda_{u_i}^t$ and $\lambda_{u_i}^t$ it follows that the first of the system (5.3.26) - (5.3.27) can be written as

$$\begin{bmatrix} I_{u_i} \\ \cdots \\ \underline{0}^t \end{bmatrix} \begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \\ \vdots \\ \dot{y}_{u_i}(t) \end{bmatrix} = \begin{bmatrix} \underline{0}^t \\ \cdots \\ I_{u_i} \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{u_i}(t) \end{bmatrix},$$

which is equivalent to

$$\begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ y_3(t) \\ \vdots \\ y_{u_i}(t) \end{bmatrix}. \tag{5.3.28}$$

whereby we have that $\underline{y}_{u_i}(t) = \underline{0}$, $t \geq t_o$. The result follows by the assumption on the initial value function $\underline{\varphi}_{u_i}(t)$. □

Proposition 5.3.4 *The initial value problem*

$$\underline{y}_p'(t) = J_p \underline{y}_p(t), \quad t > t_o \tag{5.3.29}$$

with $\underline{y}_p(t_o) = \underline{\varphi}_p(t_o) \in \mathbb{C}^p$

has a unique solution

$$\underline{y}_p(t) = e^{J_p(t-t_o)} \underline{\varphi}_p(t_o), \tag{5.3.30}$$

where

$$e^{J_p(t-t_o)} = block\ diag\left\{ e^{J_{p_1}(a_1)(t-t_o)}, e^{J_{p_2}(a_2)(t-t_o)}, \ldots, e^{J_{p_v}(a_v)(t-t_o)} \right\} \ \forall t \geq t_o, \ p = \sum_{i=1}^{v} p_i$$

and

$$e^{J_{p_i}(a_i)(t-t_o)}=\begin{bmatrix} e^{a_i(t-t_o)} & \frac{t}{1!}e^{a_i(t-t_o)} & \cdots & \frac{t^{p_i-1}}{(p_i-1)!}e^{a_i(t-t_o)} \\ 0 & e^{a_i(t-t_o)} & \cdots & \frac{t^{p_i-2}}{(p_i-2)!}e^{a_i(t-t_o)} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & e^{a_i(t-t_o)} \end{bmatrix}\in \mathbb{C}^{p_i\times p_i},\ i=1,2,\ldots,v. \quad \square$$

The proof may be treated by the known classical methods, see Kalogeropoulos (1985), Dai (1989), Grispos (1992) etc.

Preposition 5.3.5 *The initial value problem*

$$H_q\underline{y}_q'(t)=\underline{y}_q(t),\ t>t_o \tag{5.3.31}$$

with

$$\underline{y}_q(t_o)=\underline{\varphi}_q(t_o)\in\mathbb{C}^q$$

has a unique solution

$$\underline{y}_q(t)=\underline{0}_q. \tag{5.3.32}$$

□

The proof may be also found in Kalogeropoulos (1985), Dai (1989), Grispos (1992) etc. As a normal consequence of the above propositions, we can state the following theorem.

Theorem 5.3.3 *The initial value problem for the linear generalized neutral dds of the type*

$$\tilde{F}\underline{x}'(t)=\tilde{G}\underline{x}(t)-\sum_{i=1}^{\rho}M_i\underline{x}'(t-\tau_i)+\sum_{i=1}^{\rho}N_i\underline{x}(t-\tau_i),\ t>t_o,\ 0<\tau_1<\tau_2<\ldots<\tau_r$$

and the initial condition

$$\underline{x}(t)=\underline{f}(t),\ \ t_o-\tau_r\le t\le t_o$$

or equivalently

$$F\underline{z}'(t;\tau_i,i=1,2,\ldots,\rho)=G\underline{z}(t;\tau_i,i=1,2,\ldots,\rho)$$

if we set

$$F \triangleq \left[M_\rho \vdots \cdots \vdots M_1 \vdots \tilde{F} \right] \in \mathbb{C}^{n \times \rho n},\ G \triangleq \left[N_\rho \vdots \cdots \vdots N_1 \vdots \tilde{G} \right] \in \mathbb{C}^{n \times \rho n},$$

and

$$\underline{z}\left(t; \tau_i, i = 1, 2, \ldots, \rho\right) \triangleq \begin{bmatrix} \underline{x}\left(t - \tau_\rho\right) \\ \underline{x}\left(t - \tau_{\rho-1}\right) \\ \vdots \\ \underline{x}\left(t - \tau_1\right) \\ \underline{x}\left(t\right) \end{bmatrix}$$

has solution (given by Proposition 5.3.1 – 5.3.5) provided that the initial conditions are

$$\underline{\varphi}(t) = \left[\underline{\varphi}_g^t(t) \ \vdots \ \underline{\varphi}_v^t(t) \ \vdots \ \underline{\varphi}_u^t(t) \ \vdots \ \underline{\varphi}_p^t(t) \ \vdots \ \underline{\varphi}_q^t(t) \ \right]^t \in C\left(\left[t_o - \tau_\rho, t_o \right] \right). \quad \square$$

5.4 An Illustrative Example

Suppose that the matrix pencil $sF-G\in\mathcal{L}_{10,20}^{s}$ has the following set of invariants:

- f.e.d. : $(s-1)^2$
- i.e.d. : $(\hat{s})^2$
- c.m.i. : $\underbrace{0,0,\ldots,0}_{11},2$
- r. m. i. : 0, 2

To each of these invariants, the corresponding block of the Kronecker canonical form is

$$(s-1)^2\rightarrow sI_2-J_2(1),\qquad (\hat{s})^2\rightarrow sH_2-I_2$$

$$\underbrace{0,0,\ldots,0}_{11}\ \text{c.m.i.}\ \text{and}\ 0\ \text{r.m.i.}\rightarrow\mathbb{O}_{1,11}=\begin{bmatrix}0 & 0 & \cdots & 0\end{bmatrix}$$

$$2\ \text{c.m.i}\quad\rightarrow s\Lambda_2-\lambda_2=s\begin{bmatrix}1 & 0 & \vdots & 0\\ 0 & 1 & \vdots & 0\end{bmatrix}-\begin{bmatrix}0 & \vdots & 1 & 0\\ 0 & \vdots & 0 & 1\end{bmatrix}=\begin{bmatrix}s & -1 & \vdots & 0\\ 0 & s & \vdots & -1\end{bmatrix}$$

$$2\ \text{r.m.i.}\ \rightarrow s\Lambda_2^t-\lambda_2^t=s\begin{bmatrix}1 & 0\\ 0 & 1\\ \cdots & \cdots\\ 0 & 0\end{bmatrix}-\begin{bmatrix}0 & 0\\ \cdots & \cdots\\ 1 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}s & 0\\ -1 & s\\ \cdots & \cdots\\ 0 & -1\end{bmatrix}.$$

Therefore we obtain

$$sF_k-G_k=block\ diag\left\{\mathbb{O}_{1,11},s\Lambda_2,s\Lambda_2^t,sI_2,sH_2\right\}-block\ diag\left\{\mathbb{O}_{1,11},\lambda_2,\lambda_2^t,J_2,I_2\right\}$$

and for $\underline{y}(t;1)=\begin{bmatrix}y_1(t-1) & \ldots & y_{10}(t-1) & y_1(t) & \ldots & y_{10}(t)\end{bmatrix}^t$.

By equation (5.3.16) we obtain the following system

$$\Lambda_2 \underline{y}_2'(t) = \lambda_2 \underline{y}_2(t) \Rightarrow \begin{bmatrix} 1 & 0 & \vdots & 0 \\ 0 & 1 & \vdots & 0 \end{bmatrix} \begin{bmatrix} \dot{y}_2(t) \\ \dot{y}_3(t) \\ \dot{y}_4(t) \end{bmatrix} = \begin{bmatrix} 0 & \vdots & 1 & 0 \\ 0 & \vdots & 0 & 1 \end{bmatrix} \begin{bmatrix} y_2(t) \\ y_3(t) \\ y_4(t) \end{bmatrix} \Rightarrow \begin{bmatrix} \dot{y}_2(t) \\ \dot{y}_3(t) \end{bmatrix} = \begin{bmatrix} y_3(t) \\ y_4(t) \end{bmatrix}$$

According to Proposition 5.3.2, if we obtain that $y_4(t) = f(t)$, which is an arbitrary 2-times integrable function, then $y_3(t) = \int_{t_o}^{t} f(s)\,ds$ and $y_2(t) = \int_{t_o}^{t}\int_{t_o}^{t} f(s)\,ds$, respectively. Additionally, equation (5.3.17) gives

$$\Lambda_2^t \underline{y}_2'(t) = \lambda_2^t \underline{y}_2(t) \Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \cdots & \cdots \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{y}_5(t) \\ \dot{y}_6(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \cdots & \cdots \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_5(t) \\ y_6(t) \end{bmatrix} \Rightarrow \begin{bmatrix} \dot{y}_5(t) \\ \dot{y}_6(t) \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ y_5(t) \\ y_6(t) \end{bmatrix},$$

where the solution is $y_5(t) = y_6(t) = 0 \quad \forall t \geq t_o$.

Furthermore by expressions (5.3.16) and (5.3.17), we obtain

$$\underline{y}_2'(t) = J_2(1)\,\underline{y}_2(t) \Rightarrow \begin{bmatrix} \dot{y}_7(t) \\ \dot{y}_8(t) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_7(t) \\ y_8(t) \end{bmatrix}$$

where $\dot{y}_8(t) = y_8(t) \Rightarrow y_8(t) = e^t y_8(t_o)$,

and $\dot{y}_7(t) = y_7(t) + y_8(t) \Rightarrow \dot{y}_7(t) = y_7(t) + e^t y_8(t_o) \Rightarrow y_7(t) = y_7(t_o)e^{-t} + \frac{1}{2} y_8(t_o) e^t$,

while $y_9(t) = y_{10}(t) = 0 \quad \forall t \geq t_o$.

