HAL open science

State Space Modelling

Thierry Miquel

▶ **To cite this version:**

Thierry Miquel. State Space Modelling. Master. ENAC, France. 2022. hal-02987750v3

HAL Id: hal-02987750

https://cel.hal.science/hal-02987750v3

Submitted on 18 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

State Space Modelling

Lecture notes - Draft

Thierry Miquel
thierry.miquel@enac.fr

October 12, 2022

Course overview

Classical control theory is intrinsically linked to the frequency domain and the s-plane. The main drawback of classical control theory is the difficulty to apply it in Multi-Input Multi-Output (MIMO) systems. Rudolf Emil Kalman (Hungarian-born American, May 19, 1930 – July 2, 2016) is one of the greatest protagonist of modern control theory[1]. He has introduced the concept of state as well as linear algebra and matrices in control theory. With this formalism systems with multiple inputs and outputs could easily be treated.

The purpose of this lecture is to present an overview of modern control theory. More specifically, the objectives are the following:

- to learn how to model dynamic systems in the state-space and the state-space representation of transfer functions;
- to learn linear dynamical systems analysis in state-space: more specifically to solve the time invariant state equation and to get some insight on controllability, observability and stability;
- to learn state-space methods for observers and controllers design.

Assumed knowledge encompass linear algebra, Laplace transform and linear ordinary differential equations (ODE)

This lecture is organized as follows:

- The first chapter focuses on the state-space representation as well as state-space representation associated to system interconnection;
- The conversion from transfer functions to state-space representation is presented in the second chapter. This is also called transfer function realization;
- The analysis of linear dynamical systems is presented in the third chapter; more specifically we will concentrate on the solution of the state equation and present the notions of controllability, observability and stability;
- The fourth chapter is dedicated to observers design. This chapter focuses on Luenberger observer, state observer for SISO systems in observable canonical form, state observer for SISO systems in arbitrary state-space representation and state observer for MIMO systems will be presented.

[1] http://www.uta.edu/utari/acs/history.htm

- The fifth chapter is dedicated to observers and controllers design. As far as observers and controllers are linked through the duality principle the frame of this chapter will be similar to the previous chapter: state feedback controller for SISO systems in controllable canonical form, state feedback controller for SISO systems in arbitrary state-space representation, static state feedback controller and static output feedback controller for MIMO systems will be presented.

References

[1] *Introduction au contrôle des mini-drones : de la conception à la mise en œuvre*, Thierry Miquel, Jean-Philippe Condomines, Bordeneuve-Guibé Joël, Brière Yves, Ellipses Formations & Techniques

[2] *Modern Control Systems*, Richard C. Dorf, Robert H. Bishop, Prentice Hall

[3] *Feedback Control of Dynamic Systems*, Gene F. Franklin, J Powel, Abbas Emami-Naeini, Pearson

[4] *Linear Control System Analysis and Design*, Constantine H. Houpis, Stuart N. Sheldon, John J. D'Azzo, CRC Press

[5] *Modern Control Engineering*, P.N. Paraskevopoulos, CRC Press

[6] *Control Engineering: An Introductory Course*, Michael Johnson, Jacqueline Wilkie, Palgrave Macmillan

[7] *Control Engineering: A Modern Approach*, P. Bélanger, Oxford University Press

[8] *Multivariable Feedback Control Analysis and design*, S. Skogestad and I. Postlethwaite, Wiley

Table of contents

1 State-space representation **13**

1.1 Introduction . . . 13
1.2 State and output equations . . . 14
1.3 From ordinary differential equations to state-space representation 16
1.3.1 Brunovsky's canonical form . . . 16
1.3.2 Linearization of non-linear time-invariant state-space representation . . . 17
1.4 From state-space representation to transfer function . . . 23
1.5 Zeros of a transfer function - Rosenbrock's system matrix . . . 25
1.6 Faddeev-Leverrier's method to get $(s\mathbb{I} - \mathbf{A})^{-1}$ . . . 27
1.7 Matrix inversion lemma . . . 30
1.8 Interconnection of systems . . . 31
1.8.1 Parallel interconnection . . . 31
1.8.2 Series interconnection . . . 32
1.8.3 Feedback interconnection . . . 33

2 Realization of transfer functions **35**

2.1 Introduction . . . 35
2.2 Non-unicity of state-space representation . . . 35
2.2.1 Similarity transformations . . . 35
2.2.2 Inverse of a similarity transformation . . . 37
2.3 Realization of SISO transfer function . . . 37
2.3.1 Controllable canonical form . . . 38
2.3.2 Poles and zeros of the transfer function . . . 42
2.3.3 Similarity transformation to controllable canonical form . 42
2.3.4 Observable canonical form . . . 49
2.3.5 Similarity transformation to observable canonical form . . 52
2.3.6 Diagonal (or modal) form . . . 59
2.3.7 Algebraic and geometric multiplicity of an eigenvalue . . . 66
2.3.8 Jordan form and generalized eigenvectors . . . 67
2.3.9 Transfer function with multiple poles on the same location 69
2.4 Realization of SIMO transfer function . . . 74
2.4.1 Generic procedure . . . 75
2.4.2 Controllable canonical form . . . 76
2.5 Realization of MIMO transfer function . . . 79
2.5.1 Generic procedure . . . 80

2.5.2 Controllable canonical form 81
2.5.3 Observable canonical form 83
2.5.4 Diagonal (or modal) form 83
2.6 Minimal realization 87
2.6.1 System's dimension 87
2.6.2 Gilbert's minimal realization 90
2.6.3 Ho-Kalman algorithm 91

3 Analysis of Linear Time Invariant systems 93
3.1 Introduction 93
3.2 Solving the time invariant state equation 93
3.3 Output response 95
3.4 Impulse and unit step responses 95
3.5 Matrix exponential 96
3.5.1 Definition 96
3.5.2 Properties 98
3.5.3 Computation of $e^{\mathbf{A}t}$ using the diagonal form of $\mathbf{A}$ 99
3.5.4 Computation of $e^{\mathbf{A}t}$ using the Jordan form of $\mathbf{A}$ 103
3.5.5 Dunford (or Jordan–Chevalley) decomposition 106
3.5.6 Computation of $e^{\mathbf{A}t}$ using the Laplace transform 109
3.6 Stability 113
3.7 Controllability 114
3.7.1 Definition 114
3.7.2 Kalman's controllability rank condition 116
3.7.3 Use of the diagonal form: Gilbert's criteria 118
3.7.4 Popov-Belevitch-Hautus (PBH) test 119
3.7.5 Uncontrollable mode 120
3.7.6 Stabilizability 121
3.8 Observability 122
3.8.1 Definition 122
3.8.2 Kalman's observability rank condition 123
3.8.3 Use of the diagonal form: Gilbert's criteria 125
3.8.4 Input-output differential equation and observable modes . 126
3.8.5 Popov-Belevitch-Hautus (PBH) test 127
3.8.6 Unobservable mode 128
3.8.7 Detectability 128
3.8.8 Additional observability topic: relation between linear differential equation and state-space representation 128
3.9 Interpretation of the diagonal (or modal) decomposition 129
3.10 Duality principle 132
3.11 Kalman decomposition 132
3.11.1 Controllable / uncontrollable decomposition 132
3.11.2 Observable / unobservable decomposition 133
3.11.3 Canonical decomposition 134
3.12 Minimal realization (again!) 138

4 Observer design **139**
4.1 Introduction . . . 139
4.2 Luenberger observer . . . 140
4.2.1 State observer . . . 140
4.2.2 Estimating output derivative . . . 142
4.2.3 Output observer . . . 143
4.3 State observer for SISO systems in observable canonical form . . 144
4.4 State observer for SISO systems in arbitrary state-space representation . . . 146
4.5 Ackermann's formula . . . 148
4.6 State observer for MIMO systems - Roppenecker's formula . . . 150
4.7 Reduced-order observer . . . 154

5 Controller design **157**
5.1 Introduction . . . 157
5.2 Static state feedback controller . . . 157
5.3 Control of SISO systems . . . 159
5.3.1 State feedback controller in controllable canonical form . . 159
5.3.2 State feedback controller in arbitrary state-space representation . . . 160
5.3.3 Ackermann's formula . . . 162
5.3.4 Invariance of (transmission) zeros under state feedback . . 165
5.4 Static state feedback controller . . . 165
5.4.1 Roppenecker's formula . . . 165
5.4.2 Comments on Roppenecker's formula . . . 168
5.4.3 Solving Sylvester equation . . . 171
5.4.4 Solving general algebraic Riccati and Lyapunov equations 171
5.5 Control with integral action . . . 173
5.5.1 Adding an integrator in the state-space realization . . . 173
5.5.2 Proof of the cancellation of the steady state error through integral augmentation . . . 174
5.6 Observer-based controller . . . 176
5.6.1 Separation principle . . . 176
5.6.2 Example . . . 178
5.6.3 Feedback controller . . . 180
5.6.4 Unity feedback loop . . . 182
5.7 Static output feedback controller . . . 185
5.7.1 Partial eigenvalues assignment . . . 185
5.7.2 Changing PID controller into static output feedback . . . 188
5.7.3 Adding integrators, controllability and observability indexes 190
5.8 Dynamical output feedback controller . . . 191
5.8.1 From dynamical output feedback to observer-based control 191
5.8.2 Dynamical output feedback . . . 192
5.9 Mode decoupling controller . . . 195
5.9.1 Eigenstructure assignment . . . 195
5.9.2 Design procedure . . . 197
5.9.3 Example . . . 201

5.10 Sensitivity to additive uncertainties . . . 204
5.10.1 Sensitivity analysis . . . 204
5.10.2 Robust placement method . . . 205
5.10.3 Lyapunov approach . . . 207
5.10.4 Eigenvalues derivatives . . . 209
5.11 Input-output decoupling control . . . 210
5.11.1 Necessary and sufficient condition . . . 210
5.11.2 Relative degree . . . 212
5.11.3 Falb-Wolovich decoupling method . . . 213
5.11.4 Zeros and zero dynamics . . . 215
5.12 Frequency domain approach . . . 217
5.12.1 Hsu-Chen theorem . . . 217
5.12.2 Output-based controller . . . 220
5.12.3 Invariance of (transmission) zeros under state feedback . . . 223
5.12.4 Matrix fraction description of a transfer function . . . 224
5.12.5 Youla parameterization of all stabilizing controllers . . . 225
5.12.6 Frequency domain representation of state feedback loop . 226
5.12.7 Pre-filtering applied to SISO plants . . . 228

Appendices **231**

A Refresher on linear algebra **233**
A.1 Section overview . . . 233
A.2 Vectors . . . 233
A.2.1 Definitions . . . 233
A.2.2 Vectors operations . . . 234
A.3 Matrices . . . 235
A.3.1 Definitions . . . 235
A.3.2 Matrix Operations . . . 236
A.3.3 Properties . . . 237
A.3.4 Determinant and inverse . . . 237
A.4 Eigenvalues and eigenvectors . . . 239

B Overview of Lagrangian Mechanics **241**
B.1 Euler-Lagrange equations . . . 241
B.2 Robot arm . . . 247
B.3 Quadrotor . . . 248
B.3.1 Inertial frame and body frame . . . 249
B.3.2 Direct Cosine Matrix (DCM) and kinematic relations . . . 249
B.3.3 Forces and torques . . . 251
B.3.4 Generalized coordinates . . . 252
B.3.5 Inertia matrix . . . 252
B.3.6 Kinetic energy . . . 252
B.3.7 Potential energy . . . 253
B.3.8 Lagrangian . . . 254
B.3.9 Euler-Lagrange equations . . . 254
B.3.10 Newton-Euler equations . . . 255

B.3.11 Translational equations of motion with wind 257
B.3.12 Small angle approximation of angular dynamics 258
B.3.13 Synthesis model 259
B.3.14 BLDC motor with ESC and propeller 259

C Linear Algebraic Method for Control System Design 263
C.1 Matrix polynomial fraction 263
C.1.1 Greatest Common Divisor of two polynomials 263
C.1.2 Smith form of polynomial matrix 264
C.1.3 Smith-McMillan form of transfer matrix 270
C.1.4 System dimension, poles and zeros 273
C.1.5 Transfer matrix realization 275
C.1.6 Polynomial matrix fraction description (P-MFD) 278
C.1.7 Right and left coprimeness 282
C.1.8 Hermite form and applications 283
C.1.9 McMillan degree, poles and zeros of a transfer function . . 290
C.1.10 Column reduced and row reduced polynomial matrix . . . 290
C.1.11 Row index of transfer matrix 292
C.1.12 Flat outputs 293
C.2 ICD controllers 293
C.2.1 Controller structure 293
C.2.2 Closed loop transfer function 295
C.2.3 Proper controller 296
C.2.4 Decoupling 298
C.2.5 Two-degree-of-freedom control loop 299
C.3 Matrix polynomials and solvents of λ-matrices 300
C.3.1 Latent roots and latent vectors 300
C.3.2 Right and left solvents 300
C.3.3 Block Vandermonde matrices 301
C.3.4 Modal form based on right solvents 302
C.3.5 Modal form based on left solvents 303

D Singular perturbations and hierarchical control 305
D.1 Block triangular and block-diagonal forms 305
D.1.1 Block triangular form 305
D.1.2 Block-diagonal form 306
D.1.3 Similarity transformation 307
D.2 Two-stage design for eigenvalue placement of block-diagonal systems 308
D.3 Singularly perturbed system 313
D.3.1 Linear system 313
D.3.2 Nonlinear system 317
D.4 Two-frequency-scale transfer function 318
D.5 Hierarchical output feedback of singularly perturbed system . . . 321
D.6 Weierstrass decomposition and descriptor system 323

E Carleman linearization **329**
E.1 Reduced Kronecker product 329
E.2 Taylor series expansion of any continuous vector field 331
E.3 Time derivative 332
E.4 Carleman linearization of nonlinear system affine in control . . . 335
E.5 Application to mechanical systems 339
E.6 Control law for bilinear systems 341
E.6.1 Linear state feedback 341
E.6.2 Quadratic state feedback 342
E.6.3 Optimal control 344

F Introduction to fractional systems **347**
F.1 Pre-filtering 347
F.2 Design steps 348
F.3 Pre-filtering design for non-minimum phase feedback loop 349
F.4 Mathematical Definitions of Fractional Integral and Derivative . . 350
F.5 CRONE approximation of fractional derivative 351
F.6 State space fractional systems 353
F.7 Approximation of fractional systems based on differentiation operator 354
F.8 Approximation of fractional systems based on integration operator 355

Chapter 1

State-space representation

1.1 Introduction

This chapter focuses on the state-space representation as well as conversions from state-space representation to transfer function. The state-space representation associated to system interconnection is also presented.

The notion of state-space representation has been developed in the former Soviet Union where control engineers preferred to manipulate differential equations rather than transfer functions which originates in the United States of America. The diffusion to the Western world of state-space representation started after the first congress of the International Federation of Automatic Control (IFAC) which took place in Moscow in 1960.

One of the interest of the state-space representation is that it enables to generalize the analysis and control of Multi-Input Multi-Output (MIMO) linear systems with the same formalism than Single-Input Single-Output (SISO) linear systems.

Let's start with an example. We consider a system described by the following second-order linear differential equation with a damping ratio denoted m, an undamped natural frequency ω_0 and a static gain K :

$$\frac{1}{\omega_0^2}\frac{d^2y(t)}{dt^2} + \frac{2m}{\omega_0}\frac{dy(t)}{dt} + y(t) = Ku(t) \tag{1.1}$$

Here $y(t)$ denotes the output of the system whereas $u(t)$ is its input. The preceding relation represents the input-ouput description of the system.

The transfer function is obtained thanks to the Laplace transform and assuming that the initial conditions are zero (that is $\dot{y}(t) = \ddot{y}(t) = 0$). We get:

$$\begin{array}{l} \frac{1}{\omega_0^2}s^2Y(s) + \frac{2m}{\omega_0}sY(s) + Y(s) = KU(s) \\ \Leftrightarrow F(s) = \frac{Y(s)}{U(s)} = \frac{K\omega_0^2}{s^2+2m\omega_0 s+\omega_0^2} \end{array} \tag{1.2}$$

Now rather than computing the transfer function, let's assume that we wish to transform the preceding second order differential equation into a single first order vector differential equation. To do that we introduce two new variables,

say x_1 and x_2, which are defined for example as follows:

$$\begin{cases} y(t) = K\omega_0^2 x_1(t) \\ \dot{x}_1(t) = x_2(t) \end{cases} \tag{1.3}$$

Thanks to the new variables x_1 and x_2 the second order differential equation (1.1) can now be written as follows:

$$\begin{cases} \frac{dy(t)}{dt} = K\omega_0^2 \frac{dx_1(t)}{dt} = K\omega_0^2 x_2(t) \\ \frac{d^2y(t)}{dt^2} = K\omega_0^2 \frac{dx_2(t)}{dt} \end{cases} \\ \Rightarrow \frac{dx_2(t)}{dt} + 2m\omega_0 x_2(t) + \omega_0^2 x_1(t) = u(t) \tag{1.4}$$

The second equation of (1.3) and equation (1.4) form a system of two coupled first order linear differential equations:

$$\begin{cases} \frac{dx_1(t)}{dt} = x_2(t) \\ \frac{dx_2(t)}{dt} = -2m\omega_0 x_2(t) - \omega_0^2 x_1(t) + u(t) \end{cases} \tag{1.5}$$

In is worth noticing that variables $x_1(t)$ and $x_2(t)$ constitute a vector which is denoted $\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$: this is the state vector. Equation (1.5) can be rewritten in a vector form as follows:

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -2m\omega_0 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) \tag{1.6}$$

Furthermore using the first equation of (1.3) it is seen that the output $y(t)$ is related to the state vector $\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$ by the following relation:

$$y(t) = \begin{bmatrix} K\omega_0^2 & 0 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{1.7}$$

Equations (1.6) and (1.7) constitute the so called state-space representation of the second order system model (1.4). This representation can be generalized as follows:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{1.8}$$

The state-space representation is formed by a state vector and a state equation. This representation enables to describe the dynamics of a linear dynamical systems through n first order differential equations, where n is the size of the state vector, or equivalently through a single first order vector differential equation.

1.2 State and output equations

Any system that can be described by a finite number of nth order linear differential equations with constant coefficients, or any system that can be

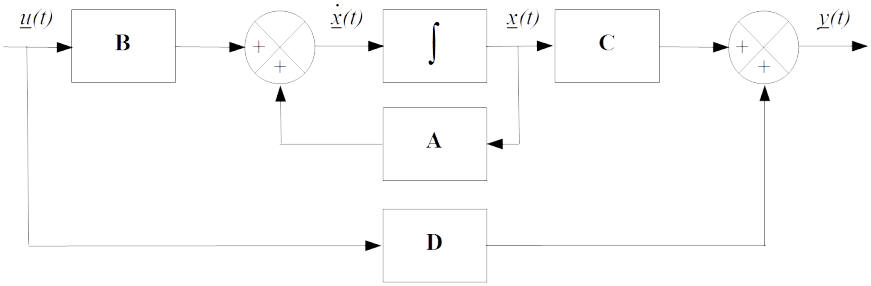

Figure 1.1: Block diagram of a state-space representation

approximated by them, can be described using the following state-space representation:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{1.9}$$

Where:

- $\underline{x}(t)$ is the state vector, which is of dimension n. The number n of the state vector components is called the order of the system;
- $\underline{u}(t)$ is the input of the system;
- $\underline{y}(t)$ is the output of the system.

State vector $\underline{x}(t)$ can be defined as a set of variables such that their knowledge at the initial time $t_0 = 0$, together with knowledge of system inputs $\underline{u}(t)$ at $t \geq 0$ are sufficient to predict the future system state and output $\underline{y}(t)$ for all time $t > 0$.

Both equations in (1.9) have a name:

- Equation $\dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t)$ is named as the state equation;
- Equation $\underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t)$ is named as the output equation.

The state equation and the output equation both constitute the state-space representation of the system.

The block diagram corresponding to state-space representation (1.9) is shown in Figure 1.1.

Furthermore matrices $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ which define the state-space representation of the system are named as follows [1]:

- $\mathbf{A}$ is the state matrix and relates how the current state affects the state change $\dot{\underline{x}}(t)$. This is a constant $n \times n$ square matrix where n is the size of the state vector;
- $\mathbf{B}$ is the control matrix and determines how the system inputs $\underline{u}(t)$ affects the state change; This is a constant $n \times m$ matrix where m is the number of system inputs;

[1] https://en.wikibooks.org/wiki/Control_Systems/State-Space_Equations

- **C** is the output matrix and determines the relation between the system state $\underline{x}(t)$ and the system outputs $\underline{y}(t)$. This is a constant $p \times n$ matrix where p is the number of system outputs;

- **D** is the feedforward matrix and allows for the system input $\underline{u}(t)$ to affect the system output $\underline{y}(t)$ directly. This is a constant $p \times m$ matrix.

1.3 From ordinary differential equations to state-space representation

1.3.1 Brunovsky's canonical form

Let's consider a Single-Input Single-Output (SISO) dynamical system modelled by the following input-output relation, which is an nth order non-linear time-invariant Ordinary Differential Equation (ODE):

$$\frac{d^n y(t)}{dt^n} = g\left(y(t), \frac{dy(t)}{dt}, \frac{d^2 y(t)}{dt^2}, \cdots, \frac{d^{n-1} y(t)}{dt^{n-1}}, u(t)\right) \tag{1.10}$$

This is a time-invariant input-output relation because time t does not explicitly appears in function g.

The usual way to get a state-space equation from the nth order non-linear time-invariant ordinary differential equation (1.10) is to choose the components $x_1(t), \cdots, x_n(t)$ of the state vector $\underline{x}(t)$ as follows:

$$\underline{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix} := \begin{bmatrix} y(t) \\ \frac{dy(t)}{dt} \\ \frac{d^2 y(t)}{dt^2} \\ \vdots \\ \frac{d^{n-2} y(t)}{dt^{n-2}} \\ \frac{d^{n-1} y(t)}{dt^{n-1}} \end{bmatrix} \tag{1.11}$$

Thus Equation (1.10) reads:

$$\dot{\underline{x}}(t) = \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_{n-1}(t) \\ \dot{x}_n(t) \end{bmatrix} = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ g\left(x_1, \cdots, x_{n-1}, u(t)\right) \end{bmatrix} := f\left(\underline{x}(t), \underline{u}(t)\right) \tag{1.12}$$

Furthermore:

$$y(t) := x_1(t) = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \underline{x}(t) \tag{1.13}$$

This special non-linear state equation is called the Brunovsky's canonical form.

1.3.2 Linearization of non-linear time-invariant state-space representation

More generally most of Multi-Input Multi-Output (MIMO) dynamical systems can be modelled by a finite number of coupled non-linear first order ordinary differential equations (ODE) as follows:

$$\dot{\underline{x}}(t) = f\left(\underline{x}(t), \underline{u}(t)\right) \tag{1.14}$$

The Brunovsky's canonical form may be used to obtain the first order ordinary differential equations.

In the preceding state equation f is called a vector field. This is a time-invariant state-space representation because time t does not explicitly appears in the vector field f.

When the vector field f is non-linear there exists quite few mathematical tools which enable to catch the intrinsic behavior of the system. Nevertheless this situation radically changes when vector field f is linear both in the state $\underline{x}(t)$ and in the control $\underline{u}(t)$. The good news is that it is quite simple to approximate a non-linear model with a linear model around an equilibrium point.

We will first define what we mean by equilibrium point and then we will see how to get a linear model from a non-linear model.

An equilibrium point is a constant value of the pair $(\underline{x}(t), \underline{u}(t))$, which will be denoted $(\underline{x}_e, \underline{u}_e)$, such that:

$$\underline{0} = f\left(\underline{x}_e, \underline{u}_e\right) \tag{1.15}$$

It is worth noticing that as soon as $(\underline{x}_e, \underline{u}_e)$ is a constant value then we have $\dot{\underline{x}}_e = \underline{0}$.

Then the linearization process consists in computing the Taylor expansion of vector field f around the equilibrium point $(\underline{x}_e, \underline{u}_e)$ and to stop it at order 1. Using the fact that $f\left(\underline{x}_e, \underline{u}_e\right) = \underline{0}$ the linearization of a vector field $f\left(\underline{x}(t), \underline{u}(t)\right)$ around the equilibrium point $(\underline{x}_e, \underline{u}_e)$ reads:

$$f\left(\underline{x}_e + \delta\underline{x}, \underline{u}_e + \delta\underline{u}\right) \approx \mathbf{A}\delta\underline{x} + \mathbf{B}\delta\underline{u} \tag{1.16}$$

Where:

$$\begin{cases} \delta\underline{x}(t) = \underline{x}(t) - \underline{x}_e \\ \delta\underline{u}(t) = \underline{u}(t) - \underline{u}_e \end{cases} \tag{1.17}$$

And where matrices $\mathbf{A}$ and $\mathbf{B}$ are constant matrices:

$$\begin{cases} \mathbf{A} = \left.\frac{\partial f(\underline{x},\underline{u})}{\partial \underline{x}}\right|_{\underline{u}=\underline{u}_e, \underline{x}=\underline{x}_e} \\ \mathbf{B} = \left.\frac{\partial f(\underline{x},\underline{u})}{\partial \underline{u}}\right|_{\underline{u}=\underline{u}_e, \underline{x}=\underline{x}_e} \end{cases} \tag{1.18}$$

Furthermore as far as $\underline{x}_e$ is a constant vector we can write:

$$\dot{\underline{x}}(t) = \dot{\underline{x}}(t) - \underline{0} = \dot{\underline{x}}(t) - \dot{\underline{x}}_e = \frac{d\left(\underline{x}(t) - \underline{x}_e\right)}{dt} = \delta\dot{\underline{x}}(t) \tag{1.19}$$

Thus the non-linear time-invariant state equation (1.14) turns to be a linear time-invariant state equation:

$$\delta\dot{\underline{x}}(t) = \mathbf{A}\delta\underline{x}(t) + \mathbf{B}\delta\underline{u}(t) \tag{1.20}$$

As far as the output equation is concerned we follow the same track. We start with the following non-linear output equation:

$$\underline{y}(t) = h\left(\underline{x}(t), \underline{u}(t)\right) \tag{1.21}$$

Proceeding as to the state equation, we approximate the vector field h by its Taylor expansion at order 1 around the equilibrium point $(\underline{x}_e, \underline{u}_e)$:

$$\underline{y}(t) = h\left(\underline{x}_e, \underline{u}_e\right) + h\left(\delta\underline{x}(t) + \underline{x}_e, \delta\underline{u}(t) + \underline{u}_e\right) \approx \underline{y}_e + \mathbf{C}\delta\underline{x} + \mathbf{D}\delta\underline{u} \tag{1.22}$$

Where:

$$\underline{y}_e = h\left(\underline{x}_e, \underline{u}_e\right) \tag{1.23}$$

And where matrices $\mathbf{C}$ and $\mathbf{D}$ are constant matrices:

$$\begin{cases} \mathbf{C} = \left.\frac{\partial h(\underline{x},\underline{u})}{\partial \underline{x}}\right|_{\underline{u}=\underline{u}_e, \underline{x}=\underline{x}_e} \\ \mathbf{D} = \left.\frac{\partial h(\underline{x},\underline{u})}{\partial \underline{u}}\right|_{\underline{u}=\underline{u}_e, \underline{x}=\underline{x}_e} \end{cases} \tag{1.24}$$

Let's introduce the difference $\delta\underline{y}(t)$ as follows:

$$\delta\underline{y}(t) = \underline{y}(t) - \underline{y}_e \tag{1.25}$$

Thus the non-linear output equation (1.21) turns to be a linear output equation:

$$\delta\underline{y}(t) = \mathbf{C}\delta\underline{x}(t) + \mathbf{D}\delta\underline{u}(t) \tag{1.26}$$

Consequently a non-linear time-invariant state representation:

$$\begin{cases} \dot{\underline{x}}(t) = f\left(\underline{x}(t), \underline{u}(t)\right) \\ \underline{y}(t) = h\left(\underline{x}(t), \underline{u}(t)\right) \end{cases} \tag{1.27}$$

can be approximated around an equilibrium point $(\underline{x}_e, \underline{u}_e)$, defined by $\underline{0} = f\left(\underline{x}_e, \underline{u}_e\right)$, by the following linear time-invariant state-space representation:

$$\begin{cases} \delta\dot{\underline{x}}(t) = \mathbf{A}\delta\underline{x}(t) + \mathbf{B}\delta\underline{u}(t) \\ \delta\underline{y}(t) = \mathbf{C}\delta\underline{x}(t) + \mathbf{D}\delta\underline{u}(t) \end{cases} \tag{1.28}$$

Nevertheless is worth noticing that the linearization process is an approximation that is only valid around a region close to the equilibrium point.

The δ notation indicates that the approximation of the non-linear state-space representation is made around an equilibrium point. This is usually omitted and the previous state-space representation will be simply rewritten as follows:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{1.29}$$

Example 1.1. *Let's consider a fictitious system whose dynamics reads:*

$$\frac{d^3y(t)}{dt^3} = \cos(\ddot{y}(t)) + e^{3\dot{y}(t)} - \tan(y(t)) + u(t) \tag{1.30}$$

Find a non-linear state-space representation of this system with the Brunovsky's choice for the components of the state vector. Then linearize the state-space representation around the equilibrium output $y_e = 0$.

As far as the differential equation which describes the dynamics of the system is of order 3, there are 3 components in the state vector:

$$\underline{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \tag{1.31}$$

The Brunovsky's canonical form is obtained by choosing the following components for the state vector:

$$\underline{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} y(t) \\ \dot{y}(t) \\ \ddot{y}(t) \end{bmatrix} \tag{1.32}$$

With this choice the dynamics of the system reads:

$$\begin{cases} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} x_2(t) \\ x_3(t) \\ \cos(x_3(t)) + e^{3x_2(t)} - \tan(x_1(t)) + u(t) \end{bmatrix} \\ y(t) = x_1(t) \end{cases} \tag{1.33}$$

The preceding relations are of the form:

$$\begin{cases} \dot{\underline{x}}(t) = f\left(\underline{x}(t), u(t)\right) \\ y(t) = h\left(\underline{x}(t), u(t)\right) \end{cases} \tag{1.34}$$

Setting the equilibrium output to be $y_e = 0$ leads to the following equilibrium point $\underline{x}_e$:

$$y_e = 0 \Rightarrow \underline{x}_e = \begin{bmatrix} y_e \\ \dot{y}_e \\ \ddot{y}_e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{1.35}$$

Similarly the value of the control u_e at the equilibrium point is obtained by solving the following equation:

$$\begin{aligned} &\tfrac{d^3y_e}{dt^3} = \cos(\ddot{y}_e) + e^{3\dot{y}_e} - \tan(y_e) + u_e \\ &\Rightarrow 0 = \cos(0) + e^{3\times 0} - \tan(0) + u_e \\ &\Rightarrow u_e = -2 \end{aligned} \tag{1.36}$$

Matrices $\mathbf{A}$ *and* $\mathbf{B}$ *are constant matrices which are computed as follows:*

$$\begin{cases} \mathbf{A} = \left.\frac{\partial f(\underline{x},\underline{u})}{\partial \underline{x}}\right|_{u=u_e,\underline{x}=\underline{x}_e} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\left(1+\tan^2(x_{1e})\right) & 3e^{3x_{2e}} & -sin(x_{3e}) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 3 & 0 \end{bmatrix} \\ \mathbf{B} = \left.\frac{\partial f(\underline{x},\underline{u})}{\partial \underline{u}}\right|_{u=u_e,\underline{x}=\underline{x}_e} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{cases} \tag{1.37}$$

Similarly matrices $\mathbf{C}$ *and* $\mathbf{D}$ *are constant matrices which are computed as follows:*

$$\begin{cases} \mathbf{C} = \left.\frac{\partial h(\underline{x},\underline{u})}{\partial \underline{x}}\right|_{\underline{u}=\underline{u}_e,\underline{x}=\underline{x}_e} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \\ \mathbf{D} = \left.\frac{\partial h(\underline{x},\underline{u})}{\partial \underline{u}}\right|_{\underline{u}=\underline{u}_e,\underline{x}=\underline{x}_e} = 0 \end{cases} \tag{1.38}$$

Consequently the non-linear time-invariant state representation $\frac{d^3y(t)}{dt^3} = \cos(\ddot{y}(t)) + e^{3\dot{y}(t)} - \tan(y(t)) + u(t)$ *can be approximated around the equilibrium output* $y_e = 0$ *by the following linear time-invariant state-space representation:*

$$\begin{cases} \delta\dot{\underline{x}}(t) = \mathbf{A}\delta\underline{x}(t) + \mathbf{B}\delta u(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 3 & 0 \end{bmatrix} \delta\underline{x}(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \delta u(t) \\ \delta y(t) = \mathbf{C}\delta\underline{x}(t) + \mathbf{D}\delta u(t) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \delta\underline{x}(t) \end{cases} \tag{1.39}$$

The Scilab code to get the state matrix $\mathbf{A}$ *around the equilibrium point* $(\underline{x}_e = \underline{0}, u_e = -2)$ *is the following:*

```
function xdot = f(x,u)
    xdot = zeros(3,1);
    xdot(1) = x(2);
    xdot(2) = x(3);
    xdot(3) = cos(x(3)) + exp(3*x(2)) - tan(x(1)) + u;
endfunction

xe = zeros(3,1);
xe(3) = 0;
ue = -2;
disp(f(xe,ue), 'f(xe,ue)=');
disp(numderivative(list(f,ue),xe),'df/dx=');
```

■

Example 1.2. *We consider the following equations which represent the dynamics of an aircraft considered as a point with constant mass*[2]*:*

$$\begin{cases} m\dot{V} = T - D - mg\sin(\gamma) \\ mV\dot{\gamma} = L\cos(\phi) - mg\cos(\gamma) \\ mV\cos(\gamma)\dot{\psi} = L\sin(\phi) \\ \dot{\phi} = p \end{cases} \tag{1.40}$$

Where:

- V *is the airspeed of the aircraft;*
- γ *is the flight path angle;*
- ψ *is the heading;*

[2]Etkin B., Dynamics of Atmospheric Flight, Dover Publications, 2005

- ϕ *is the bank angle;*
- m *is the mass (assumed constant) of the aircraft;*
- T *is the Thrust force applied by the engines on the aircraft model;*
- D *is the Drag force;*
- g *is the acceleration of gravity* ($g = 9.80665\ m/s^2$)*;*
- L *is the Lift force;*
- ϕ *is the bank angle;*
- p *is the roll rate.*

We will assume that the aircraft control vector $\underline{u}(t)$ *has the following components:*

- *The longitudinal load factor* n_x*:*

$$n_x = \frac{T - D}{mg} \tag{1.41}$$

- *The vertical load factor* n_z*:*

$$n_z = \frac{L}{mg} \tag{1.42}$$

- *The roll rate* p

Taking into account the components of the control vector $\underline{u}(t)$ *the dynamics of the aircraft model* (1.40) *reads as follows:*

$$\begin{cases} \dot{V} = g\left(n_x - \sin(\gamma)\right) \\ \dot{\gamma} = \frac{g}{V}\left(n_z \cos(\phi) - \cos(\gamma)\right) \\ \dot{\psi} = \frac{g}{V}\frac{\sin(\phi)}{\cos(\gamma)} n_z \\ \dot{\phi} = p \end{cases} \tag{1.43}$$

This is clearly a non-linear time-invariant state equation of the form:

$$\dot{\underline{x}} = f(\underline{x}, \underline{u}) \tag{1.44}$$

Where:

$$\begin{cases} \underline{x} = \begin{bmatrix} V & \gamma & \psi & \phi \end{bmatrix}^T \\ \underline{u} = \begin{bmatrix} n_x & n_z & p \end{bmatrix}^T \end{cases} \tag{1.45}$$

Let $(\underline{x}_e, \underline{u}_e)$ *be an equilibrium point defined by:*

$$f(\underline{x}_e, \underline{u}_e) = \underline{0} \tag{1.46}$$

The equilibrium point (or trim) for the aircraft model is obtained by arbitrarily setting the values of state vector $\underline{x}_e = \begin{bmatrix} V_e & \gamma_e & \psi_e & \phi_e \end{bmatrix}^T$ *which are airspeed, flight path angle, heading and bank angle, respectively. From that*

value of the state vector $\underline{x}_e$ *we get the value of the corresponding control vector* $\underline{u}_e = \begin{bmatrix} n_{xe} & n_{ze} & p_e \end{bmatrix}^T$ *by solving the following set of equations:*

$$\begin{cases} 0 = g\left(n_{xe} - \sin(\gamma_e)\right) \\ 0 = \frac{g}{V_e}\left(n_{ze}\cos(\phi_e) - \cos(\gamma_e)\right) \\ 0 = \frac{g}{V_e}\frac{\sin(\phi_e)}{\cos(\gamma_e)} n_{ze} \\ 0 = p_e \end{cases} \tag{1.47}$$

We get:

$$\begin{cases} p_e = 0 \\ \phi_e = 0 \\ n_{ze} = \frac{\cos(\gamma_e)}{\cos(\phi_e)} \quad \textit{here} \quad \phi_e = 0 \Rightarrow n_{ze} = \cos(\gamma_e) \\ n_{xe} = \sin(\gamma_e) \end{cases} \tag{1.48}$$

Let $\delta\underline{x}(t)$ *and* $\delta\underline{x}(t)$ *be defined as follows:*

$$\begin{cases} \underline{x}(t) = \underline{x}_e + \delta\underline{x}(t) \\ \underline{u}(t) = \underline{u}_e + \delta\underline{u}(t) \end{cases} \tag{1.49}$$

The linearization of the vector field f *around the equilibrium point* $(\underline{x}_e, \underline{u}_e)$ *reads:*

$$\delta\dot{\underline{x}}(t) \approx \left.\frac{\partial f(\underline{x},\underline{u})}{\partial \underline{x}}\right|_{\underline{u}=\underline{u}_e, \underline{x}=\underline{x}_e} \delta\underline{x}(t) + \left.\frac{\partial f(\underline{x},\underline{u})}{\partial \underline{u}}\right|_{\underline{u}=\underline{u}_e, \underline{x}=\underline{x}_e} \delta\underline{u}(t) \tag{1.50}$$

Assuming a level flight $(\gamma_e = 0)$ *we get the following expression of the state vector at the equilibrium:*

$$\underline{x}_e = \begin{bmatrix} V_e \\ \gamma_e = 0 \\ \psi_e \\ \phi_e = 0 \end{bmatrix} \tag{1.51}$$

Thus the control vector at the equilibrium reads:

$$\underline{u}_e = \begin{bmatrix} n_{xe} = \sin\left(\gamma_e\right) = 0 \\ n_{ze} = \cos\left(\gamma_e\right) = 1 \\ p_e = 0 \end{bmatrix} \tag{1.52}$$

Consequently:

$$\begin{aligned} \left.\frac{\partial f(\underline{x},\underline{u})}{\partial \underline{x}}\right|_{\underline{u}=\underline{u}_e, \underline{x}=\underline{x}_e} &= \left.\begin{bmatrix} 0 & -g\cos(\gamma) & 0 & 0 \\ -\frac{g}{V^2}\left(n_z\cos(\phi) - \cos(\gamma)\right) & \frac{g}{V}\sin(\gamma) & 0 & -\frac{g}{V}n_z\sin(\phi) \\ -\frac{g}{V^2}\frac{\sin(\phi)}{\cos(\gamma)}n_z & -\frac{g}{V}\frac{\sin(\phi)\sin(\gamma)}{\cos^2(\gamma)}n_z & 0 & \frac{g}{V}\frac{\cos(\phi)}{\cos(\gamma)}n_z \\ 0 & 0 & 0 & 0 \end{bmatrix}\right|_{\substack{\underline{x}=\underline{x}_e \\ \underline{u}=\underline{u}_e}} \\ &= \begin{bmatrix} 0 & -g & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{g}{V_e} \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \tag{1.53}$$

And:

$$\left.\frac{\partial f(\underline{x},\underline{u})}{\partial \underline{u}}\right|_{\underline{u}=\underline{u}_e,\underline{x}=\underline{x}_e} = \left.\begin{bmatrix} g & 0 & 0 \\ 0 & \frac{g}{V}\cos(\phi) & 0 \\ 0 & \frac{g}{V}\frac{\sin(\phi)}{\cos(\gamma)} & 0 \\ 0 & 0 & 1 \end{bmatrix}\right|_{\substack{V=V_e \\ \gamma=\gamma_e=0 \\ n_z=n_{ze}=\cos(\gamma_e)=1 \\ \phi=\phi_e=0}} = \begin{bmatrix} g & 0 & 0 \\ 0 & \frac{g}{V_e} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.54}$$

Finally using the fact that $\gamma_e = 0 \Rightarrow \delta\gamma = \gamma$, $\phi_e = 0 \Rightarrow \delta\phi = \phi$ *and* $p_e = 0 \Rightarrow \delta p = p$ *we get the following linear time-invariant state equation:*

$$\begin{bmatrix} \delta\dot{V} \\ \dot{\gamma} \\ \delta\dot{\psi} \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} 0 & -g & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{g}{V_e} \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta V \\ \gamma \\ \delta\psi \\ \phi \end{bmatrix} + \begin{bmatrix} g & 0 & 0 \\ 0 & \frac{g}{V_e} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta n_x \\ \delta n_z \\ p \end{bmatrix} \tag{1.55}$$

Obviously this is a state equation of the form $\delta\dot{\underline{x}}(t) = \mathbf{A}\delta\underline{x}(t) + \mathbf{B}\delta\underline{u}(t)$.

It can be seen that the linear aircraft model can be decoupled into longitudinal and lateral dynamics:

- *Longitudinal linearized dynamics:*

$$\begin{bmatrix} \delta\dot{V} \\ \dot{\gamma} \end{bmatrix} = \begin{bmatrix} 0 & -g \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta V \\ \delta\gamma \end{bmatrix} + \begin{bmatrix} g & 0 \\ 0 & \frac{g}{V_e} \end{bmatrix} \begin{bmatrix} \delta n_x \\ \delta n_z \end{bmatrix} \tag{1.56}$$

- *Lateral linearized dynamics:*

$$\begin{bmatrix} \delta\dot{\psi} \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} 0 & \frac{g}{V_e} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta\psi \\ \phi \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} p \tag{1.57}$$

The previous equations show that:

- *Airspeed variation is commanded by the longitudinal load factor* n_x*;*
- *Flight path angle variation is commanded by the vertical load factor* n_z*;*
- *Heading variation is commanded by the roll rate* p*.*

■

1.4 From state-space representation to transfer function

Let's consider the state-space representation (1.9) with state vector $\underline{x}(t)$, input vector $\underline{u}(t)$ and output vector $\underline{y}(t)$. The transfer function relates the relation

between the Laplace transform of the output vector, $\underline{Y}(s) = \mathcal{L}\left[\underline{y}(t)\right]$, and the Laplace transform of the input vector, $\underline{U}(s) = \mathcal{L}\left[\underline{u}(t)\right]$, assuming no initial condition, that is $\underline{x}(t)|_{t=0^+} = 0$. From (1.9) we get:

$$\underline{x}(t)|_{t=0^+} = 0 \Rightarrow \begin{cases} s\underline{X}(s) = \mathbf{A}\underline{X}(s) + \mathbf{B}\underline{U}(s) \\ \underline{Y}(s) = \mathbf{C}\underline{X}(s) + \mathbf{D}\underline{U}(s) \end{cases} \tag{1.58}$$

From the first equation of (1.58) we obtain the expression of the Laplace transform of the state vector (be careful to multiply s by the identity matrix to obtain a matrix with the same size than $\mathbf{A}$):

$$(s\mathbb{I} - \mathbf{A})\,\underline{X}(s) = \mathbf{B}\underline{U}(s) \Leftrightarrow \underline{X}(s) = (s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B}\underline{U}(s) \tag{1.59}$$

And using this result in the second equation of (1.58) leads to the expression of the transfer function $\mathbf{F}(s)$ of the system:

$$\underline{Y}(s) = \mathbf{C}\underline{X}(s) + \mathbf{D}\underline{U}(s) = \left(\mathbf{C}\,(s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B} + \mathbf{D}\right)\underline{U}(s) := \mathbf{F}(s)\underline{U}(s) \tag{1.60}$$

Where the transfer function $\mathbf{F}(s)$ of the system has the following expression:

$$\mathbf{F}(s) = \mathbf{C}\,(s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B} + \mathbf{D} \tag{1.61}$$

It is worth noticing that the denominator of the transfer function $\mathbf{F}(s)$ is also the determinant of matrix $s\mathbb{I} - \mathbf{A}$. Indeed the inverse of $s\mathbb{I} - \mathbf{A}$ is given by:

$$(s\mathbb{I} - \mathbf{A})^{-1} = \frac{1}{\det(s\mathbb{I} - \mathbf{A})}\,\mathrm{adj}(s\mathbb{I} - \mathbf{A}) \tag{1.62}$$

Where $\mathrm{adj}(s\mathbb{I} - \mathbf{A})$ is the adjugate of matrix $s\mathbb{I} - \mathbf{A}$ (that is the transpose of the matrix of cofactors [3]). Consequently, and assuming no pole-zero cancellation between $\mathrm{adj}(s\mathbb{I} - \mathbf{A})$ and $\det(s\mathbb{I} - \mathbf{A})$, the eigenvalues of matrix $\mathbf{A}$ are also the poles of the transfer function $\mathbf{F}(s)$.

From (1.62) it can be seen that the polynomials which form the numerator of $\mathbf{C}\,(s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B}$ have a degree which is strictly lower than the degree of $\det(s\mathbb{I} - \mathbf{A})$. Indeed the entry in the ith row and j^{th} column of the cofactor matrix of $s\mathbb{I} - \mathbf{A}$ (and thus the adjugate matrix) is formed by the determinant of the submatrix formed by deleting the ith row and j^{th} column of matrix $s\mathbb{I} - \mathbf{A}$; thus each determinant of those submatrices have a degree which is strictly lower than the degree of $\det(s\mathbb{I} - \mathbf{A})$. We say that $\mathbf{C}\,(s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B}$ is a strictly proper rational matrix which means that:

$$\lim_{s\to\infty} \mathbf{C}\,(s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B} = \mathbf{0} \tag{1.63}$$

In the general case of MIMO systems $\mathbf{F}(s)$ is a matrix of rational functions: the number of rows of $\mathbf{F}(s)$ is equal to the number of outputs of the system (that is the size of the output vector $\underline{y}(t)$) whereas the number of columns of $\mathbf{F}(s)$ is equal to the number of inputs of the system (that is the size of the input vector $\underline{u}(t)$).

[3] https://en.wikipedia.org/wiki/Invertible_matrix

1.5 Zeros of a transfer function - Rosenbrock's system matrix

Let $\mathbf{R}(s)$ be the so-called Rosenbrock's system matrix, as proposed in 1967 by Howard H. Rosenbrock[4]:

$$\mathbf{R}(s) = \begin{bmatrix} s\mathbb{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \tag{1.64}$$

From the fact that transfer function $\mathbf{F}(s)$ reads $\mathbf{F}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D}$, the following relation holds:

$$\begin{aligned} \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1} & \mathbb{I} \end{bmatrix} \mathbf{R}(s) &= \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1} & \mathbb{I} \end{bmatrix} \begin{bmatrix} s\mathbb{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \\ &= \begin{bmatrix} s\mathbb{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{0} & \mathbf{F}(s) \end{bmatrix} \end{aligned} \tag{1.65}$$

Matrix $\begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1} & \mathbb{I} \end{bmatrix}$ is a square matrix for which the following relation holds:

$$\det\left(\begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1} & \mathbb{I} \end{bmatrix}\right) = 1 \tag{1.66}$$

Now assume that $\mathbf{R}(s)$ is a square matrix. Using the property $\det\left(\mathbf{XY}\right) = \det\left(\mathbf{X}\right)\det\left(\mathbf{Y}\right)$, we get the following property for the Rosenbrock's system matrix $\mathbf{R}(s)$:

$$\begin{aligned} \det\left(\begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1} & \mathbb{I} \end{bmatrix} \mathbf{R}(s)\right) &= \det\left(\begin{bmatrix} s\mathbb{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{0} & \mathbf{F}(s) \end{bmatrix}\right) \\ \Rightarrow \det\left(\begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1} & \mathbb{I} \end{bmatrix}\right) \det\left(\mathbf{R}(s)\right) &= \det\left(s\mathbb{I} - \mathbf{A}\right)\det\left(\mathbf{F}(s)\right) \\ \Rightarrow \det\left(\mathbf{R}(s)\right) &= \det\left(s\mathbb{I} - \mathbf{A}\right)\det\left(\mathbf{F}(s)\right) \end{aligned} \tag{1.67}$$

For SISO systems we have $\det\left(\mathbf{F}(s)\right) = F(s)$ and consequently the preceding property reduces as follows:

$$\det\left(\mathbf{F}(s)\right) = F(s) \Rightarrow F(s) = \frac{\det\left(\mathbf{R}(s)\right)}{\det\left(s\mathbb{I} - \mathbf{A}\right)} \tag{1.68}$$

For non-square matrices, the Sylvester's rank inequality states that if $\mathbf{X}$ is a $m \times n$ matrix and $\mathbf{Y}$ is a $n \times k$ matrix, then the following relation holds:

$$\operatorname{rank}\left(\mathbf{X}\right) + \operatorname{rank}\left(\mathbf{Y}\right) - n \leq \operatorname{rank}\left(\mathbf{XY}\right) \leq \min\left(\operatorname{rank}\left(\mathbf{X}\right), \operatorname{rank}\left(\mathbf{Y}\right)\right) \tag{1.69}$$

For MIMO systems the transfer function between input i and output j is given by:

$$F_{ij}(s) = \frac{\det\left(\begin{bmatrix} s\mathbb{I} - \mathbf{A} & -\underline{b}_i \\ \underline{c}_j^T & d_{ij} \end{bmatrix}\right)}{\det(s\mathbb{I} - \mathbf{A})} \tag{1.70}$$

[4] https://en.wikipedia.org/wiki/Rosenbrock_system_matrix

where $\underline{b}_i$ is the ith column of $\mathbf{B}$ and $\underline{c}_j^T$ the jth row of $\mathbf{C}$.

Furthermore in the general case of MIMO linear time invariant systems, the (transmission) zeros of a transfer function $\mathbf{F}(s)$ are defined as the values of s such that the rank of the Rosenbrock's system matrix $\mathbf{R}(s) = \begin{bmatrix} s\mathbb{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$ is lower than its *normal* rank, meaning that the rank of $\mathbf{R}(s)$ drops.

When $\mathbf{R}(s)$ is a square matrix this means that $\mathbf{R}(s)$ is not invertible; in such a situation the (transmission) zeros are the values of s such that $\det(\mathbf{R}(s)) = 0$.

Furthermore when $\mathbf{R}(s)$ is a square matrix a (transmission) zero z in the transfer function $\mathbf{F}(s)$ indicates that there exists non-zero input vectors $\underline{u}(t)$ which produces a null output vector $\underline{y}(t)$. Let's write the state vector $\underline{x}(t)$ and input vector $\underline{u}(t)$ as follows where z is a (transmission) zero of the system:

$$\begin{cases} \underline{x}(t) = \underline{x}_0 e^{zt} \\ \underline{u}(t) = \underline{u}_0 e^{zt} \end{cases} \tag{1.71}$$

Imposing a null output vector $\underline{y}(t)$ we get from the state-space representation (1.9):

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \end{cases} \Leftrightarrow \begin{cases} z\underline{x}_0 e^{zt} = \mathbf{A}\underline{x}_0 e^{zt} + \mathbf{B}\underline{u}_0 e^{zt} \\ \underline{0} = \mathbf{C}\underline{x}_0 e^{zt} + \mathbf{D}\underline{u}_0 e^{zt} \end{cases} \tag{1.72}$$

That is:

$$\begin{cases} (z\mathbb{I} - \mathbf{A})\,\underline{x}_0 e^{zt} - \mathbf{B}\underline{u}_0 e^{zt} = \underline{0} \\ \mathbf{C}\underline{x}_0 e^{zt} + \mathbf{D}\underline{u}_0 e^{zt} = \underline{0} \end{cases} \Leftrightarrow \begin{bmatrix} s\mathbb{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}_{s=z} \begin{bmatrix} \underline{x}_0 \\ \underline{u}_0 \end{bmatrix} e^{zt} = \underline{0} \tag{1.73}$$

This relation holds for a non-zero input vector $\underline{u}(t) = \underline{u}_0 e^{zt}$ and a non-zero state vector $\underline{x}(t) = \underline{x}_0 e^{zt}$ when the values of s are chosen such that $\mathbf{R}(s)$ is not invertible ($\mathbf{R}(s)$ is assumed to be square); in such a situation the (transmission) zeros are the values of s such that $\det(\mathbf{R}(s)) = 0$. We thus retrieve Rosenbrock's result.

Example 1.3. *Let's consider the following state-space representation:*

$$\begin{cases} \dot{\underline{x}}(t) = \begin{bmatrix} -7 & -12 \\ 1 & 0 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \underline{u}(t) \\ \underline{y}(t) = \begin{bmatrix} 1 & 2 \end{bmatrix} \underline{x}(t) \end{cases} \tag{1.74}$$

From the identification with the general form of a state-space representation (1.9) *it is clear that* $\mathbf{D} = \mathbf{0}$. *Furthermore we get the following expression for the transfer function:*

$$\begin{aligned} F(s) &= \mathbf{C}\,(s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B} \\ &= \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} s+7 & 12 \\ -1 & s \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \end{bmatrix} \tfrac{1}{s(s+7)+12} \begin{bmatrix} s & -12 \\ 1 & s+7 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \tfrac{1}{s^2+7s+12} \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} s \\ 1 \end{bmatrix} \\ &= \tfrac{s+2}{s^2+7s+12} \end{aligned} \tag{1.75}$$

It can be checked the denominator of the transfer function $F(s)$ is also the determinant of matrix $s\mathbb{I}-\mathbf{A}$.

$$\det(s\mathbb{I}-\mathbf{A}) = \det\left(\begin{bmatrix} s+7 & 12 \\ -1 & s \end{bmatrix}\right) = s^2+7s+12 \tag{1.76}$$

Furthermore as far as $F(s)$ is the transfer function of a SISO system it can also be checked that its numerator of can be obtained thanks to the following relation:

$$\det\left(\begin{bmatrix} s\mathbb{I}-\mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}\right) = \det\left(\begin{bmatrix} s+7 & 12 & -1 \\ -1 & s & 0 \\ 1 & 2 & 0 \end{bmatrix}\right) = s+2 \tag{1.77}$$

Thus the only (transmission) zero for this system is $s=-2$.

■

1.6 Faddeev-Leverrier's method to get $(s\mathbb{I}-\mathbf{A})^{-1}$

Let $\mathbf{A}$ be a $n\times n$ matrix with coefficients in $\mathbb{R}$. Then matrix $(s\mathbb{I}-\mathbf{A})^{-1}$, which is called the *resolvent* of $\mathbf{A}$, may be obtained by a method proposed by D.K. Faddeev (Dmitrii Konstantinovitch Faddeev, 1907 - 1989, was a Russian mathematician). This is a modification of a method proposed by U.J.J. Leverrier (Urbain Jean Joseph Le Verrier, 1811 - 1877, was a French mathematician who specialized in celestial mechanics and is best known for predicting the existence and position of Neptune using only mathematics [5]). The starting point of the method is to relate the resolvent of matrix $\mathbf{A}$ to its characteristic polynomial $\det(s\mathbb{I}-\mathbf{A})$ through the following relation:

$$(s\mathbb{I}-\mathbf{A})^{-1} = \frac{\mathbf{N}(s)}{\det(s\mathbb{I}-\mathbf{A})} = \frac{\mathbf{F}_0 s^{n-1}+\mathbf{F}_1 s^{n-2}+\cdots+\mathbf{F}_{n-1}}{s^n-d_1 s^{n-1}-\cdots-d_n} \tag{1.78}$$

where the adjugate matrix $\mathbf{N}(s)$ is a polynomial matrix in s of degree $n-1$ with constant $n\times n$ coefficient matrices $\mathbf{F}_0,\cdots,\mathbf{F}_{n-1}$.

The Faddeev-Leverrier's method indicates that the n matrices $\mathbf{F}_k$ and coefficients d_k in (1.78) can be computed recursively as follows:

$$\begin{cases} \mathbf{F}_0 = \mathbb{I} & \\ d_1 = \operatorname{tr}(\mathbf{A}\mathbf{F}_0) & \text{and } \mathbf{F}_1 = \mathbf{A}\mathbf{F}_0 - d_1\mathbb{I} \\ d_2 = \frac{1}{2}\operatorname{tr}(\mathbf{A}\mathbf{F}_1) & \text{and } \mathbf{F}_2 = \mathbf{A}\mathbf{F}_1 - d_2\mathbb{I} \\ \vdots & \\ d_k = \frac{1}{k}\operatorname{tr}(\mathbf{A}\mathbf{F}_{k-1}) & \text{and } \mathbf{F}_k = \mathbf{A}\mathbf{F}_{k-1} - d_k\mathbb{I} \\ \vdots & \\ d_n = \frac{1}{n}\operatorname{tr}(\mathbf{A}\mathbf{F}_{n-1}) & \end{cases} \tag{1.79}$$

$$\text{and } \det(s\mathbb{I}-\mathbf{A}) = s^n - d_1 s^{n-1} - \cdots - d_n$$

[5] https://en.wikipedia.org/wiki/Urbain_Le_Verrier

To arrive at the Faddeev-Leverrier's method we shall compare coefficients of like powers of s in the following formula which is derived from (1.78):

$$(s\mathbb{I} - \mathbf{A})\left(\mathbf{F}_0 s^{n-1} + \mathbf{F}_1 s^{n-2} + \cdots + \mathbf{F}_{n-1}\right) = \mathbb{I}\left(s^n - d_1 s^{n-1} - \cdots - d_n\right) \quad (1.80)$$

and obtain immediately that matrices $\mathbf{F}_k$ are given by:

$$\begin{cases} \mathbf{F}_0 = \mathbb{I} \\ \mathbf{F}_1 = \mathbf{A}\mathbf{F}_0 - d_1\mathbb{I} \\ \mathbf{F}_2 = \mathbf{A}\mathbf{F}_1 - d_2\mathbb{I} \\ \vdots \\ \mathbf{F}_k = \mathbf{A}\mathbf{F}_{k-1} - d_k\mathbb{I} \end{cases} \quad (1.81)$$

The rest of the proof can be found in the paper of Shui-Hung Hou [6].