So, the solution of the linear generalized neutral dds is given by

$$\underline{x}(t;1) = \begin{bmatrix} \underline{x}(t-1) \\ \underline{x}(t) \end{bmatrix} = Q\underline{y}(t;1) =$$

$$[\, f_1(t-1) \ \ldots\ f_{10}(t-1)\ f_1(t)\ \int_{t_o}^{t}\int_{t_o}^{t} f_4(s)\,ds\ \int_{t_o}^{t} f_4(s)\,ds\ f_4(t)\ 0\ 0\ y_8(t_o)\ e^t y_7(t_o)e^{-t} + \frac{1}{2} y_8(t_o) e^t\ 0\ 0]^t$$

5.5 Conclusions – Further Research

In this sub-section, the generalized singular neutral differential delay system with constant coefficients is studied. These kinds of systems are inherent in many physical and engineering phenomena. Using the Matrix Pencil theory we decompose it into five subsystems, whose solutions are obtained. Moreover, the form of the initial function is given, so the corresponding initial value problem is uniquely solvable.

As a further extension of this chapter,

- the (asymptotic) stability testing problem for linear descriptor neutral delay-differential systems of type (5.1.1) will be addressed. By means of the concept of spectral radius, both delay-independent and -dependent stability criteria will be derived, see for further details Yang and Liu (2002).
- These criteria can also be extended to the neutral systems with multiple time delays.
- Finally, we will compare the derived results with the several existing stability criteria, since the stability robustness bounds are expected to get significantly improved, see Yang and Liu (2002). Some examples will be used to show the significance of our results.

Chapter 6

On Generalized Regular Stochastic Differential Delay Systems with Time Invariant Coefficients

6.1 Introduction

In many applications, the systems are considered by the causality that their future states are depende on their past states. Although, this consideration has been known for some time, the relative theory has extensively developed only recently. Additionally, since in many applications it is meaningless not to have any kind and type of perturbation, the introduction of that in delay differential systems increases dramatically the difficulties.

To the best of our knowledge, generalized stochastic delay systems have not been study by the matrix pencil theory approach. Although, the matrix pencil theory has been extensively used in control theory for the study of generalized deterministic dynamical systems without delay, see for instance Gantmacher (1959), Campbell (1980, 1982), Karcanias (1979), Karcanias and Hayton (1981), Van Dooren (1983), Kalogeropoulos (1985) et al.. Moreover, quite recently in Kalogeropoulos and Stratis (1999) and Wei (2004) research works a first discussion of generalized differential systems with delay is offered by matrix pencil and Drazin inverse matrix theory approach, respectively. Additionally, Alabert and Ferrante (2004) consider linear stochastic differential-algebraic systems with additive white noise.

Our long-term purpose is to put generalized linear regular stochastic delay differential systems (SDDSs) into the mainstream of stochastic calculus, developing as far as

possible a theory similar to that of stochastic differential equations. In this chapter, our aim is to investigate the solution of linear SDDEs with constant coefficients and an additive (fractional) white noise, that means

$$E\underline{x}'(t) = A\underline{x}(t) + B\underline{x}(t-\tau) + C\underline{u}(t) + D\underline{f}(t) + R\underline{w}(t), \tag{6.1.1}$$

where $\underline{w}$ is a (fractional) white noise of dimension s, $\underline{f} \in C^{\infty}[t_o, \infty)$ is a smooth input (column vector function of dimension k), and $\underline{u} \in C[t_o, \infty)$ is a control (column vector function of dimension l). The $E, A, B \in \mathbb{C}^{n\times n}$, with $\det E = 0$, $C \in \mathbb{C}^{n\times l}$, $D \in \mathbb{C}^{n\times k}$, and $R \in \mathbb{C}^{n\times s}$ are constant matrices. The system (6.1.1) may be reduced to studying a generalized linear regular SDDS of the form:

$$F\underline{x}'(t) = G\underline{x}(t) + K\underline{x}(t-\tau) + D\underline{f}(t) + R\underline{w}(t), \tag{6.1.2}$$

under the, usual in control theory, assumption that $rank\ C = l$ and suppose that also $\det F = 0$. When $sF - G$ is a regular pencil the system (6.1.2) is transformed using the Weierstrass form canonical decomposition of the pencil $sF - G$, in two subsystems. The irregularity of such noises as are used as inputs makes the solution processes not to be usual, but instead more generalized processes are defined, as the stochastic analogous of Schwartz generalized function.

The present section is organized as follows: Sub-section 6.2 and 6.3 contain a brief account of the required elements of the theory of systems of linear delay differential equations (DDEs), and generalized stochastic processes, respectively. Sub-section 6.3 provides the main results of this work. Finally, in the sub-section 6.4 two particular applications are discussed using Brownian motions (white noise) and fractional Brownian motions (fractional white noise), as the irregular inputs. Sub-section 6.5 concludes the whole discussion.

6.2 Preliminaries on Linear Stochastic Delay Differential Equations

In this sub-section, we briefly describe three topics: the (deterministic) differential-algebraic systems, the main elements of theory of systems of linear delay differential equations (DDEs) and the generalized stochastic processes.

6.2.1 Differential-Algebraic Systems (DASs)

Differential systems are usually used for modelling the dynamical behaviour of many physical and economical phenomena. For example, the conservation laws -such as Kirchhoff- in electrical networks, and the continuous form of the famous in multi sector economy Leontief's model are few of the most known that are consisting both differential and algebraic equations.

The most general form of differential-algebraic systems is

$$F\left(t, \underline{x}, \underline{x}'\right) = 0 \tag{6.2.1}$$

with $F = I \times D_{\underline{x}} \times D_{\underline{x}'} \to \mathbb{C}^m$, where $I \subseteq \mathbb{C}$ is a (compact) interval and $D_{\underline{x}}, D_{\underline{x}'} \in \mathbb{C}^n$ are open, $m, n \in \mathbb{N}$. The meaning of the quantity $\underline{x}'$ is ambiguous as in the case of ordinary differential systems. The reason for this ambiguity is that we want F to determine a differentiable function x that solves (6.2.1) in the sense that

$$F\left(t, \underline{x}(t), \underline{x}'(t)\right) = 0 \text{ for all } t \in I.$$

Particularly, for the linear differential-algebraic systems with time invariant coefficients, and with $\det E = 0$ (see for more details Kunkel and Mehrmann, 2006) it is named generalized differential system

$$E\underline{x}'(t) = A\underline{x}(t) + \underline{f}(t) \tag{6.2.2}$$

where $E, A \in \mathbb{C}^{n\times m}$, and $\underline{f}(t): I \to \mathbb{C}^n$ is sufficiently differentiable. The basic theory has already been established in the nineteenth century by Weierstrass, and Kronecker on matrix pencils (the relevant theory about matrix pencil has already been presented in chapter 3 and 4).

6.2.2 Linear Delay Differential Systems (DDSs)

For linear DDSs, we must provide not just the value of the solution at the initial point, but also the "history", i.e. the solution at times prior to the initial point. Thus, the main result is

Theorem 6.2.1 *Consider the system*

$$\underline{x}'(t) = A\underline{x}(t) + B\underline{x}(t-\tau) + \underline{h}(t), \qquad t > t_o,\ \tau > 0 \tag{6.2.3}$$

and the initial condition

$$\underline{x}(t) = \underline{\phi}(t), \qquad t_o - \tau \le t \le t_o\ , \tag{6.2.4}$$

for $A, B \in \mathbb{C}^{n\times n}$, *constant matrices,* $\underline{h} \in C[t_o - \tau, t_o)$ *(i.e. n-vector valued function) and* $\underline{\phi} \in C[t_o - \tau_r, t_o)$. *Then there exists a unique function*

$$\underline{x} \in C[t_o, \infty) \cap C^1(t_o, \infty) \tag{6.2.5}$$

that satisfies (6.2.3) and (6.2.4). □

The proof may be found in Bellman and Cooke (1963), Elsgolts (1966), Driver (1977), Hale (1977), and Wiener (1993).

Remark 6.2.1

- The function $\det\left(\lambda I - A - Be^{-\lambda\tau}\right)$ is called the characteristic quasi polynomial of (6.2.3), while the equation

$$\det\left(\lambda I - A - Be^{-\lambda\tau}\right) = 0, \tag{6.2.6}$$

is called the characteristic equation of (6.2.3). In general (6.2.6) has infinitely many complex solutions λ.

- The superposition principle is valid; it extends to the case of a series of solutions provided it converges and admits term-by-term differentiation

- Let (6.2.3) be written in the form

$$L\underline{x} = \underline{g} \tag{6.2.7}$$

and let $\underline{x}$ be a solution. Then $\operatorname{Re}\underline{x}$ and $\operatorname{Im}\underline{x}$ are solutions of the equations $L\underline{x} = \operatorname{Re}\underline{g}$ and $L\underline{x} = \operatorname{Im}\underline{g}$, respectively.

- To every root λ_j of (6.2.5) corresponds a particular solution of the form

- λ_j : simple real root $\rightarrow e^{\lambda_j t}$;

- λ_j : complex root $\left(\lambda_j = p_j + iq_i\right) \rightarrow e^{p_j t}\sin q_j t, e^{p_j t}\cos q_j t$;

- λ_j : real root of multiplicity $a_j \rightarrow e^{\lambda_j t}, te^{\lambda_j t}, \ldots, t^{a_j - 1}e^{\lambda_j t}$.