Example 1.4. *Compute the resolvent of matrix* $\mathbf{A}$ *where:*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (1.82)$$

Matrix $\mathbf{A}$ *is a* 2×2 *matrix. The Faddeev-Leverrier's method gives:*

$$\begin{cases} \mathbf{F}_0 = \mathbb{I} \\ d_1 = \operatorname{tr}(\mathbf{A}\mathbf{F}_0) = \operatorname{tr}(\mathbf{A}) = 0 \text{ and } \mathbf{F}_1 = \mathbf{A}\mathbf{F}_0 - d_1\mathbb{I} = \mathbf{A} \\ d_2 = \frac{1}{2}\operatorname{tr}(\mathbf{A}\mathbf{F}_1) = \frac{1}{2}\operatorname{tr}\left(\mathbf{A}^2\right) = 0 \end{cases} \quad (1.83)$$
$$\text{and } \det(s\mathbb{I} - \mathbf{A}) = s^2 - d_1 s - d_2 = s^2$$

Then:

$$(s\mathbb{I} - \mathbf{A})^{-1} = \frac{\mathbf{F}_0 s + \mathbf{F}_1}{\det(s\mathbb{I} - \mathbf{A})} = \frac{1}{s^2}\begin{bmatrix} s & 1 \\ 0 & s \end{bmatrix} = \begin{bmatrix} \frac{1}{s} & \frac{1}{s^2} \\ 0 & \frac{1}{s^2} \end{bmatrix} \quad (1.84)$$

■

Example 1.5. *Compute the resolvent of matrix* $\mathbf{A}$ *where:*

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & -5 \end{bmatrix} \quad (1.85)$$

Matrix $\mathbf{A}$ *is a* 2×2 *matrix. The Faddeev-Leverrier's method gives:*

$$\begin{cases} \mathbf{F}_0 = \mathbb{I} \\ d_1 = \operatorname{tr}(\mathbf{A}\mathbf{F}_0) = -4 \text{ and } \mathbf{F}_1 = \mathbf{A}\mathbf{F}_0 - d_1\mathbb{I} = \begin{bmatrix} 5 & 2 \\ 0 & -1 \end{bmatrix} \\ d_2 = \frac{1}{2}\operatorname{tr}(\mathbf{A}\mathbf{F}_1) = \frac{1}{2}\operatorname{tr}\left(\begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}\right) = 5 \end{cases} \quad (1.86)$$
$$\text{and } \det(s\mathbb{I} - \mathbf{A}) = s^2 - d_1 s - d_2 = s^2 + 4s - 5 = (s-1)(s+5)$$

[6] Shui-Hung Hou, A Simple Proof of the Leverrier-Faddeev Characteristic Polynomial Algorithm, SIAM Review, Vol. 40, No. 3 (Sep., 1998), pp. 706-709

Then:

$$\begin{aligned}(s\mathbb{I}-\mathbf{A})^{-1} &= \tfrac{\mathbf{F}_0 s+\mathbf{F}_1}{\det(s\mathbb{I}-\mathbf{A})} = \tfrac{1}{(s-1)(s+5)}\begin{bmatrix} s+5 & 2 \\ 0 & s-1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{s-1} & \frac{2}{(s-1)(s+5)} \\ 0 & \frac{1}{s+5} \end{bmatrix}\end{aligned} \tag{1.87}$$

■

Example 1.6. *Compute the resolvent of matrix* $\mathbf{A}$ *where:*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -2m\omega_0 \end{bmatrix} \tag{1.88}$$

Matrix $\mathbf{A}$ *is a* 2×2 *matrix. The Faddeev-Leverrier's method gives:*

$$\begin{cases} \mathbf{F}_0 = \mathbb{I} \\ d_1 = \operatorname{tr}(\mathbf{A}\mathbf{F}_0) = -2m\omega_0 \text{ and } \mathbf{F}_1 = \mathbf{A}\mathbf{F}_0 - d_1\mathbb{I} = \begin{bmatrix} 2m\omega_0 & 1 \\ -\omega_0^2 & 0 \end{bmatrix} \\ d_2 = \frac{1}{2}\operatorname{tr}(\mathbf{A}\mathbf{F}_1) = \frac{1}{2}\operatorname{tr}\left(\begin{bmatrix} -\omega_0^2 & 0 \\ -4m\omega_0^3 & -\omega_0^2 \end{bmatrix}\right) = -\omega_0^2 \end{cases} \tag{1.89}$$
$$\text{and } \det(s\mathbb{I}-\mathbf{A}) = s^2 - d_1 s - d_2 = s^2 + 2m\omega_0 s + \omega_0^2$$

Then:

$$\begin{aligned}(s\mathbb{I}-\mathbf{A})^{-1} &= \tfrac{\mathbf{F}_0 s+\mathbf{F}_1}{\det(s\mathbb{I}-\mathbf{A})} \\ &= \tfrac{1}{s^2+2m\omega_0 s+\omega_0^2}\begin{bmatrix} s+2m\omega_0 & 1 \\ -\omega_0^2 & s \end{bmatrix}\end{aligned} \tag{1.90}$$

■

Example 1.7. *Compute the resolvent of matrix* $\mathbf{A}$ *where:*

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \tag{1.91}$$

Matrix $\mathbf{A}$ *is a* 3×3 *matrix. The Faddeev-Leverrier's method gives:*

$$\begin{cases} \mathbf{F}_0 = \mathbb{I} \\ d_1 = \operatorname{tr}(\mathbf{A}\mathbf{F}_0) = 4 \text{ and } \mathbf{F}_1 = \mathbf{A}\mathbf{F}_0 - d_1\mathbb{I} = \begin{bmatrix} -2 & -1 & 0 \\ 0 & -3 & 0 \\ 1 & -1 & -3 \end{bmatrix} \\ d_2 = \frac{1}{2}\operatorname{tr}(\mathbf{A}\mathbf{F}_1) = \frac{1}{2}\operatorname{tr}\left(\begin{bmatrix} -4 & 1 & 0 \\ 0 & -3 & 0 \\ -1 & 1 & -3 \end{bmatrix}\right) = -5 \\ \mathbf{F}_2 = \mathbf{A}\mathbf{F}_1 - d_2\mathbb{I} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ -1 & 1 & 2 \end{bmatrix} \\ d_3 = \frac{1}{3}\operatorname{tr}(\mathbf{A}\mathbf{F}_2) = \frac{1}{3}\operatorname{tr}\left(\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\right) = 2 \end{cases} \tag{1.92}$$
$$\text{and } \det(s\mathbb{I}-\mathbf{A}) = s^3 - d_1 s^2 - d_2 s - d_3 = s^3 - 4s^2 + 5s - 2$$

Then:

$$
\begin{aligned}
(s\mathbb{I}-\mathbf{A})^{-1} &= \tfrac{\mathbf{F}_0 s^2+\mathbf{F}_1 s+\mathbf{F}_2}{\det(s\mathbb{I}-\mathbf{A})} \\
&= \tfrac{1}{s^3-4s^2+5s-2}\begin{bmatrix} s^2-2s+1 & -s+1 & 0 \\ 0 & s^2-3s+2 & 0 \\ s-1 & -s+1 & s^2-3s+2 \end{bmatrix} \\
&= \tfrac{1}{(s-2)(s-1)^2}\begin{bmatrix} (s-1)^2 & -(s-1) & 0 \\ 0 & (s-2)(s-1) & 0 \\ s-1 & -(s-1) & (s-2)(s-1) \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{s-2} & \frac{-1}{(s-1)(s-2)} & 0 \\ 0 & \frac{1}{s-1} & 0 \\ \frac{1}{(s-1)(s-2)} & \frac{-1}{(s-1)(s-2)} & \frac{1}{s-1} \end{bmatrix}
\end{aligned}
\tag{1.93}
$$

■

1.7 Matrix inversion lemma

Assuming that $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ are invertible matrices, the inversion of a partitioned matrix reads as follows:

$$
\begin{aligned}
\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}^{-1} &= \begin{bmatrix} \mathbf{Q}_1 & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{Q}_2 \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{Q}_1 & -\mathbf{Q}_1\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{Q}_2\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{Q}_2 \end{bmatrix}
\end{aligned}
\tag{1.94}
$$

where:

$$
\begin{cases} \mathbf{Q}_1 = \left(\mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right)^{-1} \\ \mathbf{Q}_2 = \left(\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right)^{-1} \end{cases}
\tag{1.95}
$$

We can check that:

$$
\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}\begin{bmatrix} \mathbf{Q}_1 & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{Q}_2 \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} = \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{bmatrix}
\tag{1.96}
$$

and that:

$$
\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}\begin{bmatrix} \mathbf{Q}_1 & -\mathbf{Q}_1\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{Q}_2\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{Q}_2 \end{bmatrix} = \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{bmatrix}
\tag{1.97}
$$

Matrix inversion formula can be used to compute the *resolvent* of $\mathbf{A}$, that is matrix $(s\mathbb{I}-\mathbf{A})^{-1}$.

From the preceding relations the matrix inversion lemma reads as follows:

$$
\begin{aligned}
&\left(\mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right)^{-1} \\
&\qquad = \mathbf{A}_{11}^{-1}+\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\left(\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right)^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}
\end{aligned}
\tag{1.98}
$$

In the particular case of upper triangular matrix where $\mathbf{A}_{21} = \mathbf{0}$, the preceding relations simplify as follows:

$$
\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0} & \mathbf{A}_{22}^{-1} \end{bmatrix}
\tag{1.99}
$$

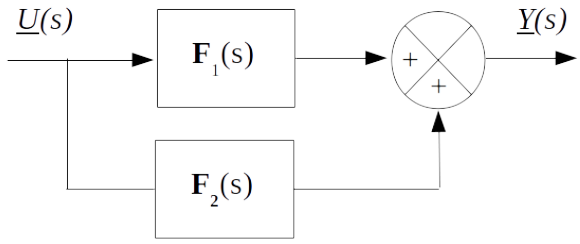

Figure 1.2: Parallel interconnection of systems

1.8 Interconnection of systems

We will consider in the following the state-space representation resulting from different systems interconnection. This will be useful to get the state-space representation of complex models.

Lets consider two linear time-invariant system with transfer functions $\mathbf{F}_1(s)$ and $\mathbf{F}_2(s)$ and state-space representations $\left(\begin{array}{c|c}\mathbf{A}_1 & \mathbf{B}_1 \\ \hline \mathbf{C}_1 & \mathbf{D}_1\end{array}\right)$ and $\left(\begin{array}{c|c}\mathbf{A}_2 & \mathbf{B}_2 \\ \hline \mathbf{C}_2 & \mathbf{D}_2\end{array}\right)$:

$$\left\{\begin{array}{l}\dot{\underline{x}}_1(t) = \mathbf{A}_1\underline{x}_1(t) + \mathbf{B}_1\underline{u}_1(t) \\ \underline{y}_1(t) = \mathbf{C}_1\underline{x}_1(t) + \mathbf{D}_1\underline{u}_1(t)\end{array}\right. \quad \text{and} \quad \left\{\begin{array}{l}\dot{\underline{x}}_2(t) = \mathbf{A}_2\underline{x}_2(t) + \mathbf{B}_2\underline{u}_2(t) \\ \underline{y}_2(t) = \mathbf{C}_2\underline{x}_2(t) + \mathbf{D}_2\underline{u}_2(t)\end{array}\right. \tag{1.100}$$

The state vector attached to the interconnection of two systems, whatever the type of interconnection, is the vector $\underline{x}(t)$ defined by:

$$\underline{x}(t) = \begin{bmatrix} \underline{x}_1(t) \\ \underline{x}_2(t) \end{bmatrix} \tag{1.101}$$

The output of the interconnection is denoted $\underline{y}(t)$ whereas the input is denoted $\underline{u}(t)$.

1.8.1 Parallel interconnection

Parallel interconnection is depicted on Figure 1.2. The transfer function $\mathbf{F}(s)$ of the parallel interconnection between two systems with transfer function $\mathbf{F}_1(s)$ and $\mathbf{F}_2(s)$ is:

$$\mathbf{F}(s) = \mathbf{F}_1(s) + \mathbf{F}_2(s) \tag{1.102}$$

Parallel interconnection is obtained when both systems have a common input and by summing the outputs assuming that the dimension of the outputs fit:

$$\left\{\begin{array}{l}\underline{u}(t) = \underline{u}_1(t) = \underline{u}_2(t) \\ \underline{y}(t) = \underline{y}_1(t) + \underline{y}_2(t)\end{array}\right. \tag{1.103}$$

The state-space representation of the parallel interconnection is the following:

$$\left\{\begin{array}{l}\dot{\underline{x}}(t) = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix}\underline{x}(t) + \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix}\underline{u}(t) \\ \underline{y}(t) = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix}\underline{x}(t) + (\mathbf{D}_1 + \mathbf{D}_2)\,\underline{u}(t)\end{array}\right. \tag{1.104}$$

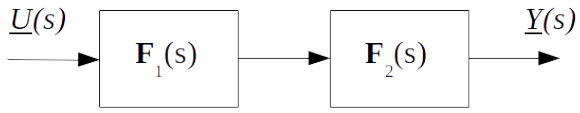

Figure 1.3: Series interconnection of systems

This result can also be easily retrieved by summing the realization of each transfer function:

$$\begin{aligned}
\mathbf{F}(s) &= \mathbf{F}_1(s) + \mathbf{F}_2(s) \\
&= \mathbf{C}_1 \left(s\mathbb{I} - \mathbf{A}_1\right)^{-1} \mathbf{B}_1 + \mathbf{D}_1 + \mathbf{C}_2 \left(s\mathbb{I} - \mathbf{A}_2\right)^{-1} \mathbf{B}_2 + \mathbf{D}_2 \\
&= \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix} \begin{bmatrix} \left(s\mathbb{I} - \mathbf{A}_1\right)^{-1} & \mathbf{0} \\ \mathbf{0} & \left(s\mathbb{I} - \mathbf{A}_2\right)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} + \mathbf{D}_1 + \mathbf{D}_2 \\
&= \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix} \begin{bmatrix} s\mathbb{I} - \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & s\mathbb{I} - \mathbf{A}_2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} + \mathbf{D}_1 + \mathbf{D}_2 \\
&= \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix} \left(s\mathbb{I} - \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix}\right)^{-1} \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} + \mathbf{D}_1 + \mathbf{D}_2
\end{aligned} \tag{1.105}$$

The preceding relation indicates that the realization of the sum $\mathbf{F}_1(s)+\mathbf{F}_2(s)$ of two transfer functions is:

$$\mathbf{F}_1(s) + \mathbf{F}_2(s) = \left(\begin{array}{cc|c} \mathbf{A}_1 & \mathbf{0} & \mathbf{B}_1 \\ \mathbf{0} & \mathbf{A}_2 & \mathbf{B}_2 \\ \hline \mathbf{C}_1 & \mathbf{C}_2 & \mathbf{D}_1 + \mathbf{D}_2 \end{array}\right) \tag{1.106}$$

1.8.2 Series interconnection

Series interconnection is depicted on Figure 1.3. The transfer function $\mathbf{F}(s)$ of the series interconnection between two systems with transfer function $\mathbf{F}_1(s)$ and $\mathbf{F}_2(s)$ is:

$$\mathbf{F}(s) = \mathbf{F}_2(s)\mathbf{F}_1(s) \tag{1.107}$$

Series interconnection is obtained when the output of the first system enters the second system as an input:

$$\begin{cases} \underline{u}_2(t) = \underline{y}_1(t) \\ \underline{y}(t) = \underline{y}_2(t) \\ \underline{u}(t) = \underline{u}_1(t) \end{cases} \tag{1.108}$$

The state-space representation of the series interconnection is the following:

$$\begin{cases} \dot{\underline{x}}(t) = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{B}_2\mathbf{C}_1 & \mathbf{A}_2 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2\mathbf{D}_1 \end{bmatrix} \underline{u}(t) \\ \underline{y}(t) = \begin{bmatrix} \mathbf{D}_2\mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix} \underline{x}(t) + \mathbf{D}_2\mathbf{D}_1\underline{u}(t) \end{cases} \tag{1.109}$$

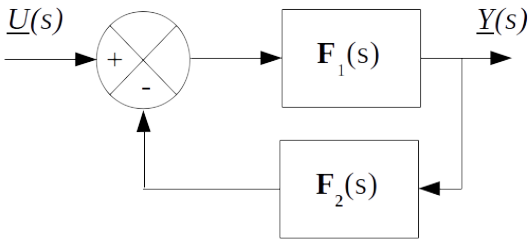

Figure 1.4: Feedback interconnection of systems

1.8.3 Feedback interconnection

Feedback interconnection is depicted on Figure 1.4. To get the transfer function $\mathbf{F}(s)$ of the feedback interconnection between two systems with transfer function $\mathbf{F}_1(s)$ and $\mathbf{F}_2(s)$ we write the relation between the Laplace transform $\underline{Y}(s)$ of the output vector and the Laplace transform of input vector $\underline{U}(s)$:

$$\begin{aligned} &\underline{Y}(s) = \mathbf{F}_1(s)\left(\underline{U}(s) - \mathbf{F}_2(s)\underline{Y}(s)\right) \\ &\Leftrightarrow \left(\mathbb{I} - \mathbf{F}_1(s)\mathbf{F}_2(s)\right)\underline{Y}(s) = \mathbf{F}_1(s)\underline{U}(s) \\ &\Leftrightarrow \underline{Y}(s) = \left(\mathbb{I} - \mathbf{F}_1(s)\mathbf{F}_2(s)\right)^{-1}\mathbf{F}_1(s)\underline{U}(s) \end{aligned} \tag{1.110}$$

We finally get:

$$\mathbf{F}(s) = \left(\mathbb{I} - \mathbf{F}_1(s)\mathbf{F}_2(s)\right)^{-1}\mathbf{F}_1(s) \tag{1.111}$$

As depicted on Figure 1.4 feedback interconnection is obtained when the output of the first system enters the second system as an input and by feeding the first system by the difference between the system input $\underline{u}(t)$ and the output of the second system (assuming that the dimension fit):

$$\begin{cases} \underline{u}_1(t) = \underline{u}(t) - \underline{y}_2(t) \Leftrightarrow \underline{u}(t) = \underline{u}_1(t) + \underline{y}_2(t) \\ \underline{y}(t) = \underline{y}_1(t) \\ \underline{u}_2(t) = \underline{y}_1(t) \end{cases} \tag{1.112}$$

Thus the state-space representation of the feedback interconnection is the following:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}_f\underline{x}(t) + \begin{bmatrix} \mathbf{B}_1 - \mathbf{B}_1\mathbf{D}_2\mathbf{M}\mathbf{D}_1 \\ \mathbf{B}_2\mathbf{D}_1 - \mathbf{B}_2\mathbf{D}_1\mathbf{D}_2\mathbf{M}\mathbf{D}_1 \end{bmatrix}\underline{u}(t) \\ \mathbf{A}_f = \begin{bmatrix} \mathbf{A}_1 - \mathbf{B}_1\mathbf{D}_2\mathbf{M}\mathbf{C}_1 & -\mathbf{B}_1\mathbf{C}_2 + \mathbf{B}_1\mathbf{D}_2\mathbf{M}\mathbf{D}_1\mathbf{C}_2 \\ \mathbf{B}_2\mathbf{C}_1 - \mathbf{B}_2\mathbf{D}_1\mathbf{D}_2\mathbf{M}\mathbf{C}_1 & \mathbf{A}_2 - \mathbf{B}_2\mathbf{D}_1\mathbf{C}_2 + \mathbf{B}_2\mathbf{D}_1\mathbf{D}_2\mathbf{M}\mathbf{D}_1\mathbf{C}_2 \end{bmatrix} \\ \mathbf{M} = \left(\mathbb{I} + \mathbf{D}_1\mathbf{D}_2\right)^{-1} \\ \underline{y}(t) = \mathbf{M}\Big(\begin{bmatrix} \mathbf{C}_1 & -\mathbf{D}_1\mathbf{C}_2 \end{bmatrix}\underline{x}(t) + \mathbf{D}_1\underline{u}(t)\Big) \end{cases} \tag{1.113}$$

In the special case of an unity feedback we have:

$$\mathbf{F}_2(s) = \mathbb{I} \Leftrightarrow \left(\begin{array}{c|c} \mathbf{A}_2 & \mathbf{B}_2 \\ \hline \mathbf{C}_2 & \mathbf{D}_2 \end{array}\right) = \left(\begin{array}{c|c} \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{K}_2 \end{array}\right) \tag{1.114}$$

Thus the preceding relations reduce as follows:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}_f \underline{x}(t) + \begin{bmatrix} \mathbf{B}_1 - \mathbf{B}_1\mathbf{K}_2 \left(\mathbb{I} + \mathbf{D}_1\right)^{-1} \mathbf{D}_1 \\ \mathbf{0} \end{bmatrix} \underline{u}(t) \\ \mathbf{A}_f = \begin{bmatrix} \mathbf{A}_1 - \mathbf{B}_1\mathbf{K}_2 \left(\mathbb{I} + \mathbf{D}_1\mathbf{K}_2\right)^{-1} \mathbf{C}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \\ \underline{y}(t) = \left(\mathbb{I} + \mathbf{D}_1\mathbf{K}_2\right)^{-1} \left(\begin{bmatrix} \mathbf{C}_1 & \mathbf{0} \end{bmatrix} \underline{x}(t) + \mathbf{D}_1 \underline{u}(t) \right) \end{cases} \tag{1.115}$$

It is clear from the preceding equation that the state vector of the system reduces to its first component $\underline{x}_1(t)$. Thus the preceding state-space realization reads:

$$\begin{cases} \dot{\underline{x}}_1(t) = \left(\mathbf{A}_1 - \mathbf{B}_1\mathbf{K}_2 \left(\mathbb{I} + \mathbf{D}_1\mathbf{K}_2\right)^{-1} \mathbf{C}_1\right) \underline{x}_1(t) + \left(\mathbf{B}_1 - \mathbf{B}_1\mathbf{K}_2 \left(\mathbb{I} + \mathbf{D}_1\right)^{-1} \mathbf{D}_1\right) \underline{u}(t) \\ \underline{y}(t) = \left(\mathbb{I} + \mathbf{D}_1\mathbf{K}_2\right)^{-1} \left(\mathbf{C}_1 \underline{x}_1(t) + \mathbf{D}_1 \underline{u}(t)\right) \end{cases} \tag{1.116}$$

Chapter 2

Realization of transfer functions

2.1 Introduction

A realization of a transfer function $\mathbf{F}(s)$ consists in finding a state-space model given the input-output description of the system through its transfer function. More specifically we call realization of a transfer function $\mathbf{F}(s)$ any quadruplet $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ such that:

$$\mathbf{F}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} \tag{2.1}$$

We said that a transfer function $\mathbf{F}(s)$ is realizable if $\mathbf{F}(s)$ is rational and proper. The state-space representation of a transfer function $\mathbf{F}(s)$ is then:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{2.2}$$

This chapter focuses on canonical realizations of transfer functions that are the controllable canonical form, the observable canonical form and the diagonal (or modal) form. Realization of SISO (Single-Input Single Output), SIMO (Single-Input Multiple-Outputs) and MIMO (Multiple-Inputs Multiple-Outputs) linear time invariant systems will be presented.

2.2 Non-unicity of state-space representation

2.2.1 Similarity transformations

Contrary to linear differential equation or transfer function which describe the dynamics of a system in a single manner the state-space representation of a system is not unique. Indeed they are several ways to choose the internal variables which describe the dynamics of the system, that is the state vector $\underline{x}(t)$, without changing the input-output representation of the system, that is both the differential equation and the transfer function.

To be more specific let's consider the state-space representation (2.2) with state vector $\underline{x}(t)$. Then choose a similarity transformation with an invertible change of basis matrix $\mathbf{P}_n$ which defines a new state vector $\underline{x}_n(t)$ as follows:

$$\underline{x}(t) = \mathbf{P}_n\underline{x}_n(t) \Leftrightarrow \underline{x}_n(t) = \mathbf{P}_n^{-1}\underline{x}(t) \tag{2.3}$$

Then take the time derivative of $\underline{x}_n(t)$:

$$\dot{\underline{x}}_n(t) = \mathbf{P}_n^{-1}\dot{\underline{x}}(t) \tag{2.4}$$

The time derivative of $\underline{x}(t)$ is obtained thanks to (2.2). By replacing $\underline{x}(t)$ by $\underline{x}_n(t)$ we get:

$$\dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) = \mathbf{A}\mathbf{P}_n\underline{x}_n(t) + \mathbf{B}\underline{u}(t) \tag{2.5}$$

Thus we finally get:

$$\begin{cases} \dot{\underline{x}}_n(t) = \mathbf{P}_n^{-1}\dot{\underline{x}}(t) = \mathbf{P}_n^{-1}\mathbf{A}\mathbf{P}_n\underline{x}_n(t) + \mathbf{P}_n^{-1}\mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\mathbf{P}_n\underline{x}_n(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{2.6}$$

We can match the preceding equations with the general form of a state-space representation (2.2) by rewriting it as follows:

$$\begin{cases} \dot{\underline{x}}_n(t) = \mathbf{A}_n\underline{x}_n(t) + \mathbf{B}_n\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}_n\underline{x}_n(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{2.7}$$

Where:

$$\begin{cases} \mathbf{A}_n = \mathbf{P}_n^{-1}\mathbf{A}\mathbf{P}_n \\ \mathbf{B}_n = \mathbf{P}_n^{-1}\mathbf{B} \\ \mathbf{C}_n = \mathbf{C}\mathbf{P}_n \end{cases} \tag{2.8}$$

It is worth noticing that the feedforward matrix $\mathbf{D}$ is independent of the choice of the state vector.

Now let's focus on the transfer function. With the new state vector $\underline{x}_n(t)$ the transfer function $\mathbf{F}(s)$ has the following expression:

$$\mathbf{F}(s) = \mathbf{C}_n\left(s\mathbb{I} - \mathbf{A}_n\right)^{-1}\mathbf{B}_n + \mathbf{D} \tag{2.9}$$

Using the expressions of (2.8) to express $\mathbf{A}_n$, $\mathbf{B}_n$ and $\mathbf{C}_n$ as a function of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ we get:

$$\mathbf{F}(s) = \mathbf{C}\mathbf{P}_n\left(s\mathbb{I} - \mathbf{P}_n^{-1}\mathbf{A}\mathbf{P}_n\right)^{-1}\mathbf{P}_n^{-1}\mathbf{B} + \mathbf{D} \tag{2.10}$$

Now use the fact that $\mathbb{I} = \mathbf{P}_n^{-1}\mathbf{P}_n$ and that $(\mathbf{XYZ})^{-1} = \mathbf{Z}^{-1}\mathbf{Y}^{-1}\mathbf{X}^{-1}$ (as soon as matrices $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are invertible) to get:

$$\begin{aligned} \mathbf{F}(s) &= \mathbf{C}\mathbf{P}_n\left(s\mathbf{P}_n^{-1}\mathbf{P}_n - \mathbf{P}_n^{-1}\mathbf{A}\mathbf{P}_n\right)^{-1}\mathbf{P}_n^{-1}\mathbf{B} + \mathbf{D} \\ &= \mathbf{C}\mathbf{P}_n\left(\mathbf{P}_n^{-1}\left(s\mathbb{I} - \mathbf{A}\right)\mathbf{P}_n\right)^{-1}\mathbf{P}_n^{-1}\mathbf{B} + \mathbf{D} \\ &= \mathbf{C}\mathbf{P}_n\mathbf{P}_n^{-1}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{P}_n\mathbf{P}_n^{-1}\mathbf{B} + \mathbf{D} \\ &= \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} \end{aligned} \tag{2.11}$$

We obviously retrieve the expression of the transfer function $\mathbf{F}(s)$ given by matrices $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$. Thus the expression of the transfer function is independent of the choice of the state vector.

2.2.2 Inverse of a similarity transformation

Let $\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n$ be the n vectors which form matrix $\mathbf{P}_n$:

$$\mathbf{P}_n = \begin{bmatrix} \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_n \end{bmatrix} \tag{2.12}$$

As far as matrix $\mathbf{P}_n$ is invertible vectors $\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n$ are independent. Let $\underline{x}(t) = \mathbf{P}_n \underline{x}_n(t)$. Denoting by $x_{n1}, x_{n2}, \cdots, x_{nn}$ the components of vector $\underline{x}_n(t)$ we get:

$$\underline{x}_n(t) = \begin{bmatrix} x_{n1} \\ x_{n2} \\ \vdots \\ x_{nn} \end{bmatrix} \Rightarrow \underline{x}(t) = \mathbf{P}_n \underline{x}_n(t) = x_{n1}\underline{v}_1 + x_{n2}\underline{v}_2 + \cdots + x_{nn}\underline{v}_n \tag{2.13}$$

Thus the state vector $\underline{x}(t)$ can be decomposed along the components of the change of basis matrix $\mathbf{P}_n$.

The inverse of the change of basis matrix $\mathbf{P}_n$ can be written in terms of rows as follows:

$$\mathbf{P}_n^{-1} = \begin{bmatrix} \underline{w}_1^T \\ \underline{w}_2^T \\ \vdots \\ \underline{w}_n^T \end{bmatrix} \tag{2.14}$$

Since $\mathbf{P}_n^{-1}\mathbf{P}_n = \mathbb{I}$ it follows that:

$$\mathbf{P}_n^{-1}\mathbf{P}_n = \begin{bmatrix} \underline{w}_1^T\underline{v}_1 & \underline{w}_1^T\underline{v}_2 & \cdots & \underline{w}_1^T\underline{v}_n \\ \underline{w}_2^T\underline{v}_1 & \underline{w}_2^T\underline{v}_2 & \cdots & \underline{w}_2^T\underline{v}_n \\ \vdots & \vdots & & \vdots \\ \underline{w}_n^T\underline{v}_1 & \underline{w}_n^T\underline{v}_2 & \cdots & \underline{w}_n^T\underline{v}_n \end{bmatrix} = \mathbb{I} \tag{2.15}$$

Hence the relation between vectors $\underline{w}_i$ and $\underline{v}_j$ is the following:

$$\underline{w}_i^T\underline{v}_j = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} \tag{2.16}$$

2.3 Realization of SISO transfer function

We have seen that a given transfer function $\mathbf{F}(s)$ can be obtained by an infinity number of state-space representations. We call realization of a transfer function $\mathbf{F}(s)$ any quadruplet $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ such that:

$$\mathbf{F}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} \tag{2.17}$$

The preceding relation is usually written as follows:

$$\mathbf{F}(s) = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) \tag{2.18}$$

We consider the following Single-Input Single-Output (SISO) transfer function which is composed of the sum between a strictly proper rational function and a constant value d:

$$\frac{Y(s)}{U(s)} = F(s) = \frac{N(s)}{D(s)} + d \tag{2.19}$$

Where $N(s)$ and $D(s)$ are polynomials in s such that the degree of $N(s)$ is strictly lower than the degree of $D(s)$:

$$\begin{cases} D(s) = a_0 + a_1 s + \cdots + a_{n-1}s^{n-1} + 1 \times s^n \\ N(s) = n_0 + n_1 s + \cdots + n_{n-1}s^{n-1} \end{cases} \tag{2.20}$$

It is worth noticing that polynomial $D(s)$ is assumed to be a monic polynomial without loss of generally. This means that the leading coefficient (that is the coefficient of s^n) of $D(s)$ is 1. Indeed $D(s)$ is identified to $\det(s\mathbb{I} - \mathbf{A})$.

When identifying (2.19) with (2.17) we get:

$$d = \mathbf{D} = \lim_{s\to\infty} \mathbf{F}(s) \tag{2.21}$$

Thus all we need now is to find a triplet $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ such that:

$$\frac{N(s)}{D(s)} = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} \tag{2.22}$$

2.3.1 Controllable canonical form

Let $N(s) = N_1(s)N_2(s)$. Then one solution of the realization problem is the following quadruplet:

$$N(s) = N_1(s)N_2(s) \Rightarrow \begin{cases} \mathbf{A}_c = \begin{bmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ & & \ddots & \ddots & 0 \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \\ \mathbf{B}_c = \mathbf{T}^{-1} \begin{bmatrix} \vdots \\ \vdots \\ \frac{N_2''(0)}{2!} \\ \frac{N_2'(0)}{1!} \\ N_2(0) \end{bmatrix} \\ \mathbf{C}_c = \begin{bmatrix} N_1(0) & \frac{N_1'(0)}{1!} & \frac{N_1''(0)}{2!} & \cdots & \cdots \end{bmatrix} \\ \mathbf{D} = d \end{cases} \tag{2.23}$$

where $\mathbf{T}$ is the following Toeplitz matrix:

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ a_{n-1} & 1 & \ddots & \ddots & 0 \\ a_{n-2} & a_{n-1} & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ a_1 & a_2 & \cdots & a_{n-1} & 1 \end{bmatrix} \tag{2.24}$$

In the case where we choose $N_1(s) := N(s)$ and $N_2(s) := 1$ we get:

$$\begin{cases} N_1(s) := N(s) \\ N_2(s) := 1 \end{cases} \Rightarrow \begin{cases} \mathbf{A}_c = \begin{bmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ & & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \\ \mathbf{B}_c = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \\ \mathbf{C}_c = \begin{bmatrix} n_0 & n_1 & \cdots & n_{n-2} & n_{n-1} \end{bmatrix} \\ \mathbf{D} = d \end{cases} \tag{2.25}$$

The quadruplet $(\mathbf{A}_c, \mathbf{B}_c, \mathbf{C}_c, d)$ is called the controllable canonical form of the SISO transfer function $F(s)$.

Alternatively the following realization is also called the controllable canonical form of the SISO transfer function $F(s)$. Compared with (2.25) value 1 appears in the lower diagonal of the state matrix which is obtained by choosing a similarity transformation with value 1 on the antidiagonal (or

counter diagonal):

$$\begin{cases} \mathbf{A}_{ca} = \begin{bmatrix} 0 & 0 & 0 & & -a_0 \\ 1 & 0 & 0 & \ddots & -a_1 \\ 0 & \ddots & \ddots & \ddots & -a_2 \\ & \ddots & & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{bmatrix} \\ \mathbf{B}_{ca} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \\ \mathbf{C}_{ca} = \begin{bmatrix} n_{n-1} & n_{n-2} & \cdots & n_1 & n_0 \end{bmatrix} \\ \mathbf{D} = d \end{cases} \tag{2.26}$$

To get the realization (2.25) we start by expressing the output $Y(s)$ of SISO system (2.19) as follows:

$$Y(s) = N(s)\frac{U(s)}{D(s)} + dU(s) \tag{2.27}$$

Now let's focus on the following intermediate variable $Z(s)$ which is defined as follows:

$$Z(s) = \frac{U(s)}{D(s)} = \frac{U(s)}{a_0 + a_1 s + a_2 s^2 + \cdots + a_{n-1}s^{n-1} + s^n} \tag{2.28}$$

That is:

$$a_0 Z(s) + a_1 s Z(s) + a_2 s^2 Z(s) + \cdots + a_{n-1}s^{n-1}Z(s) + s^n Z(s) = U(s) \tag{2.29}$$

Then we define the components of the state vector $\underline{x}(t)$ as follows:

$$\begin{cases} x_1(t) := z(t) \\ x_2(t) := \dot{x}_1(t) = \dot{z}(t) \\ x_3(t) := \dot{x}_2(t) = \ddot{z}(t) \\ \vdots \\ x_n(t) := \dot{x}_{n-1}(t) = z^{(n-1)}(t) \end{cases} \tag{2.30}$$

Coming back in the time domain Equation (2.29) is rewritten as follows:

$$\begin{aligned} & a_0 x_1(t) + a_1 x_2(t) + a_2 x_3(t) + \cdots + a_{n-1}x_n(t) + \dot{x}_n(t) = u(t) \\ & \Leftrightarrow \dot{x}_n(t) = -a_0 x_1(t) - a_1 x_2(t) - a_2 x_3(t) - \cdots - a_{n-1}x_n(t) + u(t) \end{aligned} \tag{2.31}$$

The intermediate variable $Z(s)$ allows us to express the output $Y(s)$ as follows:

$$Y(s) = N(s)Z(s) + dU(s) = \left(n_0 + \cdots + n_{n-1}s^{n-1}\right) Z(s) + dU(s) \tag{2.32}$$

That is, coming back if the time domain:

$$y(t) = n_0 z(t) + \cdots + n_{n-1} z^{(n-1)}(t) + du(t) \tag{2.33}$$

The use of the components of the state vector which have been previously defined leads to the following expression of the output $y(t)$:

$$y(t) = n_0 x_1(t) + \cdots + n_{n-1} x_n(t) + du(t) \tag{2.34}$$

By combining in vector form Equations (2.30), (2.31) and (2.34) we retrieve the state-space representation (2.25).

Thus by ordering the numerator and the denominator of the transfer function $F(s)$ according to the increasing power of s and taking care that the leading coefficient of the polynomial in the denominator is 1, the controllable canonical form (2.25) of a SISO transfer function $F(s)$ is immediate.

Example 2.1. *Let's consider the following transfer function:*

$$F(s) = \frac{(s+1)(s+2)}{2(s+3)(s+4)} = \frac{s^2+3s+2}{2s^2+14s+24} \tag{2.35}$$

We are looking for the controllable canonical form of this transfer function.

First we have to set to 1 the leading coefficient of the polynomial which appears in the denominator of the transfer function $F(s)$. We get:

$$F(s) = \frac{0.5s^2+1.5s+1}{1 \times s^2+7s+12} \tag{2.36}$$

Then we decompose $F(s)$ as a sum between a strictly proper rational function and a constant coefficient d. Constant coefficient d is obtained thanks to the following relation:

$$d = \lim_{s\to\infty} F(s) = \lim_{s\to\infty} \frac{0.5s^2+1.5s+1}{1 \times s^2+7s+12} = 0.5 \tag{2.37}$$

Thus the strictly proper transfer function $N(s)/D(s)$ is obtained by subtracting d to $F(s)$:

$$\frac{N(s)}{D(s)} = F(s) - d = \frac{0.5s^2+1.5s+1}{1 \times s^2+7s+12} - 0.5 = \frac{-2s-5}{s^2+7s+12} \tag{2.38}$$

We finally get:

$$F(s) = \frac{N(s)}{D(s)} + d = \frac{-2s-5}{s^2+7s+12} + 0.5 \tag{2.39}$$

Then we apply Equation (2.25) *to get the controllable canonical form of* $F(s)$*:*

$$\begin{cases} \mathbf{A}_c = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -12 & -7 \end{bmatrix} \\ \mathbf{B}_c = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \mathbf{C}_c = \begin{bmatrix} n_0 & n_1 \end{bmatrix} = \begin{bmatrix} -5 & -2 \end{bmatrix} \\ \mathbf{D} = 0.5 \end{cases} \tag{2.40}$$

■

2.3.2 Poles and zeros of the transfer function

It is worth noticing that the numerator of the transfer function only depends on matrices $\mathbf{B}$ and $\mathbf{C}$ whereas the denominator of the transfer function is built from the characteristic polynomial coming from the eigenvalues of the state matrix $\mathbf{A}$.

As far as the transfer function does not depend on the state space realization which is used, we can get this result by using the controllable canonical form. Indeed we can check that transfer function $\mathbf{C}_c\left(s\mathbb{I} - \mathbf{A}_c\right)^{-1}\mathbf{B}_c$ has a denominator which only depends on the state matrix $\mathbf{A}_c$ whereas its numerator only depends on $\mathbf{C}_c$, which provides the coefficients of the numerator:

$$\begin{aligned} \left(s\mathbb{I} - \mathbf{A}_c\right)^{-1}\mathbf{B}_c &= \left(\frac{1}{\det\left(s\mathbb{I} - \mathbf{A}_c\right)}\begin{bmatrix} * & * & 1 \\ * & * & s \\ \vdots & \vdots & \vdots \\ * & * & s^{n-1} \end{bmatrix}\right)\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \\ \Rightarrow \mathbf{C}_c\left(s\mathbb{I} - \mathbf{A}_c\right)^{-1}\mathbf{B}_c &= \frac{\mathbf{C}_c}{\det\left(s\mathbb{I} - \mathbf{A}_c\right)}\begin{bmatrix} 1 \\ s \\ \vdots \\ s^{n-1} \end{bmatrix} \end{aligned} \tag{2.41}$$

More generally, the characteristic polynomial of the state matrix $\mathbf{A}$ sets the denominator of the transfer function whereas the product $\mathbf{B}\,\mathbf{C}$ sets the coefficients of the numerator of a strictly proper transfer function (that is a transfer function where $\mathbf{D} = \mathbf{0}$). Consequently state matrix $\mathbf{A}$ sets the poles of a transfer function whereas product $\mathbf{B}\,\mathbf{C}$ sets its zeros.

2.3.3 Similarity transformation to controllable canonical form

We consider the following general state-space representation:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}u(t) \\ y(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}u(t) \end{cases} \tag{2.42}$$

where the size of the state vector $\underline{x}(t)$ is n.

Use of the controllability matrix

The controllable canonical form (2.25) exists if and only if the following matrix $\mathbf{Q}_c$, which is called the controllability matrix, has full rank:

$$\mathbf{Q}_c = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \tag{2.43}$$

As soon as the characteristic polynomial of matrix $\mathbf{A}$ is computed the state matrix $\mathbf{A}_c$ as well as the control matrix $\mathbf{B}_c$ corresponding to the controllable canonical form are known. Thus the controllability matrix in the controllable canonical basis, which will be denoted $\mathbf{Q}_{cc}$, can be computed as follows:

$$\mathbf{Q}_{cc} = \begin{bmatrix} \mathbf{B}_c & \mathbf{A}_c\mathbf{B}_c & \cdots & \mathbf{A}_c^{n-1}\mathbf{B}_c \end{bmatrix} \tag{2.44}$$

At that point matrices $\mathbf{A}_c$ and $\mathbf{B}_c$ are known. The only matrix which need to be computed is the output matrix $\mathbf{C}_c$. Let $\mathbf{P}_c$ be the change of basis matrix which defines the new state vector in the controllable canonical basis. From (2.8) we get:

$$\mathbf{C}_c = \mathbf{C}\mathbf{P}_c \tag{2.45}$$

And:

$$\begin{cases} \mathbf{A}_c = \mathbf{P}_c^{-1}\mathbf{A}\mathbf{P}_c \\ \mathbf{B}_c = \mathbf{P}_c^{-1}\mathbf{B} \end{cases} \tag{2.46}$$

Using these two last equations within (2.44) and the fact that $\left(\mathbf{P}_c^{-1}\mathbf{A}\mathbf{P}_c\right)^k = \underbrace{\mathbf{P}_c^{-1}\mathbf{A}\mathbf{P}_c \cdots \mathbf{P}_c^{-1}\mathbf{A}\mathbf{P}_c}_{\text{k-times}} = \mathbf{P}_c^{-1}\mathbf{A}^k\mathbf{P}_c$, we get the following expression of matrix $\mathbf{Q}_{cc}$:

$$\begin{aligned} \mathbf{Q}_{cc} &= \begin{bmatrix} \mathbf{B}_c & \mathbf{A}_c\mathbf{B}_c & \cdots & \mathbf{A}_c^{n-1}\mathbf{B}_c \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{P}_c^{-1}\mathbf{B} & \mathbf{P}_c^{-1}\mathbf{A}\mathbf{P}_c\mathbf{P}_c^{-1}\mathbf{B} & \cdots & \left(\mathbf{P}_c^{-1}\mathbf{A}\mathbf{P}_c\right)^{n-1}\mathbf{P}_c^{-1}\mathbf{B} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{P}_c^{-1}\mathbf{B} & \mathbf{P}_c^{-1}\mathbf{A}\mathbf{B} & \cdots & \mathbf{P}_c^{-1}\mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \\ &= \mathbf{P}_c^{-1}\begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \\ &= \mathbf{P}_c^{-1}\mathbf{Q}_c \end{aligned} \tag{2.47}$$

We finally get:

$$\mathbf{P}_c^{-1} = \mathbf{Q}_{cc}\mathbf{Q}_c^{-1} \Leftrightarrow \mathbf{P}_c = \mathbf{Q}_c\mathbf{Q}_{cc}^{-1} \tag{2.48}$$

Furthermore the controllable canonical form (2.25) is obtained by the following similarity transformation:

$$\underline{x}(t) = \mathbf{P}_c\underline{x}_c(t) \Leftrightarrow \underline{x}_c(t) = \mathbf{P}_c^{-1}\underline{x}(t) \tag{2.49}$$

Alternatively the constant nonsingular matrix $\mathbf{P}_c^{-1}$ can be obtained through the state matrix $\mathbf{A}$ and the last row $\underline{q}_c^T$ of the inverse of the controllability matrix $\mathbf{Q}_c$ as follows:

$$\mathbf{Q}_c^{-1} = \begin{bmatrix} * \\ \vdots \\ * \\ \underline{q}_c^T \end{bmatrix} \Rightarrow \mathbf{P}_c^{-1} = \begin{bmatrix} \underline{q}_c^T \\ \underline{q}_c^T\mathbf{A} \\ \vdots \\ \underline{q}_c^T\mathbf{A}^{n-1} \end{bmatrix} \tag{2.50}$$

To get this result we write from (2.8) the following similarity transformation:

$$\mathbf{A}_c = \mathbf{P}_c^{-1}\mathbf{A}\mathbf{P}_c \Leftrightarrow \mathbf{A}_c\mathbf{P}_c^{-1} = \mathbf{P}_c^{-1}\mathbf{A} \tag{2.51}$$

Let's denote $\det(s\mathbb{I} - \mathbf{A})$ as follows:

$$\det(s\mathbb{I} - \mathbf{A}) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \tag{2.52}$$

Thus the coefficients a_i of the state matrix $\mathbf{A}_c$ corresponding to the controllable canonical form are known and matrix $\mathbf{A}_c$ is written as follows:

$$\mathbf{A}_c = \begin{bmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ & & \ddots & \ddots & 0 \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \tag{2.53}$$

Furthermore let's write the unknown matrix $\mathbf{P}_c^{-1}$ as follows:

$$\mathbf{P}_c^{-1} = \begin{bmatrix} \underline{r}_1^T \\ \vdots \\ \underline{r}_n^T \end{bmatrix} \tag{2.54}$$

Thus the rows of the unknown matrix $\mathbf{P}_c^{-1}$ can be obtained thanks to the following similarity transformation:

$$\begin{aligned} &\mathbf{A}_c\mathbf{P}_c^{-1} = \mathbf{P}_c^{-1}\mathbf{A} \\ \Leftrightarrow &\begin{bmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ & & \ddots & \ddots & 0 \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \begin{bmatrix} \underline{r}_1^T \\ \vdots \\ \underline{r}_n^T \end{bmatrix} = \begin{bmatrix} \underline{r}_1^T \\ \vdots \\ \underline{r}_n^T \end{bmatrix} \mathbf{A} \end{aligned} \tag{2.55}$$

Working out with the first $n-1^{th}$ rows gives the following equations:

$$\begin{cases} \underline{r}_2^T = \underline{r}_1^T\mathbf{A} \\ \underline{r}_3^T = \underline{r}_2^T\mathbf{A} = \underline{r}_1^T\mathbf{A}^2 \\ \vdots \\ \underline{r}_n^T = \underline{r}_{n-1}^T\mathbf{A} = \underline{r}_1^T\mathbf{A}^{n-1} \end{cases} \tag{2.56}$$

Furthermore from (2.8) we get the relation $\mathbf{B}_c = \mathbf{P}_c^{-1}\mathbf{B}$ which is rewritten as follows:

$$\mathbf{P}_c^{-1}\mathbf{B} = \mathbf{B}_c \Leftrightarrow \begin{bmatrix} \underline{r}_1^T \\ \vdots \\ \underline{r}_n^T \end{bmatrix} \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \Leftrightarrow \begin{cases} \underline{r}_1^T\mathbf{B} = 0 \\ \vdots \\ \underline{r}_{n-1}^T\mathbf{B} = 0 \\ \underline{r}_n^T\mathbf{B} = 1 \end{cases} \tag{2.57}$$

Combining (2.56) and (2.57) we get:

$$\begin{cases} \underline{r}_1^T\mathbf{B} = 0 \\ \underline{r}_2^T\mathbf{B} = \underline{r}_1^T\mathbf{AB} = 0 \\ \vdots \\ \underline{r}_{n-1}^T\mathbf{B} = \underline{r}_1^T\mathbf{A}^{n-2}\mathbf{B} = 0 \\ \underline{r}_n^T\mathbf{B} = \underline{r}_1^T\mathbf{A}^{n-1}\mathbf{B} = 1 \end{cases} \tag{2.58}$$

These equations can in turn be written in matrix form as:

$$\underline{r}_1^T \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-2}\mathbf{B} & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{2.59}$$

Let's introduce the controllability matrix $\mathbf{Q}_c$:

$$\mathbf{Q}_c = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \tag{2.60}$$

Assuming that matrix $\mathbf{Q}_c$ has full rank we get:

$$\underline{r}_1^T\mathbf{Q}_c = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \Leftrightarrow \underline{r}_1^T = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix}\mathbf{Q}_c^{-1} \tag{2.61}$$

From the preceding equation it is clear that $\underline{r}_1^T$ is the last row of the inverse of the controllability matrix $\mathbf{Q}_c$. We will denote it $\underline{q}_c^T$:

$$\underline{r}_1^T := \underline{q}_c^T \tag{2.62}$$

Having the expression of $\underline{r}_1^T$ we can then go back to (2.56) and construct all the rows of $\mathbf{P}_c^{-1}$.

Example 2.2. *We consider the following general state-space representation:*

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}u(t) \\ y(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}u(t) \end{cases} \tag{2.63}$$

where:

$$\begin{cases} \mathbf{A} = \begin{bmatrix} 28.5 & -17.5 \\ 58.5 & -35.5 \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \\ \mathbf{C} = \begin{bmatrix} 7 & -4 \end{bmatrix} \\ \mathbf{D} = 0.5 \end{cases} \tag{2.64}$$

We are looking for the controllable canonical form of this state-space representation.

First we build the controllability matrix $\mathbf{Q}_c$ *from* (2.43)*:*

$$\mathbf{Q}_c = \begin{bmatrix} \mathbf{B} & \mathbf{AB} \end{bmatrix} = \begin{bmatrix} 2 & -13 \\ 4 & -25 \end{bmatrix} \tag{2.65}$$

To build matrix $\mathbf{Q}_{cc}$ *let's compute* $\det(s\mathbb{I}-\mathbf{A})$*:*

$$\det(s\mathbb{I}-\mathbf{A}) = s^2 + a_1 s + a_0 = s^2 + 7s + 12 \tag{2.66}$$

As soon as matrix $\mathbf{A}_c$ *is built from the denominator of the transfer function, that is from* $\det(s\mathbb{I}-\mathbf{A})$*, we get:*

$$\mathbf{A}_c = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -12 & -7 \end{bmatrix} \tag{2.67}$$

Furthermore matrix $\mathbf{B}_c$ *is straightforward for the controllable canonical form:*

$$\mathbf{B}_c = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{2.68}$$

Thus we are in position to compute matrix $\mathbf{Q}_{cc}$ *:*

$$\mathbf{Q}_{cc} = \begin{bmatrix} \mathbf{B}_c & \mathbf{A}_c\mathbf{B}_c \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -7 \end{bmatrix} \tag{2.69}$$

Then we use (2.48) *to build the similarity transformation:*

$$\begin{aligned} \mathbf{P}_c &= \mathbf{Q}_c\mathbf{Q}_{cc}^{-1} = \begin{bmatrix} 2 & -13 \\ 4 & -25 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -7 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 2 & -13 \\ 4 & -25 \end{bmatrix} \begin{bmatrix} 7 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \end{aligned} \tag{2.70}$$

Alternatively we can use (2.50) *to build the similarity transformation:*

$$\begin{aligned} &\mathbf{Q}_c^{-1} = \tfrac{1}{2}\begin{bmatrix} -25 & 13 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} * \\ \underline{q}_c^T \end{bmatrix} \\ &\Rightarrow \underline{q}_c^T = \tfrac{1}{2}\begin{bmatrix} -4 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \end{bmatrix} \end{aligned} \tag{2.71}$$

And:

$$\mathbf{P}_c^{-1} = \begin{bmatrix} \underline{q}_c^T \\ \underline{q}_c^T\mathbf{A} \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix} \tag{2.72}$$

Using the similarity relations (2.8) *we finally get the following controllable canonical form of the state-space representation:*

$$\begin{cases} \mathbf{A}_c = \mathbf{P}_c^{-1}\mathbf{A}\mathbf{P}_c = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix} \begin{bmatrix} 28.5 & -17.5 \\ 58.5 & -35.5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -12 & -7 \end{bmatrix} \\ \mathbf{B}_c = \mathbf{P}_c^{-1}\mathbf{B} = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \mathbf{C}_c = \mathbf{C}\mathbf{P}_c = \begin{bmatrix} 7 & -4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} -5 & -2 \end{bmatrix} \end{cases} \tag{2.73}$$

■

Iterative method

Equivalently change of basis matrix $\mathbf{P}_c$ of the similarity transformation can be obtained as follows:

$$\mathbf{P}_c = \begin{bmatrix} \underline{c}_1 & \underline{c}_2 & \cdots & \underline{c}_n \end{bmatrix} \tag{2.74}$$

where:

$$\begin{cases} \det(s\mathbb{I} - \mathbf{A}) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \\ \underline{c}_n = \mathbf{B} \\ \underline{c}_k = \mathbf{A}\underline{c}_{k+1} + a_k\mathbf{B} \quad \forall\, n-1 \geq k \geq 1 \end{cases} \tag{2.75}$$

To get this result we write from (2.8) the following similarity transformation:

$$\mathbf{A}_c = \mathbf{P}_c^{-1}\mathbf{A}\mathbf{P}_c \Leftrightarrow \mathbf{P}_c\mathbf{A}_c = \mathbf{A}\mathbf{P}_c \tag{2.76}$$

Let's denote $\det(s\mathbb{I} - \mathbf{A})$ as follows:

$$\det(s\mathbb{I} - \mathbf{A}) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \tag{2.77}$$

Thus the coefficients a_i of the state matrix $\mathbf{A}_c$ corresponding to the controllable canonical form are known and matrix $\mathbf{A}_c$ written as follows:

$$\mathbf{A}_c = \begin{bmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ & & \ddots & \ddots & 0 \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \tag{2.78}$$

Furthermore let's write the unknown change of basis matrix $\mathbf{P}_c$ as follows:

$$\mathbf{P}_c = \begin{bmatrix} \underline{c}_1 & \underline{c}_2 & \cdots & \underline{c}_n \end{bmatrix} \tag{2.79}$$

Thus the columns of the unknown matrix $\mathbf{P}_c$ can be obtained thanks to the similarity transformation:

$$\begin{aligned} &\mathbf{P}_c\mathbf{A}_c = \mathbf{A}\mathbf{P}_c \\ \Leftrightarrow &\begin{bmatrix} \underline{c}_1 & \underline{c}_2 & \cdots & \underline{c}_n \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ & & \ddots & \ddots & 0 \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} = \mathbf{A} \begin{bmatrix} \underline{c}_1 & \underline{c}_2 & \cdots & \underline{c}_n \end{bmatrix} \end{aligned} \tag{2.80}$$

That is:

$$\begin{cases} \underline{0} = a_0\underline{c}_n + \mathbf{A}\underline{c}_1 \\ \underline{c}_1 = a_1\underline{c}_n + \mathbf{A}\underline{c}_2 \\ \vdots \\ \underline{c}_{n-1} = a_{n-1}\underline{c}_n + \mathbf{A}\underline{c}_n \end{cases} \Leftrightarrow \begin{cases} \underline{0} = a_0\underline{c}_n + \mathbf{A}\underline{c}_1 \\ \underline{c}_k = \mathbf{A}\underline{c}_{k+1} + a_k\underline{c}_n \quad \forall\, n-1 \geq k \geq 1 \end{cases} \tag{2.81}$$

Furthermore from (2.8) we get the relation $\mathbf{B}_c = \mathbf{P}_c^{-1}\mathbf{B}$ which is rewritten as follows:

$$\mathbf{P}_c\mathbf{B}_c = \mathbf{B} \Leftrightarrow \begin{bmatrix} \underline{c}_1 & \underline{c}_2 & \cdots & \underline{c}_n \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \mathbf{B} \Rightarrow \underline{c}_n = \mathbf{B} \tag{2.82}$$

Combining the last equation of (2.81) with (2.82) gives the proposed result:

$$\begin{cases} \underline{c}_n = \mathbf{B} \\ \underline{c}_k = \mathbf{A}\underline{c}_{k+1} + a_k\mathbf{B} \quad \forall\, n-1 \geq k \geq 1 \end{cases} \tag{2.83}$$

Example 2.3. *We consider the following general state-space representation:*

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}u(t) \\ y(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}u(t) \end{cases} \tag{2.84}$$

where:

$$\begin{cases} \mathbf{A} = \begin{bmatrix} 28.5 & -17.5 \\ 58.5 & -35.5 \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \\ \mathbf{C} = \begin{bmatrix} 7 & -4 \end{bmatrix} \\ \mathbf{D} = 0.5 \end{cases} \tag{2.85}$$

This is the same state-space representation than the one which has been used in the previous example. We have seen that the similarity transformation which leads to the controllable canonical form is the following:

$$\mathbf{P}_c^{-1} = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix} \tag{2.86}$$

It is easy to compute matrix $\mathbf{P}_c$, that is the inverse of $\mathbf{P}_c^{-1}$. We get the following expression:

$$\mathbf{P}_c = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \tag{2.87}$$

We will check the expression of matrix $\mathbf{P}_c$ thanks to the iterative method proposed in (2.75)*. We get:*

$$\begin{cases} \det(s\mathbb{I} - \mathbf{A}) = s^2 + 7s + 12 \\ \underline{c}_2 = \mathbf{B} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \\ \underline{c}_1 = \mathbf{A}\underline{c}_2 + a_1\mathbf{B} = \begin{bmatrix} 28.5 & -17.5 \\ 58.5 & -35.5 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} + 7 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \end{cases} \tag{2.88}$$

Thus we fortunately retrieve the expression of matrix $\mathbf{P}_c$*:*

$$\mathbf{P}_c = \begin{bmatrix} \underline{c}_1 & \underline{c}_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \tag{2.89}$$

■

2.3.4 Observable canonical form

Another solution of the realization problem is the following quadruplet:

$$\begin{cases} \mathbf{A}_o = \begin{bmatrix} 0 & 0 & 0 & & -a_0 \\ 1 & 0 & 0 & \ddots & -a_1 \\ 0 & 1 & 0 & \ddots & -a_2 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & & 0 & 1 & -a_{n-1} \end{bmatrix} \\ \mathbf{B}_o = \begin{bmatrix} n_0 \\ n_1 \\ \vdots \\ n_{n-2} \\ n_{n-1} \end{bmatrix} \\ \mathbf{C}_o = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \\ \mathbf{D} = d \end{cases} \tag{2.90}$$

The quadruplet $(\mathbf{A}_o, \mathbf{B}_o, \mathbf{C}_o, d)$ is called the observable canonical form of the SISO transfer function $F(s)$.