- For a detailed study of the characteristic quasi polynomial and the form of the solutions of (6.2.3), we refer to Elsgolts (1966). □

Remark 6.2.2 In our case, all roots λ_j of (6.2.5) have negative real part. Therefore, by Theorem B, section 28 of Driver (1977), if $\underline{\phi} \in C^{\infty}[t_o - \tau, t_o)$ is bounded, then the solution of (6.2.3) and (6.2.4) is also bounded. □

Remark 6.2.3 In the variation of parameters method, the solution $\underline{x}\left(t; t_o, \underline{\phi}\right)$ is expressed in terms of solutions of the homogeneous equation:

$$\underline{y}'(t) = A\underline{y}(t) + B\underline{y}(t-\tau) \tag{6.2.8}$$

However, firstly we must consider u being the unit step function on $[-\tau, 0]$ (remind that τ is the delay time):

$$u(\theta) \triangleq \begin{cases} 0, & -\tau_r \le \theta < 0 \\ 1, & \theta = 0 \end{cases}$$

Moreover, let $\underline{y}(t;t_o,\underline{\phi})$ be the (unique) solution of the homogeneous problem (6.2.6) and the initial condition

$$\underline{x}(t) = \underline{\phi}(t), \qquad t_o - \tau \le t \le t_o \tag{6.2.7}$$

Then, for $\underline{h} \in C[t_o - \tau, t_o)$, the non-homogeneous system (6.2.1) and (6.2.2) has a unique solution $\underline{x}(t)$ given by

$$\underline{x}(t) = \underline{y}(t;t_o,\underline{\phi}) + \int_{t_o}^{t} \underline{y}(t;s,\underline{h}(s)u)\,ds, \qquad t \ge t_o - \tau \tag{6.2.8}$$

□

Remark 6.2.4

- As far as the form of solution is concerned, we refer to Theorems 6.3-6.6 of Bellman and Cooke (1963).
- The asymptotic behaviour of solutions is studied from Bellman and Cooke (1963) and Hale (1977). □

6.3 Generalized Stochastic (Random) Processes

In this sub-section, the relation between stochastic processes and Schwartz distribution is discussed. It is well known that a random distribution on $I \subset \mathbb{R}$ that is defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ is a measurable mapping $X : (\Omega, \mathcal{F}) \to (\mathcal{D}, \mathcal{B}(\mathcal{D}))$, where $\mathcal{D}$ is the space of distributions (generalized functions) on open set $I \subset \mathbb{R}$, which is the dual of the space $C^{\infty}(I)$ (i.e. that is the smooth real functions with compact support defined on $I \subset \mathbb{R}$) and $\mathcal{B}(\mathcal{D})$ is the Borel σ-field, relative to the strong dual topology (equivalently, the weak topology).

Moreover, we denote

$$\Phi(\varphi) \triangleq \langle X(\omega), \varphi \rangle = \int_{-\infty}^{+\infty} X(\omega)\varphi(\omega)\, d\omega, \tag{6.3.1}$$

to be the action of the distribution $X(\omega) \in \mathcal{D}$ on the test function $\varphi \in C^{\infty}(I)$, which it holds that the mapping $\omega \to \langle X(\omega), \varphi \rangle$ is measurable from $(\Omega, \mathcal{F}) \to (\mathcal{D}, \mathcal{B}(\mathcal{D}))$, hence a real random variable $\langle X(\omega), \varphi \rangle$ is on $(\Omega, \mathcal{F}, \mathcal{P})$.

The product of a real random variable a and a random distribution is defined as

$$\langle aX, \varphi \rangle \triangleq a \langle X, \varphi \rangle \tag{6.3.2}$$

is also a random distribution. Moreover, the derivative of a random distribution is also defined by the expression (6.3.3)

$$\langle X^{(k)}, \varphi \rangle \triangleq (-1)^k \langle X, \varphi^{(k)} \rangle \tag{6.3.3}$$

is again a random distribution. Given a random distribution X, the mapping $C^{\infty}(I) \to L^0(\Omega)$ defined by $\varphi \to \langle X, \varphi \rangle$ is called generalized stochastic process. This mapping is linear and continuous with the usual topology in $C^{\infty}(I)$.

In this section, we will use as the base set $U = [t_o, T],\ 0 < T < \infty$. Further results on random distribution and generalized stochastic processes can be found to the classical papers of Urbanik (1957), Urbanik K. (1958), Gel'fand and Vilenkin (1961), Schwartz (1966), Fernique (1967), Dawson (1970), Kanwal (2004) et al.

6.4 Systems of Generalized Linear Regular Stochastic Delay Differential Equations

In this section, we deal with the initial value problem for generalized linear regular SDD systems. These systems of the form

$$E\underline{x}'(t) = A\underline{x}(t) + B\underline{x}(t-\tau) + C\underline{u}(t) + D\underline{f}(t) + R\underline{w}(t), \quad t > t_o, \tau > 0 \qquad (6.4.1)$$

$$\underline{x}(t) = \underline{\phi}(t), \qquad t_o - \tau \le t \le t_o \qquad (6.4.2)$$

where $\underline{w}$ is a (fractional) white noise of dimension s, $\underline{f} \in C^{\infty}[t_o, \infty)$ is a differentiable input (column vector function of dimension k), and $\underline{u} \in C[t_o, \infty)$ is a control (column vector function of dimension l). The $E, A, B \in \mathbb{C}^{n\times n}$, with $\det E = 0$, $C \in \mathbb{C}^{n\times l}$, $D \in \mathbb{C}^{n\times k}$ and $R \in \mathbb{C}^{n\times s}$ are constant matrices.

Lemma 6.4.1 *The system (6.4.1) may be reduced to studying a generalized linear SDDS of the form:*

$$F\underline{x}'(t) = G\underline{x}(t) + K\underline{x}(t-\tau) + L\underline{f}(t) + R\underline{w}(t) \qquad (6.4.3)$$

Proof. Assume that the state-derivative and delay feedback controller has the following format:

$$\underline{u}(t) = \tilde{E}\underline{x}'(t) - \tilde{A}\underline{x}(t) - \tilde{B}\underline{x}(t-\tau) \qquad (6.4.4)$$

where $\tilde{E}, \tilde{A}, \tilde{B} \in \mathbb{C}^{l\times n}$ are constant matrices. Then by substituting the above expression into (6.4.1), we obtain (6.4.3), where

$$F = E - C\tilde{E} \in \mathbb{C}^{n\times n}, \quad G = A - C\tilde{A} \in \mathbb{C}^{n\times n},$$

$$\text{and } K = B - C\tilde{B} \in \mathbb{C}^{n\times n} \qquad \square$$

As long as the normalized condition is satisfied, the closed-loop system (6.4.3) would become

$$\underline{x}'(t) = F^{-1}G\underline{x}(t) + F^{-1}K\underline{x}(t-\tau) + F^{-1}L\underline{f}(t) + F^{-1}R\underline{w}(t) \qquad (6.4.5)$$

and its plain feature is its finite poles, i.e. there is not any infinite pole. The solution of the above equation is partially discussed in Section 3 (see also Remark 6.4.1).

Lemma 6.4.2 *Without the hypothesis of regularity of matrix pencil* $sF-G$*, the linear stochastic DDSs possess no solution at all.*

Proof. For simplicity, consider the following system without lag

$$F\underline{x}'(t)=G\underline{x}(t)+\underline{f}(t)+\underline{w}(t) \tag{6.4.6}$$

The complex Kronecker form sF_k-Q_k of the singular pencil $sF-Q$ is defined

$$sF_k-Q_k \triangleq block\ diag\left\{\mathbb{O}_{h,g}, s\Lambda_v-\lambda_v, s\Lambda_u^t-\lambda_u^t, sI_p-J_p, sH_q-I_q\right\} \tag{6.4.7}$$

see Forney (1975), Kalogeropoulos (1985) et al.,

or equivalently, there exist non singular matrices $P\in\mathbb{C}^{\rho n\times\rho n}$ and $Q\in\mathbb{C}^{\rho n\times\rho n}$ such that:

$$PFQ=F_k \triangleq block\ diag\left\{\mathbb{O}_{h,g}, \Lambda_v, \Lambda_u^t, I_p, H_q\right\} \tag{6.4.8}$$

$$PGQ=G_k \triangleq block\ diag\left\{\mathbb{O}_{h,g}, \lambda_v, \lambda_u^t, J_p, I_q\right\} \tag{6.4.9}$$

Now, consider the transformation $\underline{x}(t)=Q\underline{y}(t)$. Under that expression, the system (6.4.6) becomes

$$FQ\underline{y}'(t)=GQ\underline{y}(t)+\underline{f}(t)+\underline{w}(t)$$

whereby, multiplying by P, we arrive at

$$PFQ\underline{y}'(t)=PGQ\underline{y}(t)+P\underline{f}(t)+P\underline{w}(t)$$

So, taking into consideration the expressions (6.4.7) - (6.4.9), the differential-algebraic system (6.4.6) may be decomposed in the equivalent set of subsystems

$$\mathbb{O}_{h,g}\underline{y}_g'(t)=\mathbb{O}_{h,g}\underline{y}_g(t)+P_g\underline{f}(t)+P_g\underline{w}(t), \tag{6.4.10}$$

$$\Lambda_v\underline{y}_v'(t)=\lambda_v\underline{y}_v(t)+P_v\underline{f}(t)+P_v\underline{w}(t), \tag{6.4.11}$$

$$\Lambda_u^t \underline{y}_u'(t) = \lambda_u^t \underline{y}_u(t) + P_u \underline{f}(t) + P_u \underline{w}(t), \tag{6.4.12}$$

$$\underline{y}_p'(t) = J_p \underline{y}_p(t) + P_p \underline{f}(t) + P_p \underline{w}(t) \ , \tag{6.4.13}$$

$$H_q \underline{y}_q'(t) = \underline{y}_q(t) + P_q \underline{f}(t) + P_q \underline{w}(t). \tag{6.4.14}$$

Profoundly, the system (6.4.10) has no solution. Consequently, the system (6.4.6) is not solvable. □

For the above result, see also Remark 2.3 in Alabert and Ferrante (2004).

Now, from the regularity of $sF - G$, there exist non-singular $\mathbb{C}^{n\times n}$ matrices P and Q such that (see also Chapter 3).

$$PFQ = F_w = \begin{bmatrix} I_p & \mathbb{O}_{p,q} \\ \mathbb{O}_{q,p} & H_q \end{bmatrix}, \tag{6.4.15}$$

$$PGQ = G_w = \begin{bmatrix} J_p & \mathbb{O}_{p,q} \\ \mathbb{O}_{q,p} & I_q \end{bmatrix}, \tag{6.4.16}$$

where I_p, J_p, H_q are known matrices.

Theorem 6.4.1 *The system (6.4.3) may be written in the form*

$$\underline{y}_p'(t) = J_p \underline{y}_p(t) + (PKQ)_{p,n} \underline{y}(t-\tau) + (PL)_{p,k} \underline{f}(t) + (PR)_{p,m} \underline{w}(t), \tag{6.4.17}$$

$$H_q \underline{y}_q'(t) = \underline{y}_q(t) + (PKQ)_{q,n} \underline{y}(t-\tau) + (PL)_{q,k} \underline{f}(t) + (PR)_{q,m} \underline{w}(t). \tag{6.4.18}$$

Proof. Consider the transformation

$$\underline{x}(t) = Q\underline{y}(t). \tag{6.4.19}$$

Substituting the previous expression into (6.4.3) we obtain

$$FQ\underline{y}'(t) = GQ\underline{y}(t) + KQ\underline{y}(t-\tau) + L\underline{f}(t) + R\underline{w}(t).$$

Whereby, multiplying by P, we arrive at

$$F_w \underline{y}'(t) = G_w Q \underline{y}(t) + PKQ \underline{y}(t-\tau) + PL \underline{f}(t) + PR \underline{w}(t).$$

By writing $\underline{y}(t)$ as

$$\underline{y}(t) = \begin{bmatrix} \underline{y}_p \\ \underline{y}_q \end{bmatrix}(t).$$

And taking into account the above expressions, we arrive easily at (6.4.17) and (6.4.18).

□

Remark 6.4.1 The system (6.4.17) is in the standard form of systems of linear stochastic delay differential equations, and the corresponding initial value problem

$$\underline{y}_p'(t) = J_p \underline{y}_p(t) + (PKQ)_{p,n} \underline{y}(t-\tau) + (PL)_{p,k} \underline{f}(t) + (PR)_{p,m} \underline{w}(t) \qquad t > t_o,\ \tau > 0 \tag{6.4.20}$$

$$\underline{y}_p(t) = \tilde{\underline{\phi}}_p(t), \qquad t_o - \tau \le t \le t_o \tag{6.4.21}$$

which may be treated by classical methods. To find the solution, we shall first solve the equation within the $[0, \tau]$ interval; then, we use this solution process as the initial data to solve the equation within the next $[\tau, 2\tau]$ interval, and so on. Obviously, this procedure allows us to construct a solution step by step, providing at any stage its uniqueness and its regularity, see Mohammed (1984, 1998) et al. □

As a solution to this problem, we shall define a process $\{\underline{y}_p(t), t \in [t_o - \tau, t_o]\}$ and for given smooth test function $\varphi \in C^\infty(\mathcal{U})$,

$$\langle \underline{y}_p(t), \varphi(t) \rangle =$$

$$\left\langle \underline{\phi}_p(0) + \int_{t_o}^{t} J_p \underline{y}_p(s)\,ds + \int_{t_o}^{t} (PKQ)_{p,n} \underline{y}(s-\tau)\,ds + \int_{t_o}^{t} \left[(PL)_{p,k} \underline{f}(s) + (PR)_{p,m} \underline{w}(s) \right] ds, \varphi \right\rangle$$

$$= \int_{t_o}^{T} \left\{ \underline{\phi}_p(0) + \int_{t_o}^{t} J_p \underline{y}_p(s)\,ds + \int_{t_o}^{t} (PKQ)_{p,n} \underline{y}(s-\tau)\,ds + \int_{t_o}^{t} \left[(PL)_{p,k} \underline{f}(s) + (PR)_{p,m} \underline{w}(s) \right] ds \right\} \varphi(t)\,dt \tag{6.4.22}$$

Additionally, as it is derived from expression (6.4.21), the initial state function (6.4.2) obtains the following general format:

$$\underline{\tilde{\phi}}(t)=\begin{bmatrix}\underline{\tilde{\phi}}_p(t)\\ \underline{\tilde{\phi}}_q(t)\end{bmatrix}=Q^{-1}\underline{\phi}(t)=Q^{-1}\begin{bmatrix}\underline{\phi}_p(t)\\ \underline{\phi}_q(t)\end{bmatrix},\ \text{as}\ \det Q\neq 0\,.$$

Theorem 6.4.2 *The solution of system (6.4.6) has the following format*

$$\left\langle \underline{y}_q(t),\varphi(t)\right\rangle=\sum_{k=1}^{q^*-1}(-1)^k\left\langle (PKQ)_{q,n}\,\underline{y}(t-\tau)+(PL)_{q,k}\,\underline{f}(t)+(PR)_{q,m}\,\underline{w}(t),\varphi^{(k)}(t)\right\rangle$$

$$t>t_o,\tau>0 \quad (6.4.23)$$

where $\underline{y}(t)=\underline{\tilde{\phi}}(t)\in C[t_o-\tau,t_o)$, *and test function* $\varphi\in C^{\infty}(\mathcal{U})$.

Proof. We start by observing that -as is well known- there exists a $q^*\in\mathbb{N}$ such that

$$H_q^{q^*}=\mathbb{O} \tag{6.4.24}$$

i.e. the q^* is the annihilation index of H_q.

Setting the test function $\varphi\in C^{\infty}(\mathcal{U})$, and obtaining the generalized process

$$\begin{aligned}\left\langle H_q\underline{y}'_q(t),\varphi(t)\right\rangle&=\left\langle \underline{y}_q(t)+(PKQ)_{q,n}\,\underline{y}(t-\tau)+(PL)_{q,k}\,\underline{f}(t)+(PR)_{q,m}\,\underline{w}(t),\varphi(t)\right\rangle\\&=\left\langle \underline{y}_q(t),\varphi(t)\right\rangle+\left\langle (PKQ)_{q,n}\,\underline{y}(t-\tau)+(PL)_{q,k}\,\underline{f}(t)+(PR)_{q,m}\,\underline{w}(t),\varphi(t)\right\rangle\end{aligned}$$

(6.4.25)

whereby differentiating (in the sense of distributions), and multiplying by H_q, we get

$$\begin{aligned}&\left\langle H_q^2\underline{y}''_q(t),\varphi(t)\right\rangle\\&=\left\langle H_q\underline{y}'_q(t),\varphi(t)\right\rangle-H_q\left\{\left\langle (PKQ)_{q,n}\,\underline{y}(t-\tau)+(PL)_{q,k}\,\underline{f}(t)+(PR)_{q,m}\,\underline{w}(t),\varphi'(t)\right\rangle\right\}\end{aligned}$$

(6.4.26)

and, substituting (6.4.25) into (6.4.26)

$$\left\langle H_q^2 \underline{y}_q''(t), \varphi(t) \right\rangle = \left\langle \underline{y}_q(t), \varphi(t) \right\rangle + \left\langle (PKQ)_{q,n}\, \underline{y}(t-\tau) + (PL)_{q,k}\, \underline{f}(t) + (PR)_{q,m}\, \underline{w}(t), \varphi(t) \right\rangle$$
$$- H_q \left\{ \left\langle (PKQ)_{q,n}\, \underline{y}(t-\tau) + (PL)_{q,k}\, \underline{f}(t) + (PR)_{q,m}\, \underline{w}(t), \varphi'(t) \right\rangle \right\}$$

(6.4.27)

and repeating this argument a sufficient number of times we end up with

$$\left\langle H_q^{q^*} \underline{y}_q^{(q^*)}(t), \varphi(t) \right\rangle = \left\langle H_q^{q^*-1} \underline{y}_q^{(q^*-1)}(t), \varphi(t) \right\rangle +$$
$$(-1)^{q^*-1} H_q^{q^*-1} \left\{ \left\langle (PKQ)_{q,n}\, \underline{y}(t-\tau) + (PL)_{q,k}\, \underline{f}(t) + (PR)_{q,m}\, \underline{w}(t), \varphi^{(q^*-1)}(t) \right\rangle \right\}$$

(6.4.28)

Adding (6.4.18), (6.4.27), and all the other similar relations up to and including (6.4.28), we arrive at

$$\left\langle H_q^{q^*} \underline{y}_q^{(q^*)}(t), \varphi(t) \right\rangle$$
$$= \left\langle \underline{y}_q(t), \varphi(t) \right\rangle - \sum_{k=1}^{q^*-1} (-1)^k \left\langle (PKQ)_{q,n}\, \underline{y}(t-\tau) + (PL)_{q,k}\, \underline{f}(t) + (PR)_{q,m}\, \underline{w}(t), \varphi^{(k)}(t) \right\rangle$$

(6.4.29)

which, by (6.4.24), gives

$$\left\langle \underline{y}_q(t), \varphi(t) \right\rangle = \sum_{k=1}^{q^*-1} (-1)^k \left\langle (PKQ)_{q,n}\, \underline{y}(t-\tau) + (PL)_{q,k}\, \underline{f}(t) + (PR)_{q,m}\, \underline{w}(t), \varphi^{(k)}(t) \right\rangle,$$
$$t > t_o, \tau > 0 \qquad (6.4.30)$$

with history, $\underline{y}(t) = \underline{\tilde{\phi}}(t) \in C[t_o - \tau, t_o)$. □

We conclude this section with the following theorem; its proof follows the proceeding discussion.

Theorem 6.4.3 *The initial value problem for the homogeneous generalized linear regular SDDS of the form:*

$$F\underline{x}'(t) = G\underline{x}(t) + K\underline{x}(t-\tau) + L\underline{f}(t) + R\underline{w}(t) \quad t > t_o,\ \tau > 0$$

and the initial condition

$$\underline{x}(t) = \underline{\phi}(t), \qquad t_o - \tau \le t \le t_o$$

has a unique solution provided that $\underline{\phi}(t) = \begin{bmatrix} \underline{\phi}_p(t) \\ \underline{\phi}_q(t) \end{bmatrix} \in C[t_o - \tau_r, t_o)$, *test function* $\varphi \in C^{\infty}(\mathcal{U})$, *and* $p = \sum_{j=1}^{\nu} p_j$ *(i.e. the sum of the degrees of the f.e.d.* $p + q = n$*).* □

6.5 The Main Results with Respect to Certain Type of Noises

6.5.1 Brownian Motion (or White Noise)

In this sub-section we will use white noise on $\mathbb{R}_+$ coincides with the Wiener integral with respect to the standard Brownian motion (sBm), $\{W(t), t \geq 0\}$, on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$. Moreover, if $\varphi \in C^{\infty}(U)$ is as a test function, then

$$\langle \xi, \varphi \rangle = \int_{t_o}^{T} \varphi(s)\, dW(s) \tag{6.5.1.1}$$

in the sense of equality in law. More precisely, the Wiener integral is defined as the extension to $L^2(\mathbb{R}_+)$ of white noise, see Kuo (1975) and Borodin and Salminen (2002) for more details about the construction of the Wiener integral as the extension of white noise. Now, integrating by parts in (6.5.1.1), we can write

$$\langle \xi', \varphi \rangle = -\int_{t_o}^{T} W(s)\varphi'(s)\, ds = -\langle W, \varphi' \rangle. \tag{6.5.1.2}$$

Thus, the ξ is the derivative of the Brownian motion W as random distributions. A random distribution is Gaussian if every finite-dimensional projection is a Gaussian random vector. This is the case of white noise and Brownian motion.