Alternatively the following realization is also called the observable canonical form of the SISO transfer function $F(s)$. Compared with (2.90) value 1 appears in the upper diagonal of the state matrix which is obtained by choosing a similarity transformation with value 1 on the antidiagonal (or

counter diagonal):

$$\begin{cases} \mathbf{A}_{oa} = \begin{bmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ & & \ddots & \ddots & 0 \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \\ \mathbf{B}_{oa} = \begin{bmatrix} n_{n-1} \\ n_{n-2} \\ \cdots \\ n_1 \\ n_0 \end{bmatrix} \\ \mathbf{C}_{oa} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \\ \mathbf{D} = d \end{cases} \tag{2.91}$$

To get the realization (2.90) we start by expressing the output $Y(s)$ of SISO system (2.19) as follows:

$$\frac{Y(s)}{U(s)} = \frac{N(s)}{D(s)} + d \Leftrightarrow (Y(s) - dU(s))\, D(s) = N(s)U(s) \tag{2.92}$$

That is:

$$\begin{aligned} &\left(a_0 + a_1 s + a_2 s^2 + \cdots + a_{n-1}s^{n-1} + s^n\right)(Y(s) - d\,U(s)) \\ &\qquad = \left(n_0 + n_1 s + \cdots + n_{n-1}s^{n-1}\right) U(s) \end{aligned} \tag{2.93}$$

Dividing by s^n we get:

$$\begin{aligned} &\left(\frac{a_0}{s^n} + \frac{a_1}{s^{n-1}} + \frac{a_2}{s^{n-2}} + \cdots + \frac{a_{n-1}}{s} + 1\right)(Y(s) - d\,U(s)) \\ &\qquad = \left(\frac{n_0}{s^n} + \frac{n_1}{s^{n-1}} + \frac{n_2}{s^{n-2}} + \cdots + \frac{n_{n-1}}{s}\right) U(s) \end{aligned} \tag{2.94}$$

When regrouping the terms according the increasing power of $\frac{1}{s}$ we obtain:

$$\begin{aligned} Y(s) = d\,U(s) + \frac{1}{s}\left(\alpha_{n-1}U(s) - a_{n-1}Y(s)\right) + \frac{1}{s^2}\left(\alpha_{n-2}U(s) - a_{n-2}Y(s)\right) + \\ \cdots + \frac{1}{s^n}\left(\alpha_0 U(s) - a_0 Y(s)\right) \end{aligned} \tag{2.95}$$

Where:

$$\alpha_i = n_i + d\,a_i \tag{2.96}$$

That is:

$$\begin{aligned} Y(s) = d\,U(s) + \frac{1}{s}\Bigg(\alpha_{n-1}U(s) - a_{n-1}Y(s) + \frac{1}{s}\Big(\alpha_{n-2}U(s) - a_{n-2}Y(s)\Big) + \\ \frac{1}{s}\Bigg(\cdots + \frac{1}{s}\Big(\alpha_0 U(s) - a_0 Y(s)\Big)\Bigg)\Bigg) \end{aligned} \tag{2.97}$$

Then we define the Laplace transform of the components of the state vector $\underline{x}(t)$ as follows:

$$\begin{cases} sX_1(s) = \alpha_0 U(s) - a_0 Y(s) \\ sX_2(s) = \alpha_1 U(s) - a_1 Y(s) + X_1(s) \\ sX_3(s) = \alpha_2 U(s) - a_2 Y(s) + X_2(s) \\ \vdots \\ sX_n(s) = \alpha_{n-1} U(s) - a_{n-1} Y(s) + X_{n-1}(s) \end{cases} \tag{2.98}$$

So we get:

$$Y(s) = d\,U(s) + \frac{1}{s}\Big(sX_n(s)\Big) = d\,U(s) + X_n(s) \tag{2.99}$$

Replacing $Y(s)$ by $X_n(s)$ and using the fact that $\alpha_i = n_i + d\,a_i$ Equation (2.98) is rewritten as follows:

$$\begin{cases} sX_1(s) & = \alpha_0 U(s) - a_0\,(d\,U(s) + X_n(s)) \\ & = -a_0 X_n(s) + n_0 U(s) \\ sX_2(s) & = \alpha_1 U(s) - a_1\,(d\,U(s) + X_n(s)) + X_1(s) \\ & = X_1(s) - a_1 X_n(s) + n_1 U(s) \\ sX_3(s) & = \alpha_2 U(s) - a_2\,(d\,U(s) + X_n(s)) + X_2(s) \\ & = X_2(s) - a_2 X_n(s) + n_2 U(s) \\ \vdots & \\ sX_n(s) & = \alpha_{n-1} U(s) - a_{n-1}\,(d\,U(s) + X_n(s)) + X_{n-1}(s) \\ & = X_{n-1}(s) - a_{n-1} X_n(s) + n_{n-1} U(s) \end{cases} \tag{2.100}$$

Coming back in the time domain we finally get:

$$\begin{cases} \dot{x}_1(t) = -a_0 x_n(t) + n_0 u(t) \\ \dot{x}_2(t) = x_1(t) - a_1 x_n(t) + n_1 u(t) \\ \dot{x}_3(t) = x_2(t) - a_2 x_n(t) + n_2 u(t) \\ \vdots \\ \dot{x}_n(t) = x_{n-1}(t) - a_{n-1} x_n(t) + n_{n-1} u(t) \end{cases} \tag{2.101}$$

And:

$$y(t) = x_n(t) + d\,u(t) \tag{2.102}$$

The preceding equations written in vector form leads to the observable canonical form of Equation (2.90).

Thus by ordering the numerator and the denominator of the transfer function $F(s)$ according to the increasing power of s and taking care that the leading coefficient of the polynomial in the denominator is 1, the observable canonical form (2.90) of a SISO transfer function $F(s)$ is immediate.

Example 2.4. *Let's consider the following transfer function:*

$$F(s) = \frac{(s+1)(s+2)}{2(s+3)(s+4)} = \frac{s^2 + 3s + 2}{2s^2 + 14s + 24} \tag{2.103}$$

We are looking for the observable canonical form of this transfer function.

As in the preceding example we first set to 1 the leading coefficient of the polynomial which appears in the denominator of the transfer function $F(s)$. We get:

$$F(s) = \frac{0.5s^2 + 1.5s + 1}{1 \times s^2 + 7s + 12} \tag{2.104}$$

Then we decompose $F(s)$ as a sum between a strictly proper rational function and a constant coefficient d. Constant coefficient d is obtained thanks to the following relation:

$$d = \lim_{s\to\infty} \mathbf{F}(s) = \lim_{s\to\infty} \frac{0.5s^2 + 1.5s + 1}{1 \times s^2 + 7s + 12} = 0.5 \tag{2.105}$$

Thus the strictly proper transfer function $N(s)/D(s)$ is obtained by subtracting d to $F(s)$:

$$\frac{N(s)}{D(s)} = F(s) - d = \frac{0.5s^2 + 1.5s + 1}{1 \times s^2 + 7s + 12} - 0.5 = \frac{-2s - 5}{s^2 + 7s + 12} \tag{2.106}$$

We finally get:

$$F(s) = \frac{N(s)}{D(s)} + d = \frac{-2s - 5}{s^2 + 7s + 12} + 0.5 \tag{2.107}$$

Then we apply Equation (2.90) to get the observable canonical form of $F(s)$:

$$\begin{cases} \mathbf{A}_o = \begin{bmatrix} 0 & -a_0 \\ 1 & -a_1 \end{bmatrix} = \begin{bmatrix} 0 & -12 \\ 1 & -7 \end{bmatrix} \\ \mathbf{B}_o = \begin{bmatrix} n_0 \\ n_1 \end{bmatrix} = \begin{bmatrix} -5 \\ -2 \end{bmatrix} \\ \mathbf{C}_o = \begin{bmatrix} 0 & 1 \end{bmatrix} \\ \mathbf{D} = 0.5 \end{cases} \tag{2.108}$$

■

2.3.5 Similarity transformation to observable canonical form

We consider the following general state-space representation:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}u(t) \\ y(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}u(t) \end{cases} \tag{2.109}$$

where the size of the state vector $\underline{x}(t)$ is n.

Use of the observability matrix

The observable canonical form (2.90) exists if and only if the following matrix $\mathbf{Q}_o$, which is called the observability matrix, has full rank:

$$\mathbf{Q}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \tag{2.110}$$

As soon as the characteristic polynomial of matrix $\mathbf{A}$ is computed the state matrix $\mathbf{A}_o$ as well as the output matrix $\mathbf{C}_o$ corresponding to the observable canonical form are known. Thus the observability matrix in the observable canonical basis, which will be denoted $\mathbf{Q}_{oo}$, can be computed as follows:

$$\mathbf{Q}_{oo} = \begin{bmatrix} \mathbf{C}_o \\ \mathbf{C}_o\mathbf{A}_o \\ \vdots \\ \mathbf{C}_o\mathbf{A}_o^{n-1} \end{bmatrix} \tag{2.111}$$

At that point matrices $\mathbf{A}_o$ and $\mathbf{C}_o$ are known. The only matrix which need to be computed is the control matrix $\mathbf{B}_o$. Let $\mathbf{P}_o$ be the change of basis matrix which defines the new state vector in the observable canonical basis. From (2.8) we get:

$$\mathbf{B}_o = \mathbf{P}_o^{-1}\mathbf{B} \tag{2.112}$$

And:

$$\begin{cases} \mathbf{A}_o = \mathbf{P}_o^{-1}\mathbf{A}\mathbf{P}_o \\ \mathbf{C}_o = \mathbf{C}\mathbf{P}_o \end{cases} \tag{2.113}$$

Using these last two equations within (2.111) leads to the following expression of matrix $\mathbf{Q}_{oo}$:

$$\begin{aligned} \mathbf{Q}_{oo} &= \begin{bmatrix} \mathbf{C}_o \\ \mathbf{C}_o\mathbf{A}_o \\ \vdots \\ \mathbf{C}_o\mathbf{A}_o^{n-1} \end{bmatrix} = \begin{bmatrix} \mathbf{C}\mathbf{P}_o \\ \mathbf{C}\mathbf{P}_o\mathbf{P}_o^{-1}\mathbf{A}\mathbf{P}_o \\ \vdots \\ \mathbf{C}\mathbf{P}_o\left(\mathbf{P}_o^{-1}\mathbf{A}\mathbf{P}_o\right)^{n-1} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{C}\mathbf{P}_o \\ \mathbf{C}\mathbf{A}\mathbf{P}_o \\ \vdots \\ \mathbf{C}\mathbf{A}^{n-1}\mathbf{P}_o \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \vdots \\ \mathbf{C}\mathbf{A}^{n-1} \end{bmatrix}\mathbf{P}_o \\ &= \mathbf{Q}_o\mathbf{P}_o \end{aligned} \tag{2.114}$$

We finally get:

$$\mathbf{P}_o = \mathbf{Q}_o^{-1}\mathbf{Q}_{oo} \Leftrightarrow \mathbf{P}_o^{-1} = \mathbf{Q}_{oo}^{-1}\mathbf{Q}_o \tag{2.115}$$

Furthermore the observable canonical form (2.90) is obtained by the following similarity transformation:

$$\underline{x}(t) = \mathbf{P}_o\underline{x}_o(t) \Leftrightarrow \underline{x}_o(t) = \mathbf{P}_o^{-1}\underline{x}(t) \tag{2.116}$$

Alternatively the constant nonsingular matrix $\mathbf{P}_o$ can be obtained through the state matrix $\mathbf{A}$ and the last column $\underline{q}_o$ of the inverse of the observability matrix $\mathbf{Q}_o$ as follows:

$$\mathbf{Q}_o^{-1} = \begin{bmatrix} * & \cdots & * & \underline{q}_o \end{bmatrix} \Rightarrow \mathbf{P}_o = \begin{bmatrix} \underline{q}_o & \mathbf{A}\underline{q}_o & \cdots & \mathbf{A}^{n-1}\underline{q}_o \end{bmatrix} \tag{2.117}$$

To get this result we write from (2.8) the following similarity transformation:

$$\mathbf{A}_o = \mathbf{P}_o^{-1}\mathbf{A}\mathbf{P}_o \Leftrightarrow \mathbf{P}_o\mathbf{A}_o = \mathbf{A}\mathbf{P}_o \tag{2.118}$$

Let's denote $\det(s\mathbb{I} - \mathbf{A})$ as follows:

$$\det(s\mathbb{I} - \mathbf{A}) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \tag{2.119}$$

Thus the coefficients a_i of the state matrix $\mathbf{A}_o$ corresponding to the observable canonical form are known and matrix $\mathbf{A}_o$ is written as follows:

$$\mathbf{A}_o = \begin{bmatrix} 0 & 0 & 0 & & -a_0 \\ 1 & 0 & 0 & \ddots & -a_1 \\ 0 & 1 & 0 & \ddots & -a_2 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & & 0 & 1 & -a_{n-1} \end{bmatrix} \tag{2.120}$$

Furthermore let's write the unknown matrix $\mathbf{P}_o$ as follows:

$$\mathbf{P}_o = \begin{bmatrix} \underline{c}_1 & \cdots & \underline{c}_n \end{bmatrix} \tag{2.121}$$

Thus the columns of the unknown change of basis matrix $\mathbf{P}_o$ can be obtained thanks to the following similarity transformation:

$$\begin{aligned} &\mathbf{P}_o\mathbf{A}_o = \mathbf{A}\mathbf{P}_o \\ \Leftrightarrow &\begin{bmatrix} \underline{c}_1 & \cdots & \underline{c}_n \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & & -a_0 \\ 1 & 0 & 0 & \ddots & -a_1 \\ 0 & 1 & 0 & \ddots & -a_2 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & & 0 & 1 & -a_{n-1} \end{bmatrix} = \mathbf{A} \begin{bmatrix} \underline{c}_1 & \cdots & \underline{c}_n \end{bmatrix} \end{aligned} \tag{2.122}$$

Working out with the first $n-1^{th}$ columns gives the following equations:

$$\begin{cases} \underline{c}_2 = \mathbf{A}\underline{c}_1 \\ \underline{c}_3 = \mathbf{A}\underline{c}_2 = \mathbf{A}^2\underline{c}_1 \\ \vdots \\ \underline{c}_n = \mathbf{A}\underline{c}_{n-1} = \mathbf{A}^{n-1}\underline{c}_1 \end{cases} \tag{2.123}$$

Furthermore from (2.8) we get the relation $\mathbf{CP}_o = \mathbf{C}_o$ which is rewritten as follows:

$$\begin{aligned}
&\mathbf{CP}_o = \mathbf{C}_o \\
&\Leftrightarrow \mathbf{C}\begin{bmatrix} \underline{c}_1 & \cdots & \underline{c}_n \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix} \\
&\Leftrightarrow \begin{cases} \mathbf{C}\underline{c}_1 = 0 \\ \mathbf{C}\underline{c}_2 = 0 \\ \vdots \\ \mathbf{C}\underline{c}_{n-1} = 0 \\ \mathbf{C}\underline{c}_n = 1 \end{cases}
\end{aligned} \tag{2.124}$$

Combining (2.123) and (2.124) we get:

$$\begin{cases} \mathbf{C}\underline{c}_1 = 0 \\ \mathbf{C}\underline{c}_2 = \mathbf{CA}\underline{c}_1 = 0 \\ \vdots \\ \mathbf{C}\underline{c}_{n-1} = \mathbf{CA}^{n-2}\underline{c}_1 = 0 \\ \mathbf{C}\underline{c}_n = \mathbf{CA}^{n-1}\underline{c}_1 = 1 \end{cases} \tag{2.125}$$

These equations can in turn be written in matrix form as:

$$\begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-2} \\ \mathbf{CA}^{n-1} \end{bmatrix} \underline{c}_1 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{2.126}$$

Let's introduce the observability matrix $\mathbf{Q}_o$:

$$\mathbf{Q}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \tag{2.127}$$

Assuming that matrix $\mathbf{Q}_o$ has full rank we get:

$$\mathbf{Q}_o\underline{c}_1 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \Leftrightarrow \underline{c}_1 = \mathbf{Q}_o^{-1}\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{2.128}$$

From the preceding equation it is clear that $\underline{c}_1$ is the last column of the inverse of the observability matrix $\mathbf{Q}_o$. We will denote it $\underline{q}_o$:

$$\underline{c}_1 := \underline{q}_o \tag{2.129}$$

Having the expression of $\underline{c}_1$ we can then go back to (2.123) and construct all the columns of $\mathbf{P}_o$.

Example 2.5. *We consider the following general state-space representation:*

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}u(t) \\ y(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}u(t) \end{cases} \tag{2.130}$$

where:

$$\begin{cases} \mathbf{A} = \begin{bmatrix} 28.5 & -17.5 \\ 58.5 & -35.5 \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \\ \mathbf{C} = \begin{bmatrix} 7 & -4 \end{bmatrix} \\ \mathbf{D} = 0.5 \end{cases} \tag{2.131}$$

We are looking for the observable canonical form of this state-space representation.

First we build the observability matrix $\mathbf{Q}_o$ *from* (2.110)*:*

$$\mathbf{Q}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \end{bmatrix} = \begin{bmatrix} 7 & -4 \\ -34.5 & 19.5 \end{bmatrix} \tag{2.132}$$

To build matrix $\mathbf{Q}_{oo}$ *let's compute* $\det(s\mathbb{I} - \mathbf{A})$*:*

$$\det(s\mathbb{I} - \mathbf{A}) = s^2 + a_1 s + a_0 = s^2 + 7s + 12 \tag{2.133}$$

As soon as matrix $\mathbf{A}_o$ *is built from the denominator of the transfer function, that is from* $\det(s\mathbb{I} - \mathbf{A})$*, we get:*

$$\mathbf{A}_o = \begin{bmatrix} 0 & -a_0 \\ 1 & -a_1 \end{bmatrix} = \begin{bmatrix} 0 & -12 \\ 1 & -7 \end{bmatrix} \tag{2.134}$$

Furthermore matrix $\mathbf{C}_o$ *is straightforward for the observable canonical form:*

$$\mathbf{C}_o = \begin{bmatrix} 0 & 1 \end{bmatrix} \tag{2.135}$$

Thus we are in position to compute matrix $\mathbf{Q}_{oo}$ *:*

$$\mathbf{Q}_{oo} = \begin{bmatrix} \mathbf{C}_o \\ \mathbf{C}_o\mathbf{A}_o \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -7 \end{bmatrix} \tag{2.136}$$

Then we use (2.115) *to build the similarity transformation:*

$$\begin{aligned} \mathbf{P}_o &= \mathbf{Q}_o^{-1}\mathbf{Q}_{oo} = \begin{bmatrix} 7 & -4 \\ -34.5 & 19.5 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 \\ 1 & -7 \end{bmatrix} \\ &= \tfrac{1}{-1.5} \begin{bmatrix} 19.5 & 4 \\ 34.5 & 7 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -7 \end{bmatrix} \\ &= \tfrac{2}{3} \begin{bmatrix} -19.5 & -4 \\ -34.5 & -7 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -7 \end{bmatrix} \\ &= \tfrac{2}{3} \begin{bmatrix} -4 & 8.5 \\ -7 & 14.5 \end{bmatrix} \end{aligned} \tag{2.137}$$

Alternatively we can use (2.117) *to build the similarity transformation:*

$$\begin{aligned} \mathbf{Q}_o^{-1} &= \tfrac{2}{3}\begin{bmatrix} -19.5 & -4 \\ -34.5 & -7 \end{bmatrix} = \begin{bmatrix} * & \underline{q}_o \end{bmatrix} \\ \Rightarrow \underline{q}_o &= \tfrac{2}{3}\begin{bmatrix} -4 \\ -7 \end{bmatrix} \end{aligned} \tag{2.138}$$

And:

$$\mathbf{P}_o = \begin{bmatrix} \underline{q}_o & \mathbf{A}\underline{q}_o \end{bmatrix} = \frac{2}{3}\begin{bmatrix} -4 & 8.5 \\ -7 & 14.5 \end{bmatrix} \tag{2.139}$$

Using the similarity relations (2.8) *we finally get the following observable canonical form of the state-space representation:*

$$\begin{cases} \mathbf{A}_o = \mathbf{P}_o^{-1}\mathbf{A}\mathbf{P}_o = \begin{bmatrix} 14.5 & -8.5 \\ 7 & -4 \end{bmatrix}\begin{bmatrix} 28.5 & -17.5 \\ 58.5 & -35.5 \end{bmatrix}\tfrac{2}{3}\begin{bmatrix} -4 & 8.5 \\ -7 & 14.5 \end{bmatrix} = \begin{bmatrix} 0 & -12 \\ 1 & -7 \end{bmatrix} \\ \mathbf{B}_o = \mathbf{P}_o^{-1}\mathbf{B} = \begin{bmatrix} 14.5 & -8.5 \\ 7 & -4 \end{bmatrix}\begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -5 \\ -2 \end{bmatrix} \\ \mathbf{C}_o = \mathbf{C}\mathbf{P}_o = \begin{bmatrix} 7 & -4 \end{bmatrix}\tfrac{2}{3}\begin{bmatrix} -4 & 8.5 \\ -7 & 14.5 \end{bmatrix} = \begin{bmatrix} 0 & 1 \end{bmatrix} \end{cases} \tag{2.140}$$

■

Iterative method

Equivalently the inverse of the change of basis matrix $\mathbf{P}_o$ of the similarity transformation can be obtained as follows:

$$\mathbf{P}_o^{-1} = \begin{bmatrix} \underline{r}_1^T \\ \underline{r}_2^T \\ \vdots \\ \underline{r}_n^T \end{bmatrix} \tag{2.141}$$

where:

$$\begin{cases} \det(s\mathbb{I} - \mathbf{A}) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \\ \underline{r}_n^T = \mathbf{C} \\ \underline{r}_k^T = \underline{r}_{k+1}^T\mathbf{A} + a_k\mathbf{C} \quad \forall\, n-1 \geq k \geq 1 \end{cases} \tag{2.142}$$

To get this result we write from (2.8) the following similarity transformation:

$$\mathbf{A}_o = \mathbf{P}_o^{-1}\mathbf{A}\mathbf{P}_o \Leftrightarrow \mathbf{A}_o\mathbf{P}_o^{-1} = \mathbf{P}_o^{-1}\mathbf{A} \tag{2.143}$$

Let's denote $\det(s\mathbb{I} - \mathbf{A})$ as follows:

$$\det(s\mathbb{I} - \mathbf{A}) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \tag{2.144}$$

Thus the coefficients a_i of the state matrix $\mathbf{A}_o$ corresponding to the observable canonical form are known and matrix $\mathbf{A}_o$ is written as follows:

$$\mathbf{A}_o = \begin{bmatrix} 0 & 0 & 0 & & -a_0 \\ 1 & 0 & 0 & \ddots & -a_1 \\ 0 & 1 & 0 & \ddots & -a_2 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & & 0 & 1 & -a_{n-1} \end{bmatrix} \tag{2.145}$$

Furthermore let's write the inverse of the unknown change of basis matrix $\mathbf{P}_o$ as follows:

$$\mathbf{P}_o^{-1} = \begin{bmatrix} \underline{r}_1^T \\ \underline{r}_2^T \\ \vdots \\ \underline{r}_n^T \end{bmatrix} \tag{2.146}$$

Thus the columns of the unknown matrix $\mathbf{P}_o$ can be obtained thanks to the similarity transformation:

$$\begin{aligned} &\mathbf{A}_o\mathbf{P}_o^{-1} = \mathbf{P}_o^{-1}\mathbf{A} \\ \Leftrightarrow &\begin{bmatrix} 0 & 0 & 0 & & -a_0 \\ 1 & 0 & 0 & \ddots & -a_1 \\ 0 & 1 & 0 & \ddots & -a_2 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & & 0 & 1 & -a_{n-1} \end{bmatrix} \begin{bmatrix} \underline{r}_1^T \\ \underline{r}_2^T \\ \vdots \\ \underline{r}_n^T \end{bmatrix} = \begin{bmatrix} \underline{r}_1^T \\ \underline{r}_2^T \\ \vdots \\ \underline{r}_n^T \end{bmatrix} \mathbf{A} \end{aligned} \tag{2.147}$$

That is:

$$\begin{cases} -a_0\underline{r}_n^T = \underline{r}_1^T\mathbf{A} \\ \underline{r}_1^T - a_1\underline{r}_n^T = \underline{r}_2^T\mathbf{A} \\ \vdots \\ \underline{r}_{n-1}^T - a_{n-1}\underline{r}_n^T = \underline{r}_n^T\mathbf{A} \end{cases} \Leftrightarrow \begin{cases} \underline{0} = \underline{r}_1^T\mathbf{A} + a_0\underline{r}_n^T \\ \underline{r}_k^T = \underline{r}_{k+1}^T\mathbf{A} + a_k\underline{r}_n^T \quad \forall\ n-1 \geq k \geq 1 \end{cases} \tag{2.148}$$

Furthermore from (2.8) we get the relation $\mathbf{C}_o = \mathbf{C}\mathbf{P}_o$ which is rewritten as follows:

$$\mathbf{C}_o\mathbf{P}_o^{-1} = \mathbf{C} \Leftrightarrow \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} \underline{r}_1^T \\ \underline{r}_2^T \\ \vdots \\ \underline{r}_n^T \end{bmatrix} = \mathbf{C} \Rightarrow \underline{r}_n^T = \mathbf{C} \tag{2.149}$$

Combining the last equation of (2.148) with (2.149) gives the proposed result:

$$\begin{cases} \underline{r}_n^T = \mathbf{C} \\ \underline{r}_k^T = \underline{r}_{k+1}^T\mathbf{A} + a_k\mathbf{C} \quad \forall\ n-1 \geq k \geq 1 \end{cases} \tag{2.150}$$

Example 2.6. *We consider the following general state-space representation:*

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}u(t) \\ y(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}u(t) \end{cases} \tag{2.151}$$

where:

$$\begin{cases} \mathbf{A} = \begin{bmatrix} 28.5 & -17.5 \\ 58.5 & -35.5 \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \\ \mathbf{C} = \begin{bmatrix} 7 & -4 \end{bmatrix} \\ \mathbf{D} = 0.5 \end{cases} \tag{2.152}$$

This is the same state-space representation than the one which has been used in the previous example. We have seen that the similarity transformation which leads to the controllable canonical form is the following:

$$\mathbf{P}_o = \frac{2}{3} \begin{bmatrix} -4 & 8.5 \\ -7 & 14.5 \end{bmatrix} \tag{2.153}$$

It is easy to compute matrix $\mathbf{P}_o^{-1}$*, that is the inverse of* $\mathbf{P}_o$*. We get the following expression:*

$$\mathbf{P}_o^{-1} = \begin{bmatrix} 14.5 & -8.5 \\ 7 & -4 \end{bmatrix} \tag{2.154}$$

We will check the expression of matrix $\mathbf{P}_o^{-1}$ *thanks to the iterative method proposed in* (2.142)*. We get:*

$$\begin{cases} \det(s\mathbb{I} - \mathbf{A}) = s^2 + 7s + 12 \\ \underline{r}_2^T = \mathbf{C} = \begin{bmatrix} 7 & -4 \end{bmatrix} \\ \underline{r}_1^T = \underline{r}_2^T \mathbf{A} + a_1 \mathbf{C} = \begin{bmatrix} 7 & -4 \end{bmatrix} \begin{bmatrix} 28.5 & -17.5 \\ 58.5 & -35.5 \end{bmatrix} + 7 \begin{bmatrix} 7 & -4 \end{bmatrix} = \begin{bmatrix} 14.5 & -8.5 \end{bmatrix} \end{cases} \tag{2.155}$$

Thus we fortunately retrieve the expression of matrix $\mathbf{P}_o^{-1}$*:*

$$\mathbf{P}_o^{-1} = \begin{bmatrix} \underline{r}_1^T \\ \underline{r}_2^T \end{bmatrix} = \begin{bmatrix} 14.5 & -8.5 \\ 7 & -4 \end{bmatrix} \tag{2.156}$$

■

2.3.6 Diagonal (or modal) form

One particular useful canonical form is called the (block) diagonal or modal form. The (block) diagonal form is obtained thanks to the partial fraction expansion of transfer function $F(s)$. This is a diagonal representation of the state-space model when all the poles of $F(s)$ are distinct. Otherwise this is a Jordan representation.

Transfer function with distinct poles

Assume for now that transfer function $F(s)$ has distinct poles λ_i. Then the partial fraction expansion of $F(s)$ reads:

$$\begin{aligned} \frac{Y(s)}{U(s)} = F(s) &= \frac{N(s)}{D(s)} + d \\ &= \frac{N(s)}{(s-\lambda_1)(s-\lambda_2)\cdots(s-\lambda_n)} + d \\ &= \frac{r_1}{s-\lambda_1} + \frac{r_2}{s-\lambda_2} + \cdots + \frac{r_n}{s-\lambda_n} + d \end{aligned} \tag{2.157}$$

Number r_i is called the residue of transfer function $F(s)$ in λ_i. When the multiplicity of the pole (or eigenvalue) λ_i is 1 it is clear from the preceding relation that the residue r_i can be obtained thanks to the following formula:

$$r_i = (s-\lambda_i)F(s)|_{s=\lambda_i} \tag{2.158}$$

Now we define constants b_i and c_i such that the product $b_i c_i$ is equal to r_i:

$$r_i = c_i\, b_i \tag{2.159}$$

Consequently transfer function $F(s)$ can be written as follows:

$$F(s) = \frac{c_1\, b_1}{s-\lambda_1} + \frac{c_2\, b_2}{s-\lambda_2} + \cdots + \frac{c_n\, b_n}{s-\lambda_n} + d \tag{2.160}$$

Then we define the Laplace transform of the components $x_1(t), x_2(t), \cdots, x_n(t)$ of the state vector $\underline{x}(t)$ as follows:

$$\begin{cases} \frac{X_1(s)}{U(s)} = \frac{b_1}{s-\lambda_1} \\ \frac{X_2(s)}{U(s)} = \frac{b_2}{s-\lambda_2} \\ \vdots \\ \frac{X_n(s)}{U(s)} = \frac{b_n}{s-\lambda_n} \end{cases} \tag{2.161}$$

Using (2.161) transfer function $F(s)$ can be written as follows:

$$\begin{aligned} &\frac{Y(s)}{U(s)} = F(s) = c_1\frac{X_1(s)}{U(s)} + c_2\frac{X_2(s)}{U(s)} + \cdots + c_n\frac{X_n(s)}{U(s)} + d \\ &\Rightarrow Y(s) = c_1X_1(s) + c_2X_2(s) + \cdots + c_nX_n(s) + dU(s) \end{aligned} \tag{2.162}$$

Coming back to the time domain we get:

$$y(t) = c_1x_1(t) + c_2x_2(t) + \cdots + c_nx_n(t) + du(t) \tag{2.163}$$

Whereas in the time domain (2.161) reads:

$$\begin{cases} \dot{x}_1(t) = \lambda_1x_1(t) + b_1u(t) \\ \dot{x}_2(t) = \lambda_2x_2(t) + b_2u(t) \\ \vdots \\ \dot{x}_n(t) = \lambda_nx_n(t) + b_nu(t) \end{cases} \tag{2.164}$$

Equations (2.164) and (2.163) lead to the following state-space representation, which is called the diagonal or modal form:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}u(t) \\ y(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}u(t) \end{cases} \tag{2.165}$$

where:

$$\begin{cases} \mathbf{A} = \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\ \mathbf{C} = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix} \\ \mathbf{D} = d \end{cases} \tag{2.166}$$

Example 2.7. *Let's consider the following transfer function:*

$$F(s) = \frac{(s+1)(s+2)}{2(s+3)(s+4)} = \frac{s^2+3s+2}{2s^2+14s+24} \tag{2.167}$$

We are looking for the diagonal form of this transfer function.