In that particular case, since the $W(t)$ is a s-dimensional standard Wiener process, the expressions (6.4.22) and (6.4.30) can be expressed as follows:

$$\langle \underline{y}_p(t), \varphi(t) \rangle =$$
$$\int_{t_o}^{T} \left\{ \underline{\phi}_p(0) + J_p \int_{t_o}^{t} \underline{y}_p(s)\, ds + (PKQ)_{p,n} \int_{t_o}^{t} \underline{y}(s-\tau)\, ds + (PL)_{p,k} \int_{t_o}^{t} \underline{f}(s)\, ds + (PR)_{p,m} \int_{t_o}^{t} dW(s) \right\} \varphi(t)\, dt, \tag{6.5.1.3}$$

and

$$\left\langle \underline{y}_q(t), \varphi(t) \right\rangle =$$

$$\sum_{k=1}^{q^*-1} (-1)^k \left[(PKQ)_{q,n} \int_0^\infty \underline{y}(t-\tau)\varphi^{(k)}(t)\,dt + (PL)_{q,k} \int_0^\infty \underline{f}(t)\varphi^{(k)}(t)\,dt + (PR)_{q,m} \int_0^\infty \varphi^{(k)}(t)\,dW(s) \right]. \tag{6.5.1.4}$$

6.5.2 Fractional Brownian Motion (or Fractional White Noise)

Due to their important applications, the fBm have been studied by many authors in recent years, as a consequence several kinds of stochastic calculus have been developed, see Descreusefond and Üstünel (1995), Duncan, Hu and Pasik – Duncan (2000), Alòs, Mazet and Nualart (2001), Hu and Øksendal (2003), Yan and Mohammed (2005) etc.

In this subsection, we show a way to adapt the traditional white noise calculus to the fractional white noise case. Firstly, we recall that if $\{W(t), t \geq 0\}$ is a standard Brownian motion (sBm) on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$, then it is defined

$$W^{\mathcal{H}}(t) = \int_{t_o}^{T} Z_{\mathcal{H}}(t,s)\,dW(s),\ t \geq 0 \tag{6.5.2.1}$$

which is the representation of fBm of Hurst parameter $\mathcal{H} \in (0,1)$ on the same probability space (see Hu, 2005, for more details) , where

$$Z_{\mathcal{H}}(t,s) = \begin{cases} k_{\mathcal{H}} \left[\left(\frac{t}{s}\right)^{\mathcal{H}-\frac{1}{2}} (t-s)^{\mathcal{H}-\frac{1}{2}} - \left(\mathcal{H}-\frac{1}{2}\right) s^{\frac{1}{2}-\mathcal{H}} \int_s^t u^{\mathcal{H}-\frac{3}{2}} (u-s)^{\mathcal{H}-\frac{1}{2}}\,du \right], & \text{if } 0 < \mathcal{H} < 1/2 \\ \left(\mathcal{H}-\frac{1}{2}\right) k_H s^{\frac{1}{2}-\mathcal{H}} \int_s^t u^{\mathcal{H}-\frac{1}{2}} (u-s)^{\mathcal{H}-\frac{3}{2}}\,du, & \text{if } 1/2 < \mathcal{H} < 1 \end{cases}$$

also $k_{\mathcal{H}} = \sqrt{\dfrac{2\mathcal{H}\Gamma\left(\frac{3}{2}-\mathcal{H}\right)}{\Gamma\left(\mathcal{H}+\frac{1}{2}\right)\Gamma(2-2\mathcal{H})}}$, $\Gamma(a) = \int_o^\infty s^{a-1}e^{-s}ds$ is the gamma function (6.5.2.2)

Now, if we formally differentiate (6.5.2.1) with respect to t, then we obtain the following heuristic equality

$$\dot{W}^{\mathcal{H}}(t)=\frac{d}{dt}\int_{t_o}^{t}Z_{\mathcal{H}}(t,s)dW(s)=\frac{d}{dt}\int_{t_o}^{t}Z_{\mathcal{H}}(t,s)\dot{W}(s)ds \tag{6.5.2.3}$$

Thus, the above equation (6.5.2.3) suggests that formally $\mathrm{I}\Gamma_{\mathcal{H}}$ is a transformation which transforms the white noise (the derivative of sBm) to fractional noise (the derivative of fBm), where

$$\mathrm{I}\Gamma_{\mathcal{H}}\{g(t)\}=\frac{d}{dt}\int_{t_o}^{t}Z_{H}(t,s)g(s)ds,\ t\geq 0 \tag{6.5.2.4}$$

and

$$\mathrm{I}\Gamma_{\mathcal{H}}\dot{W}=\dot{W}^{\mathcal{H}} \tag{6.5.2.5}$$

if $\varphi\in C^{\infty}(\mathcal{U})$ is as test function, then

$$\begin{aligned}\langle\xi,\varphi\rangle&=\int_{t_o}^{T}\varphi(s)\dot{W}^{H}(s)ds=\int_{t_o}^{T}\varphi(s)\mathrm{I}\Gamma_{\mathcal{H}}(\dot{W})(s)ds\\&=\int_{t_o}^{T}(\mathrm{I}\Gamma_{\mathcal{H}}^{*}\varphi)(s)\dot{W}(s)ds=\int_{t_o}^{T}(\mathrm{I}\Gamma_{\mathcal{H}}^{*}\varphi)(s)dW(s)\end{aligned} \tag{6.5.2.6}$$

in the sense of equality in law, where for $f\in C^{\infty}(\mathcal{U})$

$$\mathrm{I}\Gamma_{\mathcal{H}}^{*}f(s)=-k_{\mathcal{H}}s^{\frac{1}{2}-\mathcal{H}}\frac{d}{ds}\int_{s}^{T}t^{\mathcal{H}-\frac{1}{2}}(t-s)^{\mathcal{H}-\frac{1}{2}}f(s)dt,\ s\geq 0. \tag{6.5.2.7}$$

More details about the construction of the fractional white noise see Hu (2005). Now, integrating by parts in (6.5.2.6), we can write

$$\langle\xi',\varphi\rangle=-\int_{t_o}^{T}W(s)(\mathrm{I}\Gamma_{\mathcal{H}}^{*}\varphi)'(s)ds=-\left\langle W,(\mathrm{I}\Gamma_{\mathcal{H}}^{*}\varphi)'\right\rangle. \tag{6.5.2.8}$$

Now, the ξ is the derivative of the fBm $W^{\mathcal{H}}$ as random distributions.

In that particular case, where $W(t)$ is a s-dimensional standard Wiener process. The expressions (6.4.22) and (6.4.30) can be expressed as follows:

$$\left\langle \underline{y}_p(t), \varphi(t) \right\rangle =$$
$$\int_{t_o}^{T} \left\{ \underline{\phi}_p(0) + J_p \int_{t_o}^{t} \underline{y}_p(s)\,ds + (PKQ)_{p,n} \int_{t_o}^{t} \underline{y}(s-\tau)\,ds + (PL)_{p,k} \int_{t_o}^{t} \underline{f}(s)\,ds + (PR)_{p,m} \int_{t_o}^{t} dW(s) \right\} \left(\Pi\Gamma_{\mathcal{H}}^{*}\varphi\right)(t)\,dt \tag{6.5.1.3}$$

and

$$\left\langle \underline{y}_q(t), \varphi(t) \right\rangle = \sum_{k=1}^{q^*-1} (-1)^k \left[\begin{array}{l} (PKQ)_{q,n} \int_{t_o}^{T} \underline{y}(t-\tau) \left(\Pi\Gamma_{\mathcal{H}}^{*}\varphi\right)^{(k)}(t)\,dt \\ + (PL)_{q,k} \int_{t_o}^{T} \underline{f}(t) \left(\Pi\Gamma_{\mathcal{H}}^{*}\varphi\right)^{(k)}(t)\,dt + (PR)_{q,m} \int_{t_o}^{T} \left(\Pi\Gamma_{\mathcal{H}}^{*}\varphi\right)^{(k)}(t)\,dW(t) \end{array} \right] \tag{6.5.1.4}$$

A very good reference for Malliavin calculus and stochastic delay equations is Bell and Mohammed (1991).

Remark 6.5.1 As we observe from the subsections 6.5.1 and 6.5.2, we can also consider other stochastic noises, since the only requirement is to can define a Wiener integral with respect to such noise. □

6.6 Conclusion – Further Research

In this chapter, we consider the generalized linear regular stochastic differential delay system with constant coefficients and two simultaneous external differentiable and non differentiable perturbations. These kinds of systems are inherent in many application fields; among them we mention fluid dynamics, the modelling of multi body mechanisms, finance and the problem of protein folding. Using regular Matrix Pencil theory, we decompose it into two subsystems, whose solutions are obtained as generalized processes (in the sense of distributions). Moreover, the form of the initial function is given, so the corresponding initial value problem is uniquely solvable. Finally, two illustrative applications are presented using white noise and fractional white noise, respectively.

The results of the 6th section can be further extended into several interesting directions.

- First, as it has already been discussed in the 4th chapter, we want to investigate the special properties of the control input. Thus, several known controllability and stability criteria (see for instance Klamka and Socha, 1977, Zabczyk, 1981, Ehrhard and Kliemann, 1982, Mahmudov, 2001 etc) can be further extended. Furthermore, the derived results can be transferred into the special cases of standard and fractional Brownian motions.

- These criteria can also be extended to the stochastic differential systems with multiple time delays and different kind of irregular noises-processes (for instance, we can use some special Lévy and Jump processes, which have several applications into Actuarial/Financial science).

- Additionally, as it has been already proposed in the 4th chapter, the results of the 2nd chapter can be applied and further extended into such kind of differential stochastic systems. Thus, the change of the state in zero time, and the related impulsive behaviour can be also combined with the introduction of special *normalized bounded variation (**NBV**) functions (or distributions)* $\mu:[t_o,t_o+\tau]\to\mathbb{C}$, i.e.

$$E\underline{x}'(t) = A\underline{x}(t) + B\int_{t_o}^{t_o+\tau} \underline{x}(t-s)d\mu(s) + C\underline{u}(t) + D\underline{f}(t) + R\underline{w}(t).$$

Finally,

- we want to consider a more general system, see (4.1.2), i.e.

$$E\underline{x}'(t) = A\underline{x}(t) + B\underline{x}(t-\tau) + C\underline{u}(t) + D\underline{f}(t) + R\underline{w}(t),$$

where the matrices E and A are time-invariant coefficients with a special structure (symmetric, skew symmetric, Toeplitz, non-negative etc). Then some more special canonical forms, like Tompson etc should be applied. In several applications, see Kalogeropoulos, Karageorgos and Pantelous (2009) and references therein, analytical solutions are also required, where some elements of ODEs and Operator Theory have to be applied.

- secondly, we want to investigate the approximation of the solutions, see (6.4.22) and (6.4.23). Consequently, the derived -practical useful- results will be used in several significant applications in Actuarial and Financial science.

Chapter 7

Conclusions – Further Research

In this chapter, we want to conclude and present the numerous, basic extensions of the present PhD thesis. Analytically,

A) distributional solutions and behaviour enter the study of many areas in systems and control such as:

(i) Controllability, Observability,

(ii) Infinite zero characteristic behaviour,

(iii) Almost invariant subspaces, almost controllability spaces,

(iv) Dynamics of singular systems, etc.

The distributional characterization is also linked to solution of a number of control problems. Although such solutions have theoretical significance, their value is limited from the practical (implementation of solutions) viewpoint, since impulses represent distributions and cannot be constructed. Only functions can.

Thus, we want to develop a theory for approximating distributions with different families of smooth functions. This should involve defining also a ‘metric’ to measure how good the approximation is and ways to parameterise these families. For instance, if we use the Gauss function and its derivatives, then we may parameterise the families in terns of σ. Can we link σ to the distance from the distribution? Can we associate σ to the corresponding energy? 2-norm (Euclidean norm)? What are advantages / disadvantages of the different approximations?

Now, consider the problem of transferring the origin of a controllable system to any point within a hyper sphere of radius $\mathbb{R}$. We know that this can be done in 0 – time with impulses. What is the minimal time required for achieving this if we use an approximation to the distributional solution, by using a specifying families? If these are restrictions on the energy of the input signal, can we achieve this transfer within the $\mathbb{R}$-sphere? If yes, what is the required time?

Clearly similar problems can be defined for the dual problem of reconstructibility.

Impulsive solution of implicit system descriptions of the:

Pencil type: $F\dot{\underline{x}} = G\underline{x}$,

Autoregressive type: $T(p)\underline{x} = 0$.

Here we have to clarify the fundamental system motion of the significance of the approximation.

One way of handling this may be to pose the question: ‘is there another system for which the approximation of the distribution is an autonomous smooth solution?’ Difficult question that needs a lot of thinking. Essentially, we ask for ‘system deformations’ that express natural correspondence between solutions.

Given that autonomous solutions are expressed as exponentials, the link between exponentials and distributions is worth examining.

Consider the problems of almost (A,B)- invariance and almost Controllability subspaces. In the first case we use distributions to help trajectories in the subspace. If we use approximations, how close can we keep the trajectories to the subspace? If we use Gauss approximations, what is the link of σ and the distance from the subspace? What happens if we impose conditions on the energy of the signal? What is the effect on the distance? Repeat the same for the almost controllability case.

Moreover, if $\boldsymbol{\mathcal{V}}$ is an almost (A,B)- invariant, or almost Controllability subspace, can we define spaces for the same system $\tilde{\boldsymbol{\mathcal{V}}}$ which are (A,B)- invariant, controllability subspaces? Which are those which are closest to them? Can we relate this to approximation of distributions?

Infinite zero solutions. Output zeroing problem for distributions, characterize the infinite zero structure. Define approximate smooth output willing solutions.

In the next lines, we will present some related questions

PROBLEM (A1): Distance between differential systems descriptions.

Consider two pencil models: $F\dot{\underline{x}} = G\underline{x}$ and $F'\dot{\underline{x}}' = G'\underline{x}'$ where (F,G), (F',G') are pairs of the same dimension.

(i) Define distance functions between (F,G), (F',G') pairs

(ii) Investigate relations between kronecker structures of $sF-G$, $sF'-G'$ as a function of the distance.

(iii) Topology of (F,G) pairs and spectra, indices, Plücker invariants properties.

Extension from pencils to polynomial models

$$A(p)\underline{x} = 0,$$

where $$A(p) = p^m A_m + p^{m-1}A_{m-1} + \ldots + A_o.$$

PROBLEM (A2): Define distance functions between distributions and differential families of functions.

B) As a further extension of the 3rd chapter, we are interested in extending the presenting results to the complex case, where $\lambda_1, \lambda_2, ..., \lambda_m \in \mathbb{C}$.

Moreover, based on our approach, we want to extend Martinez and Peña (1998b) and Eisinberg, Franzé and Salerno (2001) research works. In the first case, i.e. Eisinberg, Franzé and Salerno (2001), we have a special type of $\lambda_i = \cos\left[\frac{2i-1}{2n}\pi\right]$ for $i = 1, 2, ..., n$ (Chebychev nodes) and in the next case, i.e. Martinez and Peña (1998b), we

want to calculate the appropriate complete symmetric function, in order to determine the LU factorization of the rectangular Vandermonde matrix.

C) As a further extension of the 4th chapter, we want to investigate the special properties of the control input. Thus, several known controllability and stability criteria (see Wei, 2004) can be further extended.

The introduction of special *normalized bounded variation (**NBV**) functions (or distributions)* $\mu:[t_o,t_o+\tau]\to\mathbb{C}$ is also of a great mathematical interest and importance. In order to be able to investigate the existence and the uniqueness of the solution, some elements of the Functional Analysis are required, see for instance Yosida (1966), Hirch and Lacombe (1999), and Pedersen (2000).

Additionally, the results of the 2nd chapter can be applied and further extended into such kind of differential systems. Thus, the change of the state in zero time, and the related impulsive behaviour can be combined with the special *normalized bounded variation (**NBV**) functions (or distributions)* $\mu:[t_o,t_o+\tau]\to\mathbb{C}$.

Moreover, we want to consider a more general system, see (4.1.2), i.e.

$$E\underline{x}'(t)=A\int_{t_o}^{t_o+\tau}\underline{x}(t-s)d\mu(s)+B\underline{u}(t),$$

where the matrices E and A are rectangular time-invariant coefficients or with a special structure (symmetric, skew symmetric, Toeplitz, non-negative etc). Then some

more special canonical forms, like Kronecker or Tompson etc should be applied. In several applications, see Kalogeropoulos, Karageorgos and Pantelous (2009) and references therein, analytical solutions are also required, where some elements of ODEs and Operator Theory have to be applied.

Finally, we want to consider the stochastic version of the system (4.1.2), see also 6th section of the present thesis. Under the introduction of irregular inputs, several other aspects can be further discussed.

D) As a further extension of the 5th chapter, the (asymptotic) stability testing problem for linear descriptor neutral delay-differential systems of type (5.1.1) will be addressed. By means of the concept of spectral radius, both delay-independent and -dependent stability criteria will be derived, see for further details Yang and Liu (2002).

These criteria can also be extended to the neutral systems with multiple time delays.

Finally, we will compare the derived results with the several existing stability criteria, since the stability robustness bounds are expected to get significantly improved, see Yang and Liu (2002). Some examples will be used to show the significance of our results.

E) The results of the 6th section can be further extended into several interesting directions.

First, as it has already been discussed in the 4th chapter, we want to investigate the special properties of the control input. Thus, several known controllability and stability criteria (see for instance Klamka and Socha, 1977, Zabczyk, 1981, Ehrhard and Kliemann, 1982, Mahmudov, 2001 etc) can be further extended. Furthermore, the derived results can be transferred into the special cases of standard and fractional Brownian motions.

These criteria can also be extended to the stochastic differential systems with multiple time delays and different kind of irregular noises-processes (for instance, we can use some special Lévy and Jump processes, which have several applications into Actuarial/Financial science).

Additionally, as it has been already proposed in the 4th chapter, the results of the 2nd chapter can be applied and further extended into such kind of differential stochastic systems. Thus, the change of the state in zero time, and the related impulsive behaviour can be also combined with the introduction of special *normalized bounded variation (**NBV**) functions (or distributions)* $\mu:[t_o,t_o+\tau]\to\mathbb{C}$, i.e.

$$E\underline{x}'(t)=A\underline{x}(t)+B\int_{t_o}^{t_o+\tau}\underline{x}(t-s)d\mu(s)+C\underline{u}(t)+D\underline{f}(t)+R\underline{w}(t)$$

Finally, we want to consider a more general system, see (4.1.2), i.e.

$$E\underline{x}'(t)=A\underline{x}(t)+B\underline{x}(t-\tau)+C\underline{u}(t)+D\underline{f}(t)+R\underline{w}(t),$$

where the matrices E and A are time-invariant coefficients with a special structure (symmetric, skew symmetric, Toeplitz, non-negative etc). Then some more special ca-

nonical forms, like Tompson etc should be applied. In several applications, see Kalogeropoulos, Karageorgos and Pantelous (2009) and references therein, analytical solutions are also required, where some elements of ODEs and Operator Theory have to be applied.