First we have to set to 1 the leading coefficient of the polynomial which appears in the denominator of the transfer function $F(s)$. We get:

$$F(s) = \frac{0.5s^2+1.5s+1}{1 \times s^2+7s+12} \tag{2.168}$$

Then we decompose $F(s)$ as a sum between a strictly proper rational function and a constant coefficient d. Constant coefficient d is obtained thanks to the following relation:

$$d = \lim_{s\to\infty} \mathbf{F}(s) = \lim_{s\to\infty} \frac{0.5s^2+1.5s+1}{1 \times s^2+7s+12} = 0.5 \tag{2.169}$$

Thus the strictly proper transfer function $N(s)/D(s)$ is obtained by subtracting d to $F(s)$:

$$\frac{N(s)}{D(s)} = F(s) - d = \frac{0.5s^2+1.5s+1}{1 \times s^2+7s+12} - 0.5 = \frac{-2s-5}{s^2+7s+12} \tag{2.170}$$

The two poles of $F(s)$ are -3 and -4. Thus the partial fraction expansion of $F(s)$ reads:

$$F(s) = \frac{r_1}{s+3} + \frac{r_2}{s+4} + d = \frac{r_1}{s+3} + \frac{r_2}{s+4} - 0.5 \tag{2.171}$$

where the residues r_1 and r_2 are:

$$\begin{cases} r_1 = (s+3)F(s)|_{s=-3} \left.\frac{(s+1)(s+2)}{2(s+4)}\right|_{s=-3} = \frac{(-3+1)(-3+2)}{2(-3+4)} = 1 \\ r_2 = (s+4)F(s)|_{s=-4} \left.\frac{(s+1)(s+2)}{2(s+3)}\right|_{s=-4} = \frac{(-4+1)(-4+2)}{2(-4+3)} = -3 \end{cases} \tag{2.172}$$

We finally get:

$$F(s) = \frac{N(s)}{D(s)} + d = \frac{1}{s+3} + \frac{-3}{s+4} + 0.5 \tag{2.173}$$

Residues r_1 and r_2 are expressed for example as follows:

$$\begin{cases} r_1 = 1 = 1 \times 1 = c_1 \times b_1 \\ r_2 = -3 = -3 \times 1 = c_2 \times b_2 \end{cases} \tag{2.174}$$

Then we apply Equation (2.166) to get the diagonal canonical form of $F(s)$:

$$\begin{cases} \mathbf{A} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 0 & -4 \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \end{bmatrix} \\ \mathbf{C} = \begin{bmatrix} c_1 & c_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \end{bmatrix} \\ \mathbf{D} = d = 0.5 \end{cases} \tag{2.175}$$

■

Similarity transformation to diagonal form

Assume that state matrix $\mathbf{A}$ has distinct eigenvalues λ_i. Starting from a realization $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & d \end{array}\right)$ let $\mathbf{P}_m$ be the change of basis matrix such that:

$$\mathbf{A}_m = \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} = \mathbf{P}_m^{-1}\mathbf{A}\mathbf{P}_m \tag{2.176}$$

We will denote $\mathbf{P}_m$ as follows:

$$\mathbf{P}_m = \begin{bmatrix} \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_n \end{bmatrix} \tag{2.177}$$

It can be seen that vectors $\underline{v}_i$ are the eigenvectors of matrix $\mathbf{A}$. Indeed let λ_i be an eigenvalue of $\mathbf{A}$. Then:

$$\begin{cases} \mathbf{A}\underline{v}_1 = \lambda_1 \underline{v}_1 \\ \vdots \\ \mathbf{A}\underline{v}_n = \lambda_n \underline{v}_n \end{cases} \tag{2.178}$$

Writing this equation in vector form leads to the following relation:

$$\begin{bmatrix} \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} = \mathbf{A} \begin{bmatrix} \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_n \end{bmatrix} \tag{2.179}$$

That is:

$$\mathbf{P}_m \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} = \mathbf{A}\mathbf{P}_m \tag{2.180}$$

Or equivalently:

$$\begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} = \mathbf{P}_m^{-1}\mathbf{A}\mathbf{P}_m \tag{2.181}$$

The inverse of the change of basis matrix $\mathbf{P}_n$ can be written in terms of rows as follows:

$$\mathbf{P}_m^{-1} = \begin{bmatrix} \underline{w}_1^T \\ \underline{w}_2^T \\ \vdots \\ \underline{w}_n^T \end{bmatrix} \tag{2.182}$$

It can be seen that vectors $\underline{w}_i$ are the eigenvectors of matrix $\mathbf{A}^T$. Indeed let λ_i be an eigenvalue of $\mathbf{A}$, which is also an eigenvalue of $\mathbf{A}^T$ as far as $\det(s\mathbb{I} - \mathbf{A}) = \det(s\mathbb{I} - \mathbf{A})^T = \det\left(s\mathbb{I} - \mathbf{A}^T\right)$. Then:

$$\mathbf{A}^T\underline{w}_i = \lambda_i\underline{w}_i \Rightarrow \underline{w}_i^T\mathbf{A} = \lambda_i\underline{w}_i^T \tag{2.183}$$

Thus by multiplying by $\underline{v}_j$ and using the fact that $\underline{v}_j$ is an eigenvector of $\mathbf{A}$, that is $\mathbf{A}\underline{v}_j = \lambda_j\underline{v}_j$, we get:

$$\lambda_i\underline{w}_i^T\underline{v}_j = \underline{w}_i^T\mathbf{A}\underline{v}_j = \lambda_j\underline{w}_i^T\underline{v}_j \Rightarrow (\lambda_i - \lambda_j)\,\underline{w}_i^T\underline{v}_j = 0 \tag{2.184}$$

Since $\lambda_i \neq \lambda_j\ \forall i \neq j$ we finally get:

$$\underline{w}_i^T\underline{v}_j = 0 \text{ if } i \neq j \tag{2.185}$$

As far as $\underline{w}_i$ and $\underline{v}_j$ are defined to within a constant we impose $\underline{w}_i^T\underline{v}_j = 1$ if $i = j$. Consequently:

$$\underline{w}_i^T\underline{v}_j = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} \tag{2.186}$$

Transfer function with complex conjugate pair of poles

If some of the poles are complex so are the residues and so is the diagonal form. This may be inconvenient. We will see hereafter how to retrieve real matrices corresponding to the diagonal form.

Assume that λ and $\overline{\lambda}$ is a complex conjugate pair of poles of $F(s)$:

$$F(s) = \frac{r_1}{s-\lambda} + \frac{\overline{r}_1}{s-\overline{\lambda}} \tag{2.187}$$

Let α be the real part of the pole λ and β its imaginary part:

$$\lambda = \alpha + j\beta \Leftrightarrow \overline{\lambda} = \alpha - j\beta \tag{2.188}$$

According to the preceding section the state-space representation of $F(s)$ is the following:

$$\begin{cases} \mathbf{A}_m = \begin{bmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{bmatrix} = \begin{bmatrix} \alpha + j\beta & 0 \\ 0 & \alpha - j\beta \end{bmatrix} \\ \mathbf{B}_m = \begin{bmatrix} b_1 \\ \overline{b}_1 \end{bmatrix} \\ \mathbf{C}_m = \begin{bmatrix} c_1 & \overline{c}_1 \end{bmatrix} \\ \mathbf{D} = 0 \end{cases} \tag{2.189}$$

Where:

$$r_1 = b_1 c_1 \Rightarrow \overline{r}_1 = \overline{b}_1 \overline{c}_1 \tag{2.190}$$

It is clear that the diagonal form of transfer function $F(s)$ is complex. From the preceding realization we get the following equations:

$$\begin{cases} \dot{x}_1(t) = (\alpha + j\beta)\, x_1(t) + b_1 u(t) \\ \dot{x}_2(t) = (\alpha - j\beta)\, x_2(t) + \overline{b}_1 u(t) \end{cases} \tag{2.191}$$

We deduce from the preceding equation that the state components $x_1(t)$ and $x_2(t)$ are complex conjugate. Let $x_R(t)$ be the real part of $x_1(t)$ and $x_I(t)$ its imaginary part:

$$x_1(t) = x_R(t) + jx_I(t) \Rightarrow x_2(t) = \overline{x}_1(t) = x_R(t) - jx_I(t) \tag{2.192}$$

Thus Equation (2.191) reads:

$$\begin{cases} \dot{x}_R(t) + j\dot{x}_I(t) = (\alpha + j\beta)\,(x_R(t) + jx_I(t)) + b_1 u(t) \\ \dot{x}_R(t) - j\dot{x}_I(t) = (\alpha - j\beta)\,(x_R(t) - jx_I(t)) + \overline{b}_1 u(t) \end{cases} \tag{2.193}$$

We deduce two new equations from the two preceding equations as follows: the first new equation is obtained by adding the two preceding equations and

dividing the result by 2 whereas the second new equation is obtained by subtracting the two preceding equations and dividing the result by $2j$. We get:

$$\begin{cases} \dot{x}_R(t) = \alpha x_R(t) - \beta x_I(t) + \frac{b_1+\bar{b}_1}{2} u(t) \\ \dot{x}_I(t) = \beta x_R(t) + \alpha x_I(t) + \frac{b_1-\bar{b}_1}{2j} u(t) \end{cases} \tag{2.194}$$

As far as the output $y(t)$ is concerned we can express it as a function of the new components $x_R(t)$ and $x_I(t)$ of the state vector:

$$\begin{aligned} y(t) &= c_1 x_1(t) + \bar{c}_1 \bar{x}_1(t) \\ &= c_1 \left(x_R(t) + j x_I(t)\right) + \bar{c}_1 \left(x_R(t) - j x_I(t)\right) \\ &= \left(c_1 + \bar{c}_1\right) x_R(t) + j \left(c_1 - \bar{c}_1\right) x_I(t) \end{aligned} \tag{2.195}$$

Consequently the complex diagonal form in Equation (2.189) is rendered real by using the real part and the imaginary part of the complex state component which appear in the state vector rather than the complex state component and its conjugate. Indeed Equations (2.194) and (2.195) lead to the following state-space representation where matrices $(\mathbf{A}_m, \mathbf{B}_m, \mathbf{C}_m, \mathbf{D})$ are real:

$$\begin{cases} \mathbf{A}_m = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \\ \mathbf{B}_m = \begin{bmatrix} \frac{b_1+\bar{b}_1}{2} \\ \frac{b_1-\bar{b}_1}{2j} \end{bmatrix} \\ \mathbf{C}_m = \begin{bmatrix} (c_1 + \bar{c}_1) & j\,(c_1 - \bar{c}_1) \end{bmatrix} \\ \mathbf{D} = 0 \end{cases} \tag{2.196}$$

It can be seen that complex matrix $\mathbf{A}$ has the same determinant than the real matrix $\mathbf{A}_m$:

$$\det\left(s\mathbb{I} - \begin{bmatrix} \alpha + j\beta & 0 \\ 0 & \alpha - j\beta \end{bmatrix}\right) = \det\left(s\mathbb{I} - \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}\right) \tag{2.197}$$

Example 2.8. *Let's consider the following transfer function:*

$$F(s) = \frac{s+2}{s^2 - 2s + 5} \tag{2.198}$$

The two poles of $F(s)$ are $\lambda_1 = 1 + 2j$ and $\overline{\lambda}_1 = 1 - 2j$. Thus the partial fraction expansion of $F(s)$ reads:

$$F(s) = \frac{s+2}{s^2 - 2s + 5} = \frac{s+2}{(s-\lambda_1)(s-\overline{\lambda}_1)} = \frac{r_1}{s-\lambda_1} + \frac{r_2}{s-\overline{\lambda}_1} \quad \textit{where } \lambda_1 = 1 + 2j \tag{2.199}$$

where the residues r_1 and r_2 are:

$$\begin{cases} r_1 = (s-\lambda_1)F(s)\big|_{s=\lambda_1} = \left.\frac{s+2}{s-\overline{\lambda}_1}\right|_{s=\lambda_1} = \frac{3+2j}{4j} = \frac{2-3j}{4} \\ r_2 = (s-\overline{\lambda}_1)F(s)\big|_{s=\overline{\lambda}_1} = \left.\frac{s+2}{s-\lambda_1}\right|_{s=\overline{\lambda}_1} = \frac{3-2j}{-4j} = \frac{2+3j}{4} = \overline{r}_1 \end{cases} \tag{2.200}$$

We finally get:

$$F(s) = \frac{r_1}{s-\lambda_1} + \frac{r_2}{s-\overline{\lambda}_1} = \frac{\frac{2-3j}{4}}{s-(1+2j)} + \frac{\frac{2+3j}{4}}{s-(1-2j)} \tag{2.201}$$

Residues r_1 and r_2 are expressed for example as follows:

$$\begin{cases} r_1 = \frac{2-3j}{4} = (2-3j) \times \frac{1}{4} = c_1 \times b_1 \\ r_2 = \frac{2+3j}{4} = (2+3j) \times \frac{1}{4} = \bar{c}_1 \times \bar{b}_1 \end{cases} \tag{2.202}$$

Then we apply Equation (2.166) *to get the diagonal canonical form of $F(s)$:*

$$\begin{cases} \mathbf{A}_m = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \overline{\lambda}_1 \end{bmatrix} = \begin{bmatrix} \alpha + j\beta & 0 \\ 0 & \alpha - j\beta \end{bmatrix} = \begin{bmatrix} 1+2j & 0 \\ 0 & 1-2j \end{bmatrix} \\ \mathbf{B}_m = \begin{bmatrix} b_1 \\ \overline{b}_1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \mathbf{C}_m = \begin{bmatrix} c_1 & \overline{c}_1 \end{bmatrix} = \begin{bmatrix} 2-3j & 2+3j \end{bmatrix} \\ \mathbf{D} = 0 \end{cases} \tag{2.203}$$

This complex diagonal form realization is rendered real by using (2.196)*:*

$$\begin{cases} \mathbf{A}_m = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \\ \mathbf{B}_m = \begin{bmatrix} \frac{b_1+\overline{b}_1}{2} \\ \frac{b_1-\overline{b}_1}{2j} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \mathbf{C}_m = \begin{bmatrix} (c_1+\overline{c}_1) & j\,(c_1-\overline{c}_1) \end{bmatrix} = \begin{bmatrix} 4 & 6 \end{bmatrix} \\ \mathbf{D} = 0 \end{cases} \tag{2.204}$$

For both realizations we can check that $F(s) = \mathbf{C}_m\,(s\mathbb{I} - \mathbf{A}_m)^{-1}\,\mathbf{B}_m + \mathbf{D}$ but in the last realzation matrices $(\mathbf{A}_m, \mathbf{B}_m, \mathbf{C}_m, \mathbf{D})$ are real.

■

2.3.7 Algebraic and geometric multiplicity of an eigenvalue

The *algebraic multiplicity* n_i of an eigenvalue λ_i of matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is the number of times λ_i appears as a root of the characteristic polynomial $\det(s\mathbb{I}-\mathbf{A})$.

The *geometric multiplicity* q_i of an eigenvalue λ_i of matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is the dimension of the kernel of $\lambda_i\mathbb{I} - \mathbf{A}$.

If for every eigenvalue of $\mathbf{A}$ the geometric multiplicity equals the algebraic multiplicity, then matrix $\mathbf{A}$ is said to be diagonalizable.

Example 2.9. *Let's consider the following matrix:*

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} \tag{2.205}$$

We have:

$$\det(s\mathbb{I} - \mathbf{A}) = (s-2)^2 \tag{2.206}$$

Consequently the algebraic multiplicity *of eigenvalue* $\lambda_1 = 2$ *is* $n_1 = 2$.

In order to get the geometric multiplicity *of eigenvalue* λ_1 *we consider the following matrix:*

$$\lambda_1\mathbb{I} - \mathbf{A} = \begin{bmatrix} 0 & -3 \\ 0 & 0 \end{bmatrix} \tag{2.207}$$

The dimension of the kernel of $\lambda_1\mathbb{I} - \mathbf{A}$ *is clearly* 1. *Consequently the geometric multiplicity of eigenvalue* λ_1 *is* $q_1 = 1$.

■

2.3.8 Jordan form and generalized eigenvectors

Matrix $\mathbf{A}$ is not diagonalizable if there is at least one eigenvalue with geometric multiplicity q_i (dimension of its eigenspace) strictly less than its algebraic multiplicity n_i. Equivalently, if for every eigenvalue λ_i of $\mathbf{A}$ the geometric multiplicity equals the algebraic multiplicity then $\mathbf{A}$ is diagonalizable. If not, the diagonal form of matrix $\mathbf{A}$ is replaced by its Jordan form which is achieved through the so-called *generalized* eigenvectors.

Indeed, any square matrix $\mathbf{A}$ is similar to the a block diagonal matrix through an invertible matrix $\mathbf{P}$:

$$\mathbf{AP} = \mathbf{P}\begin{bmatrix} \mathbf{J}_{\lambda_1} & & \\ & \ddots & \\ & & \mathbf{J}_{\lambda_p} \end{bmatrix} \tag{2.208}$$

where the Jordan block $\mathbf{J}_{\lambda_i}$ of matrix $\mathbf{A}$ corresponding to eigenvalue λ_i with algebraic multiplicity n_i is the following $n_i \times n_i$ matrix:

$$\mathbf{J}_{\lambda_i} = \underbrace{\begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{bmatrix}}_{n_i \text{ terms}} = \lambda_i\,\mathbb{I}_{n_i} + \mathbf{N} \tag{2.209}$$

Matrix $\mathbf{N}$ is a nilpotent matrix. The number of 1 over the diagonal (the *superdiagonal*) of $\mathbf{J}_{\lambda_i}$ (and $\mathbf{N}$) is equal to the difference between the algebraic multiplicity n_i and the geometric multiplicity (dimension of its eigenspace) q_i of eigenvalue λ_i. When the algebraic multiplicity of eigenvalue λ_i is strictly greater than its geometric multiplicity (dimension of its eigenspace), we said that λ_i is *degenerate*.

For example, assume that matrix $\mathbf{A}$ has one eigenvalue λ_i with algebraic multiplicity $n_i = 3$ and geometric multiplicity (dimension of its eigenspace) equals to $q_i = 1$. Then λ_i is *degenerate* and matrix $\mathbf{A}$ is not diagonalizable. Its Jordan form has $n_i - q_i = 3 - 1 = 2$ values 1 over its diagonal (the *superdiagonal*):

$$\mathbf{J}_{\lambda_i} = \begin{bmatrix} \lambda_i & 1 & 0 \\ 0 & \lambda_i & 1 \\ 0 & 0 & \lambda_i \end{bmatrix} = \lambda_i \mathbb{I}_3 + \mathbf{N} \text{ where } \mathbf{N} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{2.210}$$

Then matrix $\mathbf{P}$, which has to be found, reads:

$$\mathbf{P} = \begin{bmatrix} \underline{v}_{\lambda_i,1} & \underline{v}_{\lambda_i,2} & \underline{v}_{\lambda_i,3} \end{bmatrix} \tag{2.211}$$

and (2.208) reads:

$$\mathbf{A} \begin{bmatrix} \underline{v}_{\lambda_i,1} & \underline{v}_{\lambda_i,2} & \underline{v}_{\lambda_i,3} \end{bmatrix} = \begin{bmatrix} \underline{v}_{\lambda_i,1} & \underline{v}_{\lambda_i,2} & \underline{v}_{\lambda_i,3} \end{bmatrix} \begin{bmatrix} \lambda_i & 1 & 0 \\ 0 & \lambda_i & 1 \\ 0 & 0 & \lambda_i \end{bmatrix} \tag{2.212}$$

Developing each column of the previous product leads to the following relations:

$$\begin{cases} \mathbf{A}\,\underline{v}_{\lambda_i,1} = \lambda_i\,\underline{v}_{\lambda_i,1} \\ \mathbf{A}\,\underline{v}_{\lambda_i,2} = \underline{v}_{\lambda_i,1} + \lambda_i\,\underline{v}_{\lambda_i,2} \\ \mathbf{A}\,\underline{v}_{\lambda_i,3} = \underline{v}_{\lambda_i,2} + \lambda_i\,\underline{v}_{\lambda_i,3} \end{cases} \Leftrightarrow \begin{cases} (\mathbf{A} - \lambda_i\mathbb{I})\,\underline{v}_{\lambda_i,1} = \underline{0} \\ (\mathbf{A} - \lambda_i\mathbb{I})\,\underline{v}_{\lambda_i,2} = \underline{v}_{\lambda_i,1} \\ (\mathbf{A} - \lambda_i\mathbb{I})\,\underline{v}_{\lambda_i,3} = \underline{v}_{\lambda_i,2} \end{cases} \tag{2.213}$$

The first equality indicates that $\underline{v}_{\lambda_i,1}$ belongs to the kernel of $\mathbf{A} - \lambda_i\mathbb{I}$, which is also the kernel of $\lambda_i\mathbb{I} - \mathbf{A}$. Once $\underline{v}_{\lambda_i,1}$ has been chosen, generalized eigenvectors $\underline{v}_{\lambda_i,2}$ and $\underline{v}_{\lambda_i,3}$ shall be chosen such that (2.213) holds.

Alternatively, relation (2.213) indicates that as soon as $\underline{v}_{\lambda_i,3} \neq \underline{0}$ is known, then $\underline{v}_{\lambda_i,2}$ and $\underline{v}_{\lambda_i,1}$ can be computed recursively as follows:

$$\begin{cases} \underline{v}_{\lambda_i,2} = (\mathbf{A} - \lambda_i\mathbb{I})\,\underline{v}_{\lambda_i,3} \\ \underline{v}_{\lambda_i,1} = (\mathbf{A} - \lambda_i\mathbb{I})\,\underline{v}_{\lambda_i,2} \end{cases} \tag{2.214}$$

To get $\underline{v}_{\lambda_i,3}$, it is worth noticing that by multiplying the second equation of (2.213) by $\mathbf{A} - \lambda_i\mathbb{I}$ and the third equation of (2.213) by $(\mathbf{A} - \lambda_i\mathbb{I})^2$, and using the fact that $(\mathbf{A} - \lambda_i\mathbb{I})\,\underline{v}_{\lambda_i,1} = \underline{0}$, we get:

$$\begin{cases} (\mathbf{A} - \lambda_i\mathbb{I})\,\underline{v}_{\lambda_i,1} = \underline{0} \\ (\mathbf{A} - \lambda_i\mathbb{I})^2\,\underline{v}_{\lambda_i,2} = (\mathbf{A} - \lambda_i\mathbb{I})\,\underline{v}_{\lambda_i,1} \\ (\mathbf{A} - \lambda_i\mathbb{I})^3\,\underline{v}_{\lambda_i,3} = (\mathbf{A} - \lambda_i\mathbb{I})^2\,\underline{v}_{\lambda_i,2} \end{cases} \Rightarrow \begin{cases} (\mathbf{A} - \lambda_i\mathbb{I})\,\underline{v}_{\lambda_i,1} = \underline{0} \\ (\mathbf{A} - \lambda_i\mathbb{I})^2\,\underline{v}_{\lambda_i,2} = \underline{0} \\ (\mathbf{A} - \lambda_i\mathbb{I})^3\,\underline{v}_{\lambda_i,3} = \underline{0} \end{cases} \tag{2.215}$$

From the relation $(\mathbf{A} - \lambda_i\mathbb{I})^3\,\underline{v}_{\lambda_i,3} = \underline{0}$, it is clear that $\underline{v}_{\lambda_i,3} \neq \underline{0}$ belongs to the nullspace of $(\mathbf{A} - \lambda_i\mathbb{I})^3$, where $n_i = 3$ is the algebraic multiplicity of eigenvalue λ_i. Moreover, as soon as matrix $\mathbf{P}$ shall be invertible, we shall have $\underline{v}_{\lambda_i,2} \neq \underline{0}$ and $\underline{v}_{\lambda_i,1} \neq \underline{0}$. Thus, we get from (2.214):

$$\begin{cases} \underline{v}_{\lambda_i,2} = (\mathbf{A} - \lambda_i\mathbb{I})\,\underline{v}_{\lambda_i,3} \neq \underline{0} \\ \underline{v}_{\lambda_i,1} = (\mathbf{A} - \lambda_i\mathbb{I})\,\underline{v}_{\lambda_i,2} = (\mathbf{A} - \lambda_i\mathbb{I})^2\,\underline{v}_{\lambda_i,3} \neq \underline{0} \end{cases} \tag{2.216}$$

More generally, to get the Jordan form of any $n \times n$ square matrix $\mathbf{A}$ with eigenvalues λ_i with algebraic multiplicity n_i we must compute the nullspace of $(\mathbf{A} - \lambda_i \mathbb{I})^{n_i}$ or equivalently the nullspace of $(\mathbf{A} - \lambda_i \mathbb{I})^{n}$ (in that situation there is no need to know n_i). Let $\underline{v}_{\lambda_i,n_i}$ be a vector which spans the nullspace of $(\mathbf{A} - \lambda_i \mathbb{I})^{n_i}$ but which does not belong to the nullspace of $(\mathbf{A} - \lambda_i \mathbb{I})^{n_i-1}$:

$$\begin{cases} (\mathbf{A} - \lambda_i \mathbb{I})^{n_i} \, \underline{v}_{\lambda_i,n_i} = \underline{0} \\ (\mathbf{A} - \lambda_i \mathbb{I})^{n_i-1} \, \underline{v}_{\lambda_i,n_i} \neq \underline{0} \end{cases} \tag{2.217}$$

Then the following chain of vectors can be formed:

$$\begin{cases} \underline{v}_{\lambda_i,n_i-1} = (\mathbf{A} - \lambda_i \mathbb{I}) \, \underline{v}_{\lambda_i,n_i} \\ \underline{v}_{\lambda_i,n_i-2} = (\mathbf{A} - \lambda_i \mathbb{I}) \, \underline{v}_{\lambda_i,n_i-1} = (\mathbf{A} - \lambda_i \mathbb{I})^2 \, \underline{v}_{\lambda_i,n_i} \\ \vdots \\ \underline{v}_{\lambda_i,1} = (\mathbf{A} - \lambda_i \mathbb{I})^{n_i-1} \, \underline{v}_{\lambda_i,n_i} \end{cases} \tag{2.218}$$

A nonzero vector $\underline{v}_{\lambda_i}$ which satisfies the following properties is called a *generalized* eigenvector of $\mathbf{A}$ corresponding to eigenvalue λ_i:

$$\begin{cases} (\mathbf{A} - \lambda_i \mathbb{I})^{k} \, \underline{v}_{\lambda_i} = \underline{0} \\ (\mathbf{A} - \lambda_i \mathbb{I})^{k-1} \, \underline{v}_{\lambda_i} \neq \underline{0} \end{cases} \tag{2.219}$$

It is clear that when $k = 1$ the preceding definition leads to the *usual* definition of eigenvector.

It can be shown that:

$$\ker\left((\mathbf{A} - \lambda_i \mathbb{I})^{k}\right) \subset \ker\left((\mathbf{A} - \lambda_i \mathbb{I})^{k+1}\right) \tag{2.220}$$

Furthermore if $\mathbf{A}$ is an $n \times n$ matrix with an eigenvalue λ_i with algebraic multiplicity n_i then there is some integer $\nu_i \leq n_i$ such that the following property holds:

$$dim\left(\ker\left((\mathbf{A} - \lambda_i \mathbb{I})^{\nu_i}\right)\right) = n_i \tag{2.221}$$

Let $\mathbf{P}_{\lambda_i}$ be the matrix formed by the chain of vectors $\begin{bmatrix} \underline{v}_{\lambda_i,1} & \cdots & \underline{v}_{\lambda_i,n_i} \end{bmatrix}$. Then the Jordan form $\mathbf{J}$ of matrix $\mathbf{A}$ of order n corresponding is obtained as:

$$\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \text{ where } \begin{cases} \mathbf{P} = \begin{bmatrix} \mathbf{P}_{\lambda_1} & \cdots & \mathbf{P}_{\lambda_n} \end{bmatrix} \\ \mathbf{P}_{\lambda_i} = \begin{bmatrix} \underline{v}_{\lambda_i,1} & \cdots & \underline{v}_{\lambda_i,n_i} \end{bmatrix} \end{cases} \tag{2.222}$$

2.3.9 Transfer function with multiple poles on the same location

Multiple real poles on the same location

Now assume that transfer function $F(s)$ has a pole λ of multiplicity n. Partial fraction expansion of $F(s)$ results in:

$$\frac{Y(s)}{U(s)} = F(s) = \frac{r_1}{s-\lambda} + \frac{r_2}{(s-\lambda)^2} + \cdots + \frac{r_n}{(s-\lambda)^n} + d \tag{2.223}$$

It is clear from the preceding relation that the numbers $r_i \ \forall\, n \geq i \geq 1$ can be obtained thanks to the following formula:

$$r_i = \frac{1}{(n-i)!}\,\frac{d^{n-i}}{ds^{n-i}}\left((s-\lambda)^n F(s)\right)\Big|_{s=\lambda} \quad \forall\, n \geq i \geq 1 \tag{2.224}$$

Number r_1 is called the residue of transfer function $F(s)$ in λ.