Finally, we want to investigate the approximation of the solutions, see (6.4.22) and (6.4.23). Consequently, the derived -practical useful- results will be used in several significant applications in Actuarial and Financial science.

In this part of the PhD thesis, we want to emphasize that many very interesting and significant issues are still open. Some preliminary work has been done, but much more is needed.

References

Alabert, A. and Ferrante, M. (2006), Linear stochastic differential-algebraic equations with constant coefficients, *Electronic Communications in Probability* Vol. **11**, pp. 316 - 335.

Alòs, E., Mazet, O. and Nualart, D. (2001), Stochastic calculus with respect to fractional Brownian motion with Hurst parameter lesser than ½, *Stochastic Processes and their Applications* Vol. **86**, pp. 121 - 139.

Antsaklis, P. and Michel, A (2009), Linear Systems, Springer, India, Private Ltd.

Baker, C.T.H., Paul, C.A.H. and Willé, D.R. (1995), *A bibliography on numerical solution of delay differential equations*, Numer. Anal. Rept. No. 269, Maths. Dept., Univ. of Manchester, UK.

Bell, D.R. and Mohammed, S.E.A. (1991), The Malliavin calculus and stochastic delay equations, *Journal of Functional Analysis* Vol. **99**, No. 1, pp. 75 - 99.

Bellmann, R. and Cooke, K.L. (1963), *Differential - Difference Equations*, Academic Press, New York.

Björck, A. and Pereyra, V. (1970), Solution of Vandermonde system of equations, *Mathematics of Computation* Vol. **24**, pp. 893 - 903.

Borodin, N.A. and Salminen P. (2002), *Handbook of Brownian motion – facts and formulae*, 2nd edition, Birkhäuser Verlag, Basel, Switzerland.

Bowen, J.M. (1994), Delta function terms arising from classical point source fields, *American Journal of Physics* Vol. **62**, pp. 511 - 515.

Boykin, T.B. (2003), Derivatives of the Dirac delta function by explicit construction of sequences, *American Journal of Physics* Vol. **72**, pp. 462 - 468.

Breda, D., Maset, S., and Vermiglio, R. (2004), Computing the characteristic roots for delay differential equations, *IMA Journal of Numerical Analysis* Vol. **2**, pp. 1 - 19.

Breda, D. (2006), Solution operator approximations for characteristic roots of delay differential equations, *Applied Numerical Mathematics* Vol. **56**, pp. 305 - 317.

Brenan, K.E., Campbell, S.L. and Petzold, L.R. (1996), *Numerical solution of initial-value problems in differential algebraic equations*, Vol. 14 of Classics in Applied Mathematics, 2nd edition, SIAM, Philadelphia, Pennsylvania.

Campbell, S.L. and Meyer, Jr., C.D. (1979), *Generalized inverses of linear transformations*, Dover Publications, USA.

Campbell, S.L. (1980), *Singular systems of differential equations*. I. Research Notes in Mathematics 40. Pitman, Boston.

Campbell, S.L. (1982), *Singular systems of differential equations*. II. Research Notes in Mathematics 61. Pitman, Boston.

Cobb, J.D. (1982), On the Solutions of Linear Differential Equations with Singular Coefficients, *Journal of Differential Equations*, Vol. **46**, No. 3, pp. 310 - 323.

Cobb, J.D. (1983), Descriptor Variable Systems and Optimal State Regulation, *IEEE Transactions on Automatic Control*, Vol. **28**, No. 5, pp. 601 - 611.

Dai, L. (1989), *Singular Control Systems*, Springer-Verlag, Berlin, Heidelberg.

Dawson, D. (1970), Generalized stochastic integrals and equations, *Transactions of the American Mathematical Society* Vol. **147**, pp. 473 - 506.

Descreusefond, L. and Üstünel, A.S. (1995), Application du calcul de variations stochastique au mouvement brownien fractionnaire, *C. R. Acad. Sci. Paris Sér. I Math.* Vol. **321**, no. 12, pp. 1605 - 1608.

Driver, R.D. (1977), *Ordinary and Delay Differential Equations*, Springer-Verlag, New York.

Duncan, T. E., Hu, Y. Z. and Pasik – Duncan, B. (2000), Stochastic calculus for fractional Brownian motion, I. Theory, *SIAM Journal of Control and Optimization*. Vol. **38**, pp. 582 - 612.

Diekmann, O., Van Gils, S.A., Verduyn Lunel, S.M., and Walther, H.O. (1995), *Delay Equations (Functional-, Complex-, and Nonlinear Analysis)*, Springer-Verlag, New York.

Eich-Soellner, E. and Führer, C. (1998), *Numerical methods in multi-body systems*, B.G. Teubner, Stuttgart.

Ehrhard, M. And Kliemann, W. (1982), Controllability of Stochastic Linear Systems, *System and Control Letters*, Vol. **2**, pp. 45 - 153.

Eisinberg, A., Franzé, G. and Salerno, N. (2001), Rectangular Vandermonde matrices on Chebyshev nodes, *Linear Algebra and its Applications* Vol. **338**, pp. 27 - 36.

Eisinberg, A. and Fedele, G. (2006), On the inversion of the Vandermonde matrix, *Applied Mathematics and Computation* Vol. **174**, pp. 1384 - 1396.

Elsgolts, L. E. (1966), *Introduction to the Theory of Differential Equations with Deviating Arguments*, Holden - Day, San Francisco.

Estrada, R. and Kanwal, R.P. (2000), *Singular integral equations*, Birkhauser, Boston.

Fernique, X. (1967), Processus linéaires, processus generalizes, *Ann. Inst. Fourier (Grenoble)* Vol. **17**, pp. 1 - 92.

Fleissner, P. (1990), Dynamic Leontief models on the test bed, Structural Change and Economic Dynamics Vol. **1**, No. 2, pp. 321 - 357.

Forney, D.G. (1975), Minimal bases of rational vector spaces with application to multivariable systems, *SIAM Journal of Control* Vol. **13**, pp 505 - 514.

Gantmacher, R.F. (1959), *The Theory of Matrices*, Vol. I and II, Chelsea, New York.

Gel'fand M. I. and Vilenkin Ja. N. (1961), *Generalized functions*, Vol. 4: Some applications of harmonic analysis, Fizmatgiz, Moscow; English translation, Academic Press, New York, 1964.

Griepentrog, E. and März, R. (1986), *Differential-Algebraic equations and their numerical treatment*, B.G. Teubner Texte zur Mathematik. Teubner-Verlag, Leipzig.

Grispos, E. (1992), Singular generalized autonomous linear differential systems, *Bulletin of the Greek Mathematical Society* Vol. **34**, pp. 25 - 43.

Gupta, S.C. and Hasdorff, L. (1963), Changing the state of a linear system by use of normal function and its derivatives, *International Journal of Electronics* Vol. **14**, pp. 351 - 359.

Gupta, S.C. (1966), *Transform and state variable methods in linear systems*, Wiley New York, U.S.A.

Hairer, E., Lubich, C. and Roche, C. (1989), *The numerical solution of Differential-Algebraic systems by Runge-Kutta methods*, Lecture Notes in Mathematics, No. 1409, Springer, Berlin – Heidelberg – New York.

Hairer, E. and Wanner, G. (1996), *Solving ordinary differential equations II. Stiff and Differential-Algebraic problems*, 2nd edition, Springer, Berlin – Heidelberg - New York.

Hale, J.K. (1977), *Theory of Functional Differential Equations*, Springer - Verlag, New York.

Hale, J.K. and Verduyn Lunel, S.M. (1993), *Introduction to Functional Differential Equations*, Springer - Verlag, New York.

Hirsch, F. and Lacombe, G. (1999), *Elements of Functional Analysis*, Springer - Verlag, USA.

Hu, Y.Z. and Øksendal B. (2003), Fractional white noise calculus and applications to finance, *Infin. Dimens. Anal. Quantum Prob. Relat. Top.*, Vol. **6**, pp. 1 - 32.

Hu, Y.Z. (2005), *Integral transformations and anticipative calculus for fractional Brownian motions*, Memoirs of the American Mathematical society, Vol. 175, Number 825, Providence, Rhode Island.

Jaffe, S. and Karcanias, N. (1981), Matrix pencil characterization of almost (A, B)- invariant subspaces: a classification of geometric concepts. *International Journal of Control* Vol. **33**, pp. 51 - 93.

Jodar, L. and Merello, P. (2010), Solving an analytic dynamic Leontief model with time dependent capital matrix, *Mathematical and Computer Modelling* Vol. **51**, 400 - 404.

Kanwal, R.P. (2004), *Generalized Functions: Theory and applications*, Birkäuser, 3rd edition, USA.

Kalogeropoulos, G.I. (1985), *Matrix Pencils and Linear Systems*, PhD Thesis, City University, London.

Kalogeropoulos, G.I. and Stratis, I.G. (1999), On generalized linear regular delay systems, *Journal of Mathematical Analysis and Applications* Vol. **237**, pp. 505 - 514.

Kalogeropoulos, G.I., Pantelous, A.A. and Papachristopoulos, D.P. (2008), Normalizability and feedback stabilization for second-order linear descriptor differential systems. *Proc. of 16th Mediterranean Conference on Control and Automation (MED'08)*, pp. 996 - 1001.

Kalogeropoulos, G.I., Karageorgos, A.D. and Pantelous, A.A. (2009), Higher-Order Linear Matrix Descriptor Differential Equations of Apostol-Kolodner Type, *Electronic Journal of Differential Equatio*ns Vol. **2009**, No. 25, pp. 1 - 13.

Karageorgos, A.D., Pantelous, A.A. and Kalogeropoulos, G.I. (2009), Transferring instantly the state of higher-order linear descriptor (regular) differential systems using impulsive inputs, *Journal of Control Science and Engineering* Vol. **2009,** pp. 1 - 32.

Karcanias, N. (1979), Matrix pencil approach to geometric systems theory, *Proceedings of IEE* Vol. **126**, pp. 585 - 590.

Karcanias, N. (2008), Structure evolving systems and control in integrated design, *Annual Reviews in Control* Vol. **32**, pp. 161 – 182.

Karcanias, N. and Hayton, G.E. (1982), Generalised autonomous dynamical systems, algebraic duality and geometric theory. *Proceedings of 8th IFAC World Congress*, Kyoto, Japan, Pergamon Press, pp. 289 - 294.

Karcanias, N. and Kalogeropoulos, G.I. (1989), Geometric Theory and Feedback Invariants of Generalized Linear Systems: A Matrix Pencil Approach, *Circuits, Systems and Signal Processing*, Special issue on Singular Systems Vol. **8**, No. 3, pp. 395 - 397.

Karcanias, N. and Kouvaritakis, B. (1979), Zero time adjustment of initial conditions and its relationship to controllability subspaces, *International Journal of Control* Vol. **29**, No. 5, pp. 749 - 765.

Kaufman, I. (1969), The inversion of the Vandermonde matrix and the transformation to the Jordan canonical form, *IEEE Transactions of Automatic Control* Vol. **14**, pp. 774 - 731.

Klamka, J and Socha, L. (1977), Some remarks about stochastic controllability, *IEEE Transactions on Automatic Control*, Vol. AC-**22**, pp. 880 – 881.

Klinger, A. (1967), The Vandermonde matrix, *The American Mathematical Monthly* Vol. **74**, No. 5, pp. 571 - 574.

Kunkel P. and Mehrmann V. (2006), *Differential-Algebraic equations*, European Mathematical Society, Zürich, Switzerland.

Kuo H. H. (1975), *Gaussian measures in Banach spaces*, Springer-Verlag, Berlin.

Kytagias, D. (1993), *An Algorithmic Method of Computation of the Reduced set of Quadratic Plücker Relations and Applications in Feedback Problems of Regular and Singular Control Systems*, PhD Thesis, University of Athens, Greece.

Leontief, W. (1966), *Input-Output Economics*, Oxford University Press, New York.

Lewis, F.L. (1986), A survey of linear singular systems, *Circuits Systems Signal Process* Vol. **5**, No. 1, pp. 3 - 36.

Luenberger, G and Arbel, A. (1977), Singular dynamic Leontief systems, Notes and Comments, Econometrica Vol **45** No. 4, pp. 991 - 995.

Mahmudov, N.I. (2001), Controllability of Linear Stochastic Systems, *IEEE Transactions of Automatic Control* Vol. **46**, pp. 724 - 777.

Martinez, J.J. and Peña, J.M. (1998a), Factorization of Cauchy-Vandermonde matrices, *Linear Algebra and its Applications*, Vol. **284**, pp. 229 - 237.

Martinez, J.J. and Peña, J.M. (1998b), Fast algorithms of Björck- Pereyra type for solving Cauchy-Vandermonde linear systems, *Applied Numerical Mathematics*, Vol. **26**, pp. 343 - 352.

Mohammed S.E.A (1984), *Stochastic functional differential equations*, Vol. 99 of Research Notes in Mathematics. Pitman (Advanced Publishing Program), Boston, MA.

Mohammed S.E.A. (1998), *Stochastic differential systems with memory: Theory, examples and applications*. In *Stochastic Analysis and Related Topics VI. The Geilo Workshop, 1996*, eds. L. Decreusefond, Jon Gjerde, B. Øksendal, and A. S. Ustunel, 1–77. Geilo, Norway: Birkhäuser.

Nualart D. (2006), *Fractional Brownian motion: stochastic calculus and applications*. *International Congress of Mathematicians. Vol. III*, 1541-1562, *Eur. Math. Soc., Zürich.*

O'Connor, R. and Henry, E.W. (1975), Input-Output Analysis and its Applications, Charles Griffin and Company Ltd., London.

Oruç, H. and Phillips, G.M. (2000), Explicit factorization of the Vandermonde matrix, *Linear Algebra and its Applications* Vol. **315**, pp. 113 - 123.

Pantelous, A.A., Zimbidis, A.A., and Kalogeropoulos, G.I. (2010), Linear Generalized Stochastic Systems for Insurance Portfolio, *Stochastic Analysis and Applications*, Vol. **28**, No. 6, pp. 946 - 971.

Pedersen, M. (2000), *Functional Analysis in Applied Mathematics and Engineering*, CRC Press LLC, USA.

Shampine, L.F. and Thompson, S. (2001), Solving Delay Differential Equations in MatLab, *Applied Numerical Mathematics* Vol. **37**, pp. 441 - 458.

Schwartz L. (1966), *Théorie des distributions*, Publ. Inst. Math. Univ. Strasbourg, nos. 9, 10, Hermann, Paris.

Szyld, D.B. (1985), *Conditions for the existence of a balanced growth solution for the Leontief dynamic input-output model*, Econometrica Vol. **53**, No. 6, 1411-1419.

Tang, W.P. and Golub, G.H. (1981), The block decomposition of a Vandermonde matrix and its applications, *BIT*, Vol. **2**1, pp. 505 - 517.

Urbanik K. (1957), Generalized stochastic processes, *Studia Mathematica* Vol. **16**, pp. 268 - 234.

Urbanik K. (1958), Local characteristics of generalized stochastic processes, *Studia Mathematica* Vol. **18**, pp. 199 - 266.

Van Dooren, P. (1983), *Reducing Subspaces: Definitions, Properties, and Algorithms, Matrix Pencils*, Lecture Notes in Mathematics, Vol. 973, Springer-Verlag, Berlin/New York.

Verghese, G.C. (1979), *Infinite Frequency Behaviour of Generalised Dynamical Systems*, PhD Thesis, Stanford University.

Verghese, G.C. and Kailath, T. (1979), Impulsive Behaviour and Dynamical Systems, *Proceedings of 4^{th} Symposium on Mathematical Theory of Networks and Systems*, Delft, The Netherlands, pp.162 - 168.

Wei, J. (2004), Eigenvalue and stability of singular differential delay systems, *Journal of Mathematical Analysis and Applications* Vol. **297**, pp. 305 - 316.

Wertz, H.J. (1965), On the numerical inversion of a recurrent problem: The Vandermonde matrix, *IEEE Transactions on Automatic Control*, Vol. **10**, p. 492.

Wiener J. (1993), *Generalized Solution of Functional Differential Equations*, World Scientific, New Jersey.

Willems, J.C. (1981), Almost Invariant spaces: An Approach to High Gain Feedback-Part I: Almost controlled invariant subspaces, *IEEE Transactions Automatic Control* Vol. AC-**26**, pp. 235 - 252.

Willems, J.C. (1989), *Model for dynamics*, Dynamics Report 2, 65, pp. 171 – 269.

Willems, J.C. (1991), Paradigms and Puzzles in the Theory of Dynamical Systems, *IEEE Transactions on Automatic Control* Vol. **36**, pp 259 - 294.

Yan, F. and Mohammed, S.E.A (2005), Stochastic calculus for systems with memory. *Stochastic Analysis and Applications* Vol. **23**, No. 3, pp. 613 - 657.

Yang, M.S and Liu, P.L (2002), On asymptotic stability of linear neutral delay-differential systems, *International Journal of Systems Science*, Vol. **33**, No. 11, pp. 901 - 907.

Yosida, K. (1966), *Functional Analysis*, Springer-Verlag, USA.

Zabcyck, J. (1981), Controllability of Stochastic Linear Systems, *System and Control Letters*, Vol. **1**, pp. 25 – 31.

Zadeh, Z.A. and Desoer, C.A. (1963), *Linear system theory, the state space approach*, McGraw-Hill, New York, USA.

Zemanian, A.H. (1987), *Distribution theory and transform analysis: An introduction to generalized functions with applications*, Dover Publications, Inc, New York, USA.

Zhao, X.H. and Jiang, L. (2009), Computer analysis algorithm of stability of nonlinear continuous--time Leontief input-output model, Proceedings of the Eighth International Conference on Machine Learning and Cybernetics, Baoding, pp. 2081 - 2085.

Published results of the present PhD Thesis

J[1]1. Generalized regular differential systems with distributed delay, *Systems Science* Vol. **34,** No. 1, 2008, pp. 15 - 22. (with Kalogeropoulos, G.I. and Kossak, O.)

J2. On Generalized regular stochastic differential-algebraic delay systems with time invariant coefficients**,** *Stochastic Analysis and Applications* Vol. **26**, No. 5, 2008, pp. 1076 - 1094. (with Kalogeropoulos, G.I.)

J3. On Generalized linear singular neutral delay differential systems, *Journal of the Franklin Institute* Vol. **346**, No. 7, 2009, pp. 691 - 704. (with Kalogeropoulos, G.I.)

J4. Approximating distributional behaviour of linear differential systems using Gaussian function and its derivatives *International Journal of Control* Vol. **85**, No. 7, 2012, pp. 830 - 841 (with Karcanias, N. and Halikias, G.)

J5. Generalized Inverses of the Vandermonde Matrix: Applications in Control Theory, *International Journal of Control, Automation and Systems* Vol. **11**, No. 5, 2013, pp. 1063-1070 (with Karageorgors, A.D.)

C[2]1. Generalized inverses of the general Vandermonde matrix, *Proceedings of the 16th Conference of the International Linear Algebra Society (ILAS)*, 21 - 25 June 2010, Pisa, Italy (after invitation to the mini-symposium "*Generalized Inverses and Applications*" organized by Nieves Castro-Gonzalez, and Pedro Patricio).

C2. Approximating distributional behaviour of linear systems using Gaussian function and its derivatives, *Proceeding of the 4th IFAC Symposium on System, Structure and Control (SSSC10)*, 15 - 18 September 2010, Ancora, Italy

C3. Generalized inverses for a special structure matrix appearing in Control. *Proceedings of the 6th IFNA World Congress (Global Health, Environment, Economy, and Energy Using Mathematical Sciences) Conference*, 25 June - 1July 2012, Athens, Greece, pp. 1-9 (with Karageorgos, A.D.)

[1] **J**: Journal
[2] **C**: Conference