Then we define the Laplace transform of the components $x_1(t), \cdots, x_n(t)$ of the state vector $\underline{x}(t)$ as follows:

$$\frac{X_i(s)}{U(s)} = \frac{1}{(s-\lambda)^{n-i+1}} \quad \forall\, n \geq i \geq 1 \tag{2.225}$$

Using (2.225) transfer function $F(s)$ can be written as follows:

$$\begin{array}{l} \frac{Y(s)}{U(s)} = F(s) = r_1 \frac{X_n(s)}{U(s)} + r_2 \frac{X_{n-1}(s)}{U(s)} + \cdots + r_n \frac{X_1(s)}{U(s)} + d \\ \Rightarrow Y(s) = r_1 X_n(s) + r_2 X_{n-1}(s) + \cdots + r_n X_1(s) + dU(s) \end{array} \tag{2.226}$$

Coming back to the time domain and rearranging the order of the state vector components we get:

$$y(t) = r_n x_1(t) + r_{n-1} x_2(t) + \cdots + r_1 x_n(t) + du(t) \tag{2.227}$$

The n components of the state vector $\underline{x}(t)$ defined by (2.225) reads:

$$\frac{X_i(s)}{U(s)} = \frac{1}{(s-\lambda)^{n-i+1}} \quad \forall\, 1 \leq i \leq n$$

$$\Leftrightarrow \left\{ \begin{array}{l} \frac{X_n(s)}{U(s)} = \frac{1}{s-\lambda} \\ \frac{X_{n-1}(s)}{U(s)} = \frac{1}{(s-\lambda)^2} \Rightarrow X_{n-1}(s) = \frac{X_n(s)}{s-\lambda} \\ \vdots \\ \frac{X_2(s)}{U(s)} = \frac{1}{(s-\lambda)^{n-1}} \Rightarrow X_2(s) = \frac{X_3(s)}{s-\lambda} \\ \frac{X_1(s)}{U(s)} = \frac{1}{(s-\lambda)^n} \Rightarrow X_1(s) = \frac{X_2(s)}{s-\lambda} \end{array} \right. \tag{2.228}$$

Coming back in the time domain and reversing the order of the equations we get:

$$\left\{ \begin{array}{l} \dot{x}_1(t) = \lambda x_1(t) + x_2(t) \\ \dot{x}_2(t) = \lambda x_2(t) + x_3(t) \\ \vdots \\ \dot{x}_{n-1}(t) = \lambda x_{n-1}(t) + x_n(t) \\ \dot{x}_n(t) = \lambda x_n(t) + u(t) \end{array} \right. \tag{2.229}$$

Equations (2.229) and (2.227) lead to the following state-space representation, which is called the diagonal (Jordan) or modal form:

$$\left\{ \begin{array}{l} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}u(t) \\ y(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}u(t) \end{array} \right. \tag{2.230}$$

Matrix $\mathbf{A}$ is a $n \times n$ square matrix, $\mathbf{B}$ is a vector with n rows and $\mathbf{C}$ is a row with n columns:

$$\begin{cases} \mathbf{A} = \underbrace{\begin{bmatrix} \lambda & 1 & 0 & \cdots \\ 0 & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda \end{bmatrix}}_{n \quad terms} \\ \mathbf{B} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \\ \mathbf{C} = \begin{bmatrix} r_n & r_{n-1} & \cdots & r_1 \end{bmatrix} \\ \mathbf{D} = d \end{cases} \tag{2.231}$$

Alternatively we can introduce polynomials $N_1(s)$ and $N_2(s)$ defined as follows:

$$F(s) = \frac{r_1}{s-\lambda} + \frac{r_2}{(s-\lambda)^2} + \cdots + \frac{r_n}{(s-\lambda)^n} = \frac{N_1(s)N_2(s)}{(s-\lambda)^n} \tag{2.232}$$

Then Pradin[1] has shown that equivalent diagonal form realizations of transfer function $F(s)$ are the following where $\mathbf{A}$ is a $n \times n$ square matrix, $\mathbf{B}$ a vector with n rows and $\mathbf{C}$ a row with n columns:

$$\begin{cases} \mathbf{A} = \underbrace{\begin{bmatrix} \lambda & 1 & 0 & \cdots \\ 0 & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda \end{bmatrix}}_{n \quad terms} \\ \mathbf{B} = \begin{bmatrix} \vdots \\ \frac{1}{2!}\frac{d^2}{ds^2}N_2(s)\Big|_{s=\lambda} \\ \frac{1}{1!}\frac{d}{ds}N_2(s)\Big|_{s=\lambda} \\ N_2(s)\big|_{s=\lambda} \end{bmatrix} \\ \mathbf{C} = \begin{bmatrix} N_1(s)\big|_{s=\lambda} & \frac{1}{1!}\frac{d}{ds}N_1(s)\Big|_{s=\lambda} & \frac{1}{2!}\frac{d^2}{ds^2}N_1(s)\Big|_{s=\lambda} & \cdots \end{bmatrix} \\ \mathbf{D} = d \end{cases} \tag{2.233}$$

[1] Bernard Pradin, Automatique Linéaire - Systémes multivariables, Notes de cours INSA 2000

The preceding relations can be extended to the general case where transfer function $F(s)$ has poles λ_i with multiplicity n_i:

$$\begin{aligned} F(s) &= \sum_i \frac{r_{i1}}{s-\lambda_i} + \frac{r_{i2}}{(s-\lambda_i)^2} + \cdots + \frac{r_{in_i}}{(s-\lambda_i)^{n_i}} + d \\ &= \sum_i \sum_{j=1}^{n_i} \frac{r_{ij}}{(s-\lambda_i)^j} + d \\ &= \sum_i \frac{N_{i1}(s)N_{i2}(s)}{(s-\lambda_i)^{n_i}} + d \end{aligned} \tag{2.234}$$

Then it is shown in [1] that a diagonal form realization of transfer function $F(s)$ is the following:

$$\begin{cases} \mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & & \\ & \mathbf{A}_2 & \\ & & \ddots \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \vdots \end{bmatrix} \\ \mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 & \cdots \end{bmatrix} \\ \mathbf{D} = d \end{cases} \tag{2.235}$$

Matrix $\mathbf{A}_i$ is a $n_i \times n_i$ square matrix, $\mathbf{B}_i$ is a vector with n_i rows and $\mathbf{C}_i$ is a row with n_i columns:

$$\begin{cases} \mathbf{A}_i = \underbrace{\begin{bmatrix} \lambda_i & 1 & 0 & \cdots \\ 0 & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda_i \end{bmatrix}}_{n_i \quad terms} \\ \mathbf{B}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \\ \mathbf{C}_i = \begin{bmatrix} r_{in_i} & \cdots & r_{i2} & r_{i1} \end{bmatrix} \\ \mathbf{D} = d \end{cases} \tag{2.236}$$

or equivalently:

$$\begin{cases} \mathbf{A}_i = \underbrace{\begin{bmatrix} \lambda_i & 1 & 0 & \cdots \\ 0 & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda_i \end{bmatrix}}_{n_i \quad terms} \\ \mathbf{B}_i = \begin{bmatrix} \vdots \\ \frac{1}{2!}\frac{d^2}{ds^2}N_{i2}(s)\Big|_{s=\lambda_i} \\ \frac{1}{1!}\frac{d}{ds}N_{i2}(s)\Big|_{s=\lambda_i} \\ N_{i2}(s)|_{s=\lambda_i} \end{bmatrix} \\ \mathbf{C}_i = \begin{bmatrix} N_{i1}(s)|_{s=\lambda_i} & \frac{1}{1!}\frac{d}{ds}N_{i1}(s)\Big|_{s=\lambda_i} & \frac{1}{2!}\frac{d^2}{ds^2}N_{i1}(s)\Big|_{s=\lambda_i} & \cdots \end{bmatrix} \\ \mathbf{D} = d \end{cases} \tag{2.237}$$

Multiple complex conjugate pair of poles on the same location

If some of the poles are complex so are the residues and so is the Jordan form. This may be inconvenient. Assume that λ and $\overline{\lambda}$ is a complex conjugate pair of poles of $F(s)$ with multiplicity 3:

$$F(s) = \frac{r_1}{s-\lambda} + \frac{r_2}{(s-\lambda)^2} + \frac{r_3}{(s-\lambda)^3} + \frac{\overline{r}_1}{s-\overline{\lambda}} + \frac{\overline{r}_2}{(s-\overline{\lambda})^2} + \frac{\overline{r}_3}{(s-\overline{\lambda})^3} \tag{2.238}$$

Let α be the real part of the pole λ and β its imaginary part:

$$\lambda = \alpha + j\beta \Leftrightarrow \overline{\lambda} = \alpha - j\beta \tag{2.239}$$

Using the result of the preceding section the Jordan form of transfer function

$F(s)$ is the following:

$$\begin{cases} \mathbf{A} = \left[\begin{array}{ccc|ccc} \lambda & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \overline{\lambda} & 1 & 0 \\ 0 & 0 & 0 & 0 & \overline{\lambda} & 1 \\ 0 & 0 & 0 & 0 & 0 & \overline{\lambda} \end{array}\right] \\ \mathbf{B} = \left[\begin{array}{c} 0 \\ 0 \\ 1 \\ \hline 0 \\ 0 \\ 1 \end{array}\right] \\ \mathbf{C} = \left[\begin{array}{ccc|ccc} r_3 & r_2 & r_1 & \overline{r}_3 & \overline{r}_2 & \overline{r}_1 \end{array}\right] \\ \mathbf{D} = 0 \end{cases} \tag{2.240}$$

It is clear that the Jordan form of transfer function $F(s)$ is complex. This complex Jordan form is rendered real by using the real part and the imaginary part of the complex state components which appear in the state vector rather than the complex state components and its conjugate. This is the same kind of trick which has been used in the section dealing with complex conjugate pair of poles. The real state matrix $\mathbf{A}_n$ is the following:

$$\mathbf{A}_n = \begin{bmatrix} \mathbf{J}_{ab} & \mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{ab} & \mathbb{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{J}_{ab} \end{bmatrix} \text{ where } \mathbf{J}_{ab} = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \tag{2.241}$$

It can be seen that complex matrix $\mathbf{A}$ has the same determinant than the following real matrix $\mathbf{A}_n$:

$$\det(s\mathbb{I} - \mathbf{A}) = \det(s\mathbb{I} - \mathbf{A}_n) \tag{2.242}$$

2.4 Realization of SIMO transfer function

The acronym SIMO stands for Single-Input Multiple-Output. The transfer function $\mathbf{F}(s)$ relates the relation between the Laplace transform of the output of the system, $\underline{y}(t)$, which is a vector, and the Laplace transform of the input of the system, $u(t)$, which is a scalar as in the SISO case. Thus in that situation the transfer function becomes a vector. Let $\underline{Y}(s) = \mathcal{L}\left[\underline{y}(t)\right]$ and $U(s) = \mathcal{L}\left[u(t)\right]$. Thus we write:

$$\underline{Y}(s) = \mathbf{F}(s)U(s) \tag{2.243}$$

As in the SISO case, the realization of a SIMO transfer function $\mathbf{F}(s)$ consists in finding any quadruplet $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ such that:

$$\mathbf{F}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} \tag{2.244}$$

We will consider in the following a SIMO system with p outputs. Thus $\underline{Y}(s)$ is a vector of p rows and $U(s)$ a scalar. Several kind of realizations are possible which will be presented hereafter.

2.4.1 Generic procedure

In the SIMO case we can always write the transfer function $\mathbf{F}(s)$ as a vector composed of p transfer functions of SISO systems:

$$\mathbf{F}(s) = \begin{bmatrix} F_1(s) \\ \vdots \\ F_p(s) \end{bmatrix} \tag{2.245}$$

If we realize $F_i(s)$ by $\left(\begin{array}{c|c} \mathbf{A}_i & \mathbf{B}_i \\ \hline \mathbf{C}_i & d_i \end{array}\right)$ then one realization of $\mathbf{F}(s)$ is the following:

$$F_i(s) = \left(\begin{array}{c|c} \mathbf{A}_i & \mathbf{B}_i \\ \hline \mathbf{C}_i & d_i \end{array}\right) \Rightarrow \mathbf{F}(s) = \left(\begin{array}{ccc|c} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{B}_1 \\ \mathbf{0} & \ddots & & \vdots \\ \vdots & & \mathbf{A}_p & \mathbf{B}_p \\ \hline \mathbf{C}_1 & \mathbf{0} & \cdots & d_1 \\ \mathbf{0} & \ddots & & \vdots \\ \vdots & & \mathbf{C}_p & d_p \end{array}\right) \tag{2.246}$$

To get the previous result we have to write $\mathbf{F}(s)$ as follows:

$$\begin{aligned}
\mathbf{F}(s) &= \begin{bmatrix} F_1(s) \\ \vdots \\ F_p(s) \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 \left(s\mathbb{I} - \mathbf{A}_1\right)^{-1} \mathbf{B}_1 + d_1 \\ \vdots \\ \mathbf{C}_p \left(s\mathbb{I} - \mathbf{A}_p\right)^{-1} \mathbf{B}_p + d_p \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{C}_1 & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & \\ \vdots & & \mathbf{C}_p \end{bmatrix} \begin{bmatrix} \left(s\mathbb{I} - \mathbf{A}_1\right)^{-1} \mathbf{B}_1 \\ \vdots \\ \left(s\mathbb{I} - \mathbf{A}_p\right)^{-1} \mathbf{B}_p \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_p \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{C}_1 & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & \\ \vdots & & \mathbf{C}_p \end{bmatrix} \begin{bmatrix} \left(s\mathbb{I} - \mathbf{A}_1\right)^{-1} & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & \\ \vdots & & \left(s\mathbb{I} - \mathbf{A}_p\right)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{B}_1 \\ \vdots \\ \mathbf{B}_p \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_p \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{C}_1 & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & \\ \vdots & & \mathbf{C}_p \end{bmatrix} \left(\begin{bmatrix} \left(s\mathbb{I} - \mathbf{A}_1\right) & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & \\ \vdots & & \left(s\mathbb{I} - \mathbf{A}_p\right) \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{B}_1 \\ \vdots \\ \mathbf{B}_p \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_p \end{bmatrix}
\end{aligned} \tag{2.247}$$

From the preceding relation we deduce the realization (2.246).

Example 2.10. *Let's consider the following SIMO transfer function:*

$$\mathbf{F}(s) = \begin{bmatrix} \frac{s+1}{s^2+6s+9} \\ \frac{5}{s^2+6s+9} \end{bmatrix} \tag{2.248}$$

$\mathbf{F}(s)$ is the transfer function of a system with $p = 2$ outputs, which is the number of rows of $\mathbf{F}(s)$, and one input, which is the number of columns of $\mathbf{F}(s)$.

We notice that $\mathbf{F}(s)$ is a strictly proper. Consequently:

$$\underline{d} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2.249}$$

Then we write the transfer function $\mathbf{F}_{sp}(s) := \mathbf{F}(s)$ as the ratio between a polynomial vector $\mathbf{N}(s)$ with $p = 2$ rows and a polynomial $\Psi(s)$:

$$\mathbf{F}(s) := \mathbf{F}_{sp}(s) = \frac{\mathbf{N}(s)}{\Psi(s)} = \frac{\begin{bmatrix} s+1 \\ 5 \end{bmatrix}}{s^2 + 6s + 9} \tag{2.250}$$

A realization of transfer function $F_1(s) = \frac{s+1}{s^2+6s+9}$ is for example the controllable canonical form:

$$\mathbf{F}_1(s) = \frac{s+1}{s^2+6s+9} = \left(\begin{array}{cc|c} 0 & 1 & 0 \\ -9 & -6 & 1 \\ \hline 1 & 1 & 0 \end{array}\right) \tag{2.251}$$

Similarly a realization of transfer function $F_2(s) = \frac{5}{s^2+6s+9}$ is for example the controllable canonical form:

$$\mathbf{F}_2(s) = \frac{5}{s^2+6s+9} = \left(\begin{array}{cc|c} 0 & 1 & 0 \\ -9 & -6 & 1 \\ \hline 5 & 0 & 0 \end{array}\right) \tag{2.252}$$

Applying the generic procedure we get a realization of the SIMO transfer function $\mathbf{F}(s)$:

$$\mathbf{F}(s) = \left(\begin{array}{cccc|c} 0 & 1 & 0 & 0 & 0 \\ -9 & -6 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -9 & -6 & 1 \\ \hline 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \end{array}\right) \tag{2.253}$$

■

2.4.2 Controllable canonical form

We can also write the transfer function $\mathbf{F}(s)$ as follows:

$$\mathbf{F}(s) = \mathbf{F}_{sp}(s) + \begin{bmatrix} d_1 \\ \vdots \\ d_p \end{bmatrix} = \mathbf{F}_{sp}(s) + \underline{d} \tag{2.254}$$

where $\underline{d}$ is a constant vector and $\mathbf{F}_{sp}(s)$ a strictly proper transfer function:

$$\begin{cases} \lim_{s\to\infty} \mathbf{F}(s) = \underline{d} \\ \lim_{s\to\infty} \mathbf{F}_{sp}(s) = \mathbf{0} \end{cases} \tag{2.255}$$

With the same argument than in the SISO case we have:

$$\mathbf{D} = \underline{d} = \begin{bmatrix} d_1 \\ \vdots \\ d_p \end{bmatrix} \tag{2.256}$$

Then we have to find matrices $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ such that:

$$\mathbf{F}_{sp}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} \tag{2.257}$$

To get the controllable canonical form we write the transfer function $\mathbf{F}_{sp}(s)$ as the ratio between a polynomial vector $\mathbf{N}(s)$ with p rows and a polynomial $\Psi(s)$:

$$\mathbf{F}_{sp}(s) = \frac{\mathbf{N}(s)}{\Psi(s)} = \frac{\begin{bmatrix} N_1(s) \\ \vdots \\ N_p(s) \end{bmatrix}}{\Psi(s)} \tag{2.258}$$

Then we build for each SISO transfer function $N_i(s)/\Psi(s)$ a controllable realization $(\mathbf{A}_c, \mathbf{B}_c, \mathbf{C}_i, \mathbf{0})$. Note that:

- Matrix $\mathbf{A}_c$ is common to each realization because the denominator $\Psi(s)$ of each transfer function $N_i(s)/\Psi(s)$ is the same. When we write $\Psi(s)$ as follows:

$$\Psi(s) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \tag{2.259}$$

  Then $\mathbf{A}_c$ is a $n \times n$ square matrix:

$$\mathbf{A}_c = \begin{bmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ & & \ddots & \ddots & 0 \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \tag{2.260}$$

- Vector $\mathbf{B}_c$ is common to each realization because we use the controllable canonical form of each SISO transfer function $N_i(s)/\Psi(s)$. This is a vector vector with n rows:

$$\mathbf{B}_c = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{2.261}$$

- Each vector $\mathbf{C}_i$ is dedicated to one output. This is a row vector with p columns formed with the coefficients of polynomials $N_i(s)$.

Then the controllable canonical form of the SIMO transfer function $\mathbf{F}(s)$ is the following:

$$\mathbf{F}(s) = \left(\begin{array}{c|c} \mathbf{A}_c & \mathbf{B}_c \\ \hline \mathbf{C}_1 & d_1 \\ \vdots & \vdots \\ \mathbf{C}_p & d_p \end{array}\right) \tag{2.262}$$

Example 2.11. *Let's consider the following SIMO transfer function:*

$$\mathbf{F}(s) = \begin{bmatrix} \frac{1}{s+1} \\ \frac{2}{s+2} \end{bmatrix} \tag{2.263}$$

$\mathbf{F}(s)$ *is the transfer function of a system with* $p = 2$ *outputs, which is the number of rows of* $\mathbf{F}(s)$*, and one input, which is the number of columns of* $\mathbf{F}(s)$*.*

We notice that $\mathbf{F}(s)$ *is a strictly proper. Consequently:*

$$\underline{d} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2.264}$$

Then we write the transfer function $\mathbf{F}_{sp}(s) := \mathbf{F}(s)$ *as the ratio between a polynomial vector* $\mathbf{N}(s)$ *with* $p = 2$ *rows and a polynomial* $\Psi(s)$*:*

$$\mathbf{F}(s) := \mathbf{F}_{sp}(s) = \frac{\mathbf{N}(s)}{\Psi(s)} = \frac{\begin{bmatrix} s+2 \\ 2(s+1) \end{bmatrix}}{(s+1)(s+2)} = \frac{\begin{bmatrix} s+2 \\ 2(s+1) \end{bmatrix}}{s^2+3s+2} \tag{2.265}$$

Then matrix $\mathbf{A}_c$ *of the controllable canonical form of* $\mathbf{F}(s)$ *is obtained thanks to the coefficients of the denominator* $\Psi(s)$ *whereas vector* $\mathbf{B}_c$ *is set by the controllable canonical form:*

$$\begin{cases} \mathbf{A}_c = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \\ \mathbf{B}_c = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{cases} \tag{2.266}$$

Vector $\mathbf{C}_c$ *is obtained thanks to the coefficients of the polynomial vector* $\mathbf{N}(s)$ *:*

$$\mathbf{C}_c = \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix} \tag{2.267}$$

We finally get:

$$\mathbf{F}(s) = \left(\begin{array}{cc|c} 0 & 1 & 0 \\ -2 & -3 & 1 \\ \hline 2 & 1 & 0 \\ 2 & 2 & 0 \end{array}\right) \tag{2.268}$$

■

Example 2.12. *Let's consider the following SIMO transfer function:*

$$\mathbf{F}(s) = \begin{bmatrix} \frac{s+1}{s^2+6s+9} \\ \frac{5}{s^2+6s+9} \end{bmatrix} \tag{2.269}$$

$\mathbf{F}(s)$ *is the transfer function of a system with* $p = 2$ *outputs, which is the number of rows of* $\mathbf{F}(s)$*, and one input, which is the number of columns of* $\mathbf{F}(s)$*.*

We notice that $\mathbf{F}(s)$ *is a strictly proper. Consequently:*

$$\underline{d} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2.270}$$

Then we write the transfer function $\mathbf{F}_{sp}(s) := \mathbf{F}(s)$ *as the ratio between a polynomial vector* $\mathbf{N}(s)$ *with* $p = 2$ *rows and a polynomial* $\Psi(s)$*:*

$$\mathbf{F}(s) := \mathbf{F}_{sp}(s) = \frac{\mathbf{N}(s)}{\Psi(s)} = \frac{\begin{bmatrix} s+1 \\ 5 \end{bmatrix}}{s^2 + 6s + 9} \tag{2.271}$$

Then matrix $\mathbf{A}_c$ *of the controllable canonical form of* $\mathbf{F}(s)$ *is obtained thanks to the coefficients of the denominator* $\Psi(s)$ *whereas vector* $\mathbf{B}_c$ *is set by the controllable canonical form:*

$$\begin{cases} \mathbf{A}_c = \begin{bmatrix} 0 & 1 \\ -9 & -6 \end{bmatrix} \\ \mathbf{B}_c = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{cases} \tag{2.272}$$

Vector $\mathbf{C}_c$ *is obtained thanks to the coefficients of the polynomial vector* $\mathbf{N}(s)$ *:*

$$\mathbf{C}_c = \begin{bmatrix} 1 & 1 \\ 5 & 0 \end{bmatrix} \tag{2.273}$$

We finally get:

$$\mathbf{F}(s) = \left(\begin{array}{cc|c} 0 & 1 & 0 \\ -9 & -6 & 1 \\ \hline 1 & 1 & 0 \\ 5 & 0 & 0 \end{array}\right) \tag{2.274}$$

■

2.5 Realization of MIMO transfer function

The acronym MIMO stands for Multi-Input Multiple-Output.

The transfer function $\mathbf{F}(s)$ relates the relation between the Laplace transform of the output of the system, which is a vector, and the Laplace transform of the input of the system, which is also a vector in the MIMO case. Due to the fact that the output $\underline{y}(t)$ of the system and the input $\underline{u}(t)$ of the system are no more scalars but vectors it is not possible to express the ratio between $\underline{Y}(s) = \mathcal{L}\left[\underline{y}(t)\right]$ and $\underline{U}(s) = \mathcal{L}\left[\underline{u}(t)\right]$. Thus we write:

$$\underline{Y}(s) = \mathbf{F}(s)\underline{U}(s) \tag{2.275}$$

We will consider in the following a MIMO system with p outputs and m inputs. Then $\underline{Y}(s) = \mathcal{L}\left[\underline{y}(t)\right]$ is a vector of p rows, $\underline{U}(s) = \mathcal{L}\left[\underline{u}(t)\right]$ is a vector of m rows and transfer function $\mathbf{F}(s)$ is a matrix with m columns and p rows.

As in the SIMO case, the realization of a MIMO transfer function $\mathbf{F}(s)$ consists in finding any quadruplet $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ such that:

$$\mathbf{F}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} \tag{2.276}$$

2.5.1 Generic procedure

In the MIMO case we can always write the transfer function $\mathbf{F}(s)$ as a matrix composed of $p \times m$ transfer functions of SISO systems $F_{ij}(s)$:

$$\mathbf{F}(s) = \begin{bmatrix} F_{11}(s) & \cdots & F_{1m}(s) \\ \vdots & & \vdots \\ F_{p1}(s) & \cdots & F_{pm}(s) \end{bmatrix} \tag{2.277}$$

The transfer function $\mathbf{F}(s)$ can be written as the sum of SIMO systems:

$$\mathbf{F}(s) = \begin{bmatrix} F_{11}(s) \\ \vdots \\ F_{p1}(s) \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} + \cdots + \begin{bmatrix} F_{1m}(s) \\ \vdots \\ F_{pm}(s) \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix} \tag{2.278}$$

That is:

$$\begin{aligned} \mathbf{F}(s) &= \mathbf{F}_1(s) \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} + \cdots + \mathbf{F}_m(s) \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix} \\ &= \textstyle\sum_{i=1}^{m} \mathbf{F}_i(s) \begin{bmatrix} 0 & \ldots & 0 & \underbrace{1}_{\text{i-th column}} & 0 & \ldots & 0 \end{bmatrix} \end{aligned} \tag{2.279}$$

If we realize the SIMO system $\mathbf{F}_i(s) = \begin{bmatrix} F_{1i}(s) \\ \vdots \\ F_{pi}(s) \end{bmatrix}$ in the ith column of $\mathbf{F}(s)$ by $\left(\begin{array}{c|c} \mathbf{A}_i & \mathbf{B}_i \\ \hline \mathbf{C}_i & \mathbf{D}_i \end{array}\right)$ then one realization of transfer function $\mathbf{F}(s)$ is the following:

$$\mathbf{F}_i(s) = \begin{bmatrix} F_{1i}(s) \\ \vdots \\ F_{pi}(s) \end{bmatrix} = \left(\begin{array}{c|c} \mathbf{A}_i & \mathbf{B}_i \\ \hline \mathbf{C}_i & \mathbf{D}_i \end{array}\right)$$

$$\Rightarrow \mathbf{F}(s) = \left(\begin{array}{ccc|ccc} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{B}_1 & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & & \mathbf{0} & \ddots & \\ \vdots & & \mathbf{A}_m & \vdots & & \mathbf{B}_m \\ \hline \mathbf{C}_1 & \cdots & \mathbf{C}_m & \mathbf{D}_1 & \cdots & \mathbf{D}_m \end{array}\right) \tag{2.280}$$

The state-space representation of each SIMO transfer function $\mathbf{F}_i(s)$ can be obtained thanks to the controllable canonical form (2.262). The achieved state-space representation is block diagonal but is not necessarily *minimal* (see section 2.6).

To get this result we use the same kind of demonstration than the one which has been to obtain the generic procedure seen in Equation (2.246). Indeed:

$$
\begin{aligned}
\mathbf{F}(s) &= \mathbf{F}_1(s)\begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} + \cdots + \mathbf{F}_m(s)\begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{F}_1(s) & \cdots & \mathbf{F}_m(s) \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{C}_1\left(s\mathbb{I}-\mathbf{A}_1\right)^{-1}\mathbf{B}_1+\mathbf{D}_1 & \cdots & \mathbf{C}_m\left(s\mathbb{I}-\mathbf{A}_m\right)^{-1}\mathbf{B}_m+\mathbf{D}_m \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{C}_1 & \cdots & \mathbf{C}_m \end{bmatrix}\begin{bmatrix} \left(s\mathbb{I}-\mathbf{A}_1\right)^{-1}\mathbf{B}_1 & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & \\ \vdots & & \left(s\mathbb{I}-\mathbf{A}_m\right)^{-1}\mathbf{B}_m \end{bmatrix} + \begin{bmatrix} \mathbf{D}_1 & \cdots & \mathbf{D}_m \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{C}_1 & \cdots & \mathbf{C}_m \end{bmatrix}\begin{bmatrix} \left(s\mathbb{I}-\mathbf{A}_1\right)^{-1} & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & \\ \vdots & & \left(s\mathbb{I}-\mathbf{A}_m\right)^{-1} \end{bmatrix}\begin{bmatrix} \mathbf{B}_1 & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & \\ \vdots & & \mathbf{B}_m \end{bmatrix} \\
&\qquad + \begin{bmatrix} \mathbf{D}_1 & \cdots & \mathbf{D}_m \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{C}_1 & \cdots & \mathbf{C}_m \end{bmatrix}\left(\begin{bmatrix} \left(s\mathbb{I}-\mathbf{A}_1\right) & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & \\ \vdots & & \left(s\mathbb{I}-\mathbf{A}_m\right) \end{bmatrix}\right)^{-1}\begin{bmatrix} \mathbf{B}_1 & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & \\ \vdots & & \mathbf{B}_m \end{bmatrix} \\
&\qquad + \begin{bmatrix} \mathbf{D}_1 & \cdots & \mathbf{D}_m \end{bmatrix}
\end{aligned}
\tag{2.281}
$$

2.5.2 Controllable canonical form

In the MIMO case, transfer function $\mathbf{F}(s)$ can always be expanded as follows where $p \times m$ constant matrices $\mathbf{C}_i,\ i = 1, \cdots, n-1$ and $\mathbf{D}$ are of the same size than transfer function $\mathbf{F}(s)$ with m inputs and p outputs:

$$
\mathbf{F}(s) = \frac{\mathbf{N}(s)}{\Psi(s)} + \mathbf{D} = \frac{\mathbf{C}_{n-1}s^{n-1} + \cdots + \mathbf{C}_1 s + \mathbf{C}_0}{1 \times s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0} + \mathbf{D} \tag{2.282}
$$

Following the same procedure than in the SISO case, and by denoting $\mathbb{I}_m$ the identity matrix of dimension m, the controllable canonical form of $\mathbf{F}(s)$ reads:

$$
\mathbf{F}(s) = \left(\begin{array}{c|c} \mathbf{A}_c & \mathbf{B}_c \\ \hline \mathbf{C}_c & \mathbf{D} \end{array}\right) \tag{2.283}
$$

where:

$$\begin{cases} \mathbf{A}_c = \begin{bmatrix} \mathbf{0} & \mathbb{I}_m & \mathbf{0} & & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbb{I}_m & \ddots & \mathbf{0} \\ & & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & & \mathbb{I}_m \\ -a_0\mathbb{I}_m & -a_1\mathbb{I}_m & -a_2\mathbb{I}_m & \cdots & -a_{n-1}\mathbb{I}_m \end{bmatrix} \\ \mathbf{B}_c = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbb{I}_m \end{bmatrix} \\ \mathbf{C}_c = \begin{bmatrix} \mathbf{C}_0 & \mathbf{C}_1 & \cdots & \mathbf{C}_{n-2} & \mathbf{C}_{n-1} \end{bmatrix} \\ \mathbf{D} = \lim_{s\to\infty} \mathbf{F}(s) \end{cases} \tag{2.284}$$

Example 2.13. *Let's consider the following transfer function:*

$$\mathbf{F}(s) = \begin{bmatrix} \frac{2}{s+2} & \frac{s+1}{s+3} \\ \frac{1}{s+2} & \frac{5}{s+2} \end{bmatrix} \tag{2.285}$$

Let's decompose $\mathbf{F}(s)$ *as follows:*

$$\begin{aligned} \mathbf{F}(s) &= \begin{bmatrix} \frac{2}{s+2} & \frac{-2}{s+3} \\ \frac{1}{s+2} & \frac{5}{s+2} \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\ &= \frac{\begin{bmatrix} 2(s+3) & -2(s+2) \\ s+3 & 5(s+3) \end{bmatrix}}{(s+2)(s+3)} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\ &= \frac{\begin{bmatrix} 2 & -2 \\ 1 & 5 \end{bmatrix} s + \begin{bmatrix} 6 & -4 \\ 3 & 15 \end{bmatrix}}{s^2+5s+6} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \end{aligned} \tag{2.286}$$

The system described by transfer function $\mathbf{F}(s)$ *has* $m = 2$ *inputs. Using* (2.284) *leads to the following controllable canonical realization of* $\mathbf{F}(s)$*:*

$$\left[\begin{array}{c} \dot{\underline{x}}(t) \\ \hline y(t) \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A}_c & \mathbf{B}_c \\ \hline \mathbf{C}_c & \mathbf{D} \end{array}\right] \left[\begin{array}{c} \underline{x}(t) \\ \hline \underline{u}(t) \end{array}\right] \tag{2.287}$$

where:

$$\begin{cases} \mathbf{A}_c = \left[\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline -6 & 0 & -5 & 0 \\ 0 & -6 & 0 & -5 \end{array}\right] \\ \mathbf{B}_c = \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \\ \hline 1 & 0 \\ 0 & 1 \end{array}\right] \\ \mathbf{C}_c = \left[\begin{array}{cc|cc} 6 & -4 & 2 & -2 \\ 3 & 15 & 1 & 5 \end{array}\right] \\ \mathbf{D} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \end{cases} \tag{2.288}$$

This result can be checked by using Equation (2.1).

■

2.5.3 Observable canonical form

The observable canonical form $\left(\begin{array}{c|c} \mathbf{A}_o & \mathbf{B}_o \\ \hline \mathbf{C}_o & \mathbf{D}_o \end{array}\right)$ of $\mathbf{F}(s)$ can be obtained by computing the transpose of the controllable canonical form of its transpose, that is the transpose of the controllable canonical form of $\mathbf{F}^T(s)$:

$$\begin{aligned} &\mathbf{F}^T(s) = \mathbf{C}_c\left(s\mathbb{I} - \mathbf{A}_c\right)^{-1}\mathbf{B}_c + \mathbf{D} := \left(\begin{array}{c|c} \mathbf{A}_c & \mathbf{B}_c \\ \hline \mathbf{C}_c & \mathbf{D} \end{array}\right) \\ &\Rightarrow \mathbf{F}(s) = \mathbf{B}_c^T\left(s\mathbb{I} - \mathbf{A}_c^T\right)^{-1}\mathbf{C}_c^T + \mathbf{D}^T = \left(\begin{array}{c|c} \mathbf{A}_c^T & \mathbf{C}_c^T \\ \hline \mathbf{B}_c^T & \mathbf{D}^T \end{array}\right) := \left(\begin{array}{c|c} \mathbf{A}_o & \mathbf{B}_o \\ \hline \mathbf{C}_o & \mathbf{D}_o \end{array}\right) \end{aligned} \tag{2.289}$$

2.5.4 Diagonal (or modal) form

As in the SIMO case we expand $\mathbf{F}(s)$ as follows:

$$\mathbf{F}(s) = \mathbf{F}_{sp}(s) + \mathbf{D} \tag{2.290}$$

where $\mathbf{D}$ is a constant matrix and $\mathbf{F}_{sp}(s)$ a strictly proper transfer function:

$$\begin{cases} \lim_{s\to\infty}\mathbf{F}(s) = \mathbf{D} \\ \lim_{s\to\infty}\mathbf{F}_{sp}(s) = \mathbf{0} \end{cases} \tag{2.291}$$

Then we have to find matrices $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ such that:

$$\mathbf{F}_{sp}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} \tag{2.292}$$

To get the diagonal (or modal) form we write the transfer function $\mathbf{F}_{sp}(s)$ as the sum between rational functions. Let $\lambda_1, \cdots, \lambda_r$ be the r distinct roots

of $\Psi(s)$ and n_i the multiplicity of root λ_i. Then we get the following partial fraction expansion of $\mathbf{F}_{sp}(s)$ where matrices $\mathbf{R}_{ij}$ are constant:

$$\mathbf{F}_{sp}(s) = \sum_i \sum_{j=1}^{n_i} \frac{\mathbf{R}_{ij}}{(s-\lambda_i)^j} \tag{2.293}$$

The diagonal (or modal) form of the MIMO transfer function $\mathbf{F}(s)$ is the following:

$$\mathbf{F}(s) = \left(\begin{array}{ccc|c} \mathbf{J}_1 & \mathbf{0} & \cdots & \mathbf{B}_1 \\ \mathbf{0} & \ddots & & \vdots \\ \vdots & & \mathbf{J}_r & \mathbf{B}_r \\ \hline \mathbf{C}_1 & \cdots & \mathbf{C}_r & \mathbf{D} \end{array}\right) \tag{2.294}$$

Denoting by n_i the multiplicity of the root λ_i, m the number of inputs of the system and $\mathbb{I}_m$ the identity matrix of size $m \times m$, matrices $\mathbf{J}_i$, $\mathbf{B}_i$ and $\mathbf{C}_i$ are defined as follows:

- The Jordan matrix $\mathbf{J}_i$ is a $(m \times n_i) \times (m \times n_i)$ matrix with the following expression:

$$\mathbf{J}_i = \underbrace{\begin{bmatrix} \lambda_i \mathbb{I}_m & \mathbb{I}_m & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \mathbb{I}_m \\ \mathbf{0} & \cdots & \mathbf{0} & \lambda_i \mathbb{I}_m \end{bmatrix}}_{n_i \ termes \ \lambda_i \mathbb{I}_m} \tag{2.295}$$

  It is worth noticing that matrix $(s\mathbb{I} - \mathbf{J}_i)^{-1}$ reads:

$$(s\mathbb{I} - \mathbf{J}_i)^{-1} = \begin{bmatrix} (s-\lambda_i)^{-1}\mathbb{I}_m & (s-\lambda_i)^{-2}\mathbb{I}_m & \cdots & (s-\lambda_i)^{-n_i}\mathbb{I}_m \\ \mathbf{0} & (s-\lambda_i)^{-1}\mathbb{I}_m & \ddots & (s-\lambda_i)^{-n_i+1}\mathbb{I}_m \\ \vdots & \ddots & \ddots & \ddots \\ \mathbf{0} & \cdots & \mathbf{0} & (s-\lambda_i)^{-1}\mathbb{I}_m \end{bmatrix} \tag{2.296}$$

- $\mathbf{B}_i$ is a $(m \times n_i) \times m$ matrix:

$$\mathbf{B}_i = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbb{I}_m \end{bmatrix} \tag{2.297}$$

- $\mathbf{C}_i$ is a $p \times (m \times n_i)$ matrix:

$$\mathbf{C}_i = \begin{bmatrix} \mathbf{R}_{in_i} & \cdots & \mathbf{R}_{i2} & \mathbf{R}_{i1} \end{bmatrix} \tag{2.298}$$

An alternative diagonal (or modal) form also exists. To get it first let's focus on the realization of the following $p \times m$ transfer function $\mathbf{F}_i(s)$ with a pole λ of multiplicity i:

$$\mathbf{F}_i(s) = \frac{\begin{bmatrix} \mathbb{I}_{\rho_i} & \mathbf{0}_{\rho_i \times (m-\rho_i)} \\ \mathbf{0}_{(p-\rho_i)\times \rho_i} & \mathbf{0}_{(p-\rho_i)\times(m-\rho_i)} \end{bmatrix}}{(s-\lambda)^i} \tag{2.299}$$

where $\mathbb{I}_{\rho_i}$ is the identity matrix of dimension ρ_i and $\mathbf{0}_{p\times m}$ the null matrix with p rows and m columns.

Then we recall the inverse of the following $n \times n$ bidiagonal matrix:

$$\mathbf{L} = \begin{bmatrix} \lambda & -1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & -1 \\ 0 & & & \lambda \end{bmatrix} \Rightarrow \mathbf{L}^{-1} = \begin{bmatrix} \lambda^{-1} & \lambda^{-2} & \cdots & \lambda^{-n} \\ & \lambda^{-1} & \ddots & \\ & & \ddots & \lambda^{-2} \\ 0 & & & \lambda^{-1} \end{bmatrix} \tag{2.300}$$

The alternative diagonal (or modal) form of $\mathbf{F}_i(s)$ is then the following [2]:

$$\mathbf{F}_i(s) = \left(\begin{array}{c|c} \mathbf{A}_i & \begin{bmatrix} \mathbf{B}_i & \mathbf{0}_{n\rho_i\times(m-\rho_i)} \end{bmatrix} \\ \hline \begin{bmatrix} \mathbf{C}_i \\ \mathbf{0}_{(p-\rho_i)\times n\rho_i} \end{bmatrix} & \mathbf{0}_{p\times m} \end{array}\right) \tag{2.301}$$

where $\mathbf{A}_i$ is a $(n \times \rho_i) \times (n \times \rho_i)$ square matrix, $\mathbf{B}_i^T$ (the transpose of $\mathbf{B}_i$) a $\rho_i \times (n \times \rho_i)$ matrix whose ρ_i rows are built from row vector $\underbrace{0 \quad 0 \quad \cdots \quad 1}_{n \quad terms}$ and

[2] Toshiya Morisue, Minimal Realization of a Transfer Function Matrix with Multiple Poles, Transactions of the Society of Instrument and Control Engineers, Volume 21 (1985) Issue 6 Pages 546-549

$\mathbf{C}_i$ a $\rho_i \times (n \times \rho_i)$ matrix whose ρ_i rows are built from row vector $\underbrace{1 \quad 0 \quad \cdots \quad 0}_{n \ \ terms}$:

$$\left\{\begin{array}{l}
\mathbf{A}_i = \underbrace{\begin{bmatrix} \mathbf{J}_i & \mathbf{0} & \cdots \\ \mathbf{0} & \ddots & \ddots \\ \vdots & \ddots & \mathbf{J}_i \end{bmatrix}}_{\rho_i \ \ terms} \quad \text{where} \quad \mathbf{J}_i = \underbrace{\begin{bmatrix} \lambda & 1 & 0 & \cdots \\ 0 & \ddots & \ddots & \ddots \\ & \ddots & & 1 \\ 0 & & 0 & \lambda \end{bmatrix}}_{n \ \ terms} \\
\mathbf{B}_i^T = \left.\begin{bmatrix} \underbrace{0 \quad 0 \quad \cdots \quad 1}_{n \ \ terms} & \mathbf{0}_{1\times n} & \cdots & \\ & & \ddots & \ddots \\ \mathbf{0}_{1\times n} & \cdots & & \underbrace{0 \quad 0 \quad \cdots \quad 1}_{n \ \ terms} \end{bmatrix}\right\} \rho_i \ \ terms \\
\mathbf{C}_i = \left.\begin{bmatrix} \underbrace{1 \quad 0 \quad \cdots \quad 0}_{n \ \ terms} & \mathbf{0}_{1\times n} & \cdots & \\ & & \ddots & \ddots \\ \mathbf{0}_{1\times n} & \cdots & & \underbrace{1 \quad 0 \quad \cdots \quad 0}_{n \ \ terms} \end{bmatrix}\right\} \rho_i \ \ terms
\end{array}\right. \tag{2.302}$$

Now let's consider the following transfer function $\mathbf{F}_i(s)$ where $\mathbf{N}_{i1}$ is a constant $p \times \rho_i$ matrix, $\mathbf{N}_{i2}$ a constant $\rho_i \times m$ matrix and $\mathbf{N}_{i1}\mathbf{N}_{i2}$ is a $p \times m$ constant matrix of rank ρ_i:

$$\mathbf{F}_i(s) = \frac{\mathbf{N}_{i1}\mathbf{N}_{i2}}{(s-\lambda)^i} = \frac{\mathbf{N}_{i1}\mathbb{I}_{\rho_i}\mathbf{N}_{i2}}{(s-\lambda)^i} \quad \text{where} \quad \operatorname{rank}(\mathbf{N}_{i1}\mathbf{N}_{i2}) = \rho_i \tag{2.303}$$

From the preceding realization it is clear that the alternative diagonal (or modal) form of $\mathbf{F}_i(s)$ is the following:

$$\mathbf{F}_i(s) = \left(\begin{array}{c|c} \mathbf{A}_i & \mathbf{B}_i\mathbf{N}_{i2} \\ \hline \mathbf{N}_{i1}\mathbf{C}_i & \mathbf{0}_{p\times m} \end{array}\right) \tag{2.304}$$

Finally let's consider a $p \times m$ transfer function $\mathbf{F}(s)$ which has pole λ with multiplicity n and where $\mathbf{R}(s)$ is a matrix of polynomial of degree strictly lower than n. The partial fraction expansion of $\mathbf{F}(s)$ reads:

$$\begin{aligned} \mathbf{F}(s) &= \tfrac{\mathbf{R}(s)}{(s-\lambda)^n} \\ &= \tfrac{\mathbf{R}_1}{s-\lambda} + \tfrac{\mathbf{R}_2}{(s-\lambda)^2} + \cdots + \tfrac{\mathbf{R}_n}{(s-\lambda)^n} \\ &= \textstyle\sum_{i=1}^{n} \tfrac{\mathbf{R}_i}{(s-\lambda)^i} \end{aligned} \tag{2.305}$$

Constant matrices $\mathbf{R}_i$ are defined by:

$$\mathbf{R}_i = \frac{1}{(n-i)!}\,\frac{d^{n-i}}{ds^{n-i}}\,(s-\lambda)^n\,\mathbf{F}(s)\bigg|_{s=\lambda} \tag{2.306}$$

Let ρ_i be the rank of *constant* matrix $\mathbf{R}_i$:

$$\rho_i = \text{rank}\,(\mathbf{R}_i) \tag{2.307}$$

Each term $\mathbf{R}_i$ can be expanded as a product of two *constant* matrices $\mathbf{N}_{i1}$ and $\mathbf{N}_{i2}$ where $\mathbf{N}_{i1}$ is a $p \times \rho_i$ matrix and $\mathbf{N}_{i2}$ a $\rho_i \times m$ matrix:

$$\mathbf{R}_i = \mathbf{N}_{i1}\mathbf{N}_{i2} = \mathbf{N}_{i1}\mathbb{I}_{\rho_i}\mathbf{N}_{i2} \tag{2.308}$$

Then the alternative diagonal (or modal) form of the MIMO transfer function $\mathbf{F}(s)$ is the following:

$$\mathbf{F}(s) = \sum_{i=1}^{n} \mathbf{F}_i(s) + \mathbf{D} = \left(\begin{array}{ccc|c} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{B}_1\mathbf{N}_{12} \\ \mathbf{0} & \ddots & & \vdots \\ \vdots & & \mathbf{A}_n & \mathbf{B}_n\mathbf{N}_{n2} \\ \hline \mathbf{N}_{11}\mathbf{C}_1 & \cdots & \mathbf{N}_{n1}\mathbf{C}_n & \mathbf{D} \end{array}\right) \tag{2.309}$$

This diagonal (or modal) form of $\mathbf{F}(s)$ is in general not *minimal* (see section 2.6).

2.6 Minimal realization

2.6.1 System's dimension

Let's start with an example and consider the following transfer functions:

$$\begin{cases} F_1(s) = \frac{1}{s+1} \\ F_2(s) = \frac{s+2}{s^2+3s+2} \end{cases} \tag{2.310}$$

From the preceding sections it can be seen that the controllable canonical form of transfer functions $F_1(s)$ and $F_2(s)$ are the following:

$$\begin{cases} F_1(s) = \left(\begin{array}{c|c} \mathbf{A}_1 & \mathbf{B}_1 \\ \hline \mathbf{C}_1 & \mathbf{D}_1 \end{array}\right) = \left(\begin{array}{c|c} -1 & 1 \\ \hline 1 & 0 \end{array}\right) \\ F_2(s) = \left(\begin{array}{c|c} \mathbf{A}_2 & \mathbf{B}_2 \\ \hline \mathbf{C}_2 & \mathbf{D}_2 \end{array}\right) = \left(\begin{array}{cc|c} 0 & 1 & 0 \\ -2 & -3 & 1 \\ \hline 2 & 1 & 0 \end{array}\right) \end{cases} \tag{2.311}$$

It is clear that the dimension of state matrix $\mathbf{A}_1$ is 1 and the dimension of state matrix $\mathbf{A}_2$ is 2.

On the other hand it can be seen that the poles of transfer function $F_2(s)$ are -1 and -2:

$$s^2 + 3s + 2 = (s+1)(s+2) \tag{2.312}$$

Consequently $F_2(s)$ reads:

$$F_2(s) = \frac{s+2}{s^2+3s+2} = \frac{s+2}{(s+1)(s+2)} = \frac{1}{s+1} \tag{2.313}$$

Thus we finally get:

$$F_2(s) = F_1(s) \tag{2.314}$$

Despite the fact that $F_2(s) = F_1(s)$ we have obtained two realizations with different size of the state matrix. This usually appears when pole-zero cancellation appears in the transfer function.

The order of a realization is the size of state matrix $\mathbf{A}$. So in that example the order of the realization of $F_2(s)$ is greater than the order of the realization of $F_1(s)$.

This example can be extended to the general case where the dimension of the state matrix $\mathbf{A}$ corresponding to the same transfer function $\mathbf{F}(s)$ may vary. We said that a realization of a transfer function $\mathbf{F}(s)$ is *minimal* if there exists no realization of lesser order whose transfer function is $\mathbf{F}(s)$.

For SISO systems it can be proven that a realization of transfer function $\mathbf{F}(s)$ is *minimal* if and only if the two polynomials $\mathbf{C}\,\mathrm{adj}(s\mathbb{I} - \mathbf{A})\mathbf{B}$ and $\det(s\mathbb{I} - \mathbf{A})$ are coprime.

For MIMO systems it can be proven that a realization of transfer function $\mathbf{F}(s)$ is *minimal* if and only if the characteristic polynomial of matrix $\mathbf{A}$ is equal to the Least Common Multiple (LCM), or Greatest Common Factor (GCF), of the denominators of all possible non zero minors (of all sizes) in $\mathbf{F}(s)$ [3].

We recall that minors or order k are the determinants of square sub-matrices of dimension k. More precisely if $\mathbf{F}(s)$ is a $p \times m$ matrix then the minors of order k are obtained by computing the determinant of all the square $k \times k$ sub-matrices where $p - k$ rows and $m - k$ columns of $\mathbf{F}(s)$ have been deleted.

To find the Least Common Multiple (LCM) of two polynomials simply factor each of the two polynomials completely. Then take the product of all factors (common and not common), every factor being affected with its greatest exponent. Finally multiply the obtained polynomial by a constant to obtain a monic polynomial.

Let $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$ be a *minimal* realization of a transfer function $\mathbf{F}(s)$. Then the eigenvalues of $\mathbf{A}$ are identical to the poles of $\mathbf{F}(s)$. If the realization is not *minimal* then the poles of $\mathbf{F}(s)$ are a subset of the eigenvalues of $\mathbf{A}$. It can be proven that Gilbert's diagonal realization is a *minimal* realization.

Example 2.14. *Let's consider the following transfer function:*

$$\mathbf{F}(s) = \begin{bmatrix} \frac{1}{s+1} & \frac{1}{s+2} \\ \frac{2}{s+1} & \frac{3}{s+1} \end{bmatrix} \tag{2.315}$$

A first realization of $\mathbf{F}(s)$ *is obtained by writing a realization of each SISO*

[3] Mohammed Dahleh, Munther A. Dahleh, George Verghese, Lectures on Dynamic Systems and Control, Massachuasetts Institute of Technology

transfer function:

$$\mathbf{F}(s) = \left(\begin{array}{cccc|cc} -1 & & & & 1 & 0 \\ & -1 & & & 2 & 0 \\ & & -2 & & 0 & 1 \\ & & & -1 & 0 & 3 \\ \hline 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{array}\right) \tag{2.316}$$

The characteristic polynomial of state matrix $\mathbf{A}$ *is:*

$$\det(s\mathbb{I} - \mathbf{A}) = (s+1)^3(s+2) \tag{2.317}$$

Whereas the Least Common Multiple (LCM) of the denominators of all possible non zero minors (of all sizes) in $\mathbf{F}(s)$ *is the following:*

$$\left\{\begin{array}{l} den\,(m_{11}(s)) = s+1 \\ den\,(m_{12}(s)) = s+2 \\ den\,(m_{21}(s)) = s+1 \\ den\,(m_{22}(s)) = s+1 \\ den\,(\mathbf{F}(s)) = (s+1)^2(s+2) \end{array}\right. \Rightarrow LCM = (s+1)^2(s+2) \tag{2.318}$$

As far as $\det(s\mathbb{I} - \mathbf{A}) \neq LCM$ *we conclude that the realization is not minimal. Furthermore the characteristic polynomial of any state matrix of a minimal realization shall be the LCM, that is here* $(s+1)^2(s+2)$.

An other realization of $\mathbf{F}(s)$ *can be obtained by writing* $\mathbf{F}(s)$ *in diagonal (or modal) form as explained in section 2.5.4:*

$$\mathbf{F}(s) = \frac{1}{s+1}\begin{bmatrix} 1 \\ 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix} + \frac{1}{s+1}\begin{bmatrix} 0 \\ 3 \end{bmatrix}\begin{bmatrix} 0 & 1 \end{bmatrix} + \frac{1}{s+2}\begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \end{bmatrix} \tag{2.319}$$

Then we get:

$$\mathbf{F}(s) = \left(\begin{array}{ccc|cc} -1 & & & 1 & 0 \\ & -1 & & 0 & 1 \\ & & -2 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 & 0 \\ 2 & 3 & 0 & 0 & 0 \end{array}\right) \tag{2.320}$$

Because $\mathbf{F}(s)$ *has distinct roots we can also use for this example Gilbert's realization as explained in section 2.6.2:*

$$\mathbf{F}(s) = \frac{1}{s+1}\mathbf{R}_1 + \frac{1}{s+2}\mathbf{R}_2 = \frac{1}{s+1}\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} + \frac{1}{s+2}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{2.321}$$

– *The rank of matrix* $\mathbf{R}_1 = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$ *is* $\rho_1 = 2$. *Thus we write* $\mathbf{R}_1 = \mathbf{C}_1\mathbf{B}_1$ *where* $\mathbf{C}_1$ *is a* $p \times \rho_1 = 2 \times 2$ *matrix and* $\mathbf{B}_1$ *is a* $\rho_1 \times m = 2 \times 2$ *matrix.*

We choose for example:

$$\begin{cases} \mathbf{C}_1 = \mathbf{R}_1 = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \\ \mathbf{B}_1 = \mathbb{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{cases} \tag{2.322}$$

– *The rank of matrix* $\mathbf{R}_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ *is* $\rho_2 = 1$. *Thus we write* $\mathbf{R}_2 = \mathbf{C}_2\mathbf{B}_2$ *where* $\mathbf{C}_2$ *is a* $p \times \rho_2 = 2 \times 1$ *matrix and* $\mathbf{B}_2$ *is a* $\rho_2 \times m = 1 \times 2$ *matrix. We choose for example:*

$$\begin{cases} \mathbf{C}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \mathbf{B}_2 = \mathbb{I} = \begin{bmatrix} 0 & 1 \end{bmatrix} \end{cases} \tag{2.323}$$

Then we get:

$$\mathbf{F}(s) = \left(\begin{array}{cc|c} \mathbf{\Lambda}_1 & 0 & \mathbf{B}_1 \\ \mathbf{0} & \lambda_2 & \mathbf{B}_2 \\ \hline \mathbf{C}_1 & \mathbf{C}_2 & \mathbf{D} \end{array}\right) = \left(\begin{array}{ccc|cc} -1 & & & 1 & 0 \\ & -1 & & 0 & 1 \\ & & -2 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 & 0 \\ 2 & 3 & 0 & 0 & 0 \end{array}\right) \tag{2.324}$$

For this example we get the same realization than (2.320).

With this realization we have $\det(s\mathbb{I} - \mathbf{A}) = LCM = (s+1)^2(s+2)$. *Thus we conclude that this realization is minimal.*

■

2.6.2 Gilbert's minimal realization

Let's write the $p \times m$ transfer function $\mathbf{F}(s)$ as follows:

$$\mathbf{F}(s) = \mathbf{F}_{sp}(s) + \mathbf{D} \tag{2.325}$$

where $\mathbf{D}$ is a constant matrix and $\mathbf{F}_{sp}(s)$ a strictly proper transfer function:

$$\begin{cases} \lim_{s\to\infty} \mathbf{F}(s) = \mathbf{D} \\ \lim_{s\to\infty} \mathbf{F}_{sp}(s) = \mathbf{0} \end{cases} \tag{2.326}$$

We consider in that section MIMO systems in which the denominator polynomial of the strictly proper transfer function $\mathbf{F}_{sp}(s)$ has distinct roots:

$$\mathbf{F}_{sp}(s) = \sum_i \frac{\mathbf{R}_i}{s - \lambda_i} \tag{2.327}$$

The residue $\mathbf{R}_i$ can be obtained as:

$$\mathbf{R}_i = \lim_{s\to\lambda_i} (s - \lambda_i)\,\mathbf{F}_{sp}(s) = \lim_{s\to\lambda_i} (s - \lambda_i)\,\mathbf{F}(s) \tag{2.328}$$

Now let ρ_i be the rank of $\mathbf{R}_i$:

$$\rho_i = \operatorname{rank}(\mathbf{R}_i) \tag{2.329}$$

and write $\mathbf{R}_i$ as follows where $\mathbf{C}_i$ is a $p \times \rho_i$ constant matrix and $\mathbf{B}_i$ a $\rho_i \times m$ constant matrix:

$$\mathbf{R}_i = \mathbf{C}_i \mathbf{B}_i \tag{2.330}$$

Then a realization of the transfer function $\mathbf{F}(s)$ is the following[4]:

$$\mathbf{F}(s) = \left(\begin{array}{ccc|c} \mathbf{\Lambda}_1 & \mathbf{0} & \cdots & \mathbf{B}_1 \\ \mathbf{0} & \ddots & & \vdots \\ \vdots & & \mathbf{\Lambda}_m & \mathbf{B}_m \\ \hline \mathbf{C}_1 & \cdots & \mathbf{C}_m & \mathbf{D} \end{array}\right) \tag{2.331}$$

where matrices $\mathbf{\Lambda}_i = \lambda_i \mathbb{I}_{\rho_i}$ are diagonal matrices of size ρ_i.

Moreover Gilbert's realization is minimal with order n given by:

$$n = \sum_i \rho_i \tag{2.332}$$

2.6.3 Ho-Kalman algorithm

To get a minimal realization $\left(\begin{array}{c|c} \mathbf{A}_m & \mathbf{B}_m \\ \hline \mathbf{C}_m & \mathbf{D} \end{array}\right)$ from a realization $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$ we can use the Ho-Kalman algorithm which is described hereafter:

- Let r be the dimension of the state matrix $\mathbf{A}$, which may not be minimal. First compute the observability matrix $\mathbf{Q}_o$ and the controllability matrix $\mathbf{Q}_c$ of the realization $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$:

$$\begin{cases} \mathbf{Q}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{r-1} \end{bmatrix} \\ \mathbf{Q}_c = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{r-1}\mathbf{B} \end{bmatrix} \end{cases} \tag{2.333}$$

  The realization is minimal if and only if:

$$\operatorname{rank}(\mathbf{Q}_o) = \operatorname{rank}(\mathbf{Q}_c) \tag{2.334}$$

  In all situations the dimension n of the system is given by:

$$n = \min(\operatorname{rank}(\mathbf{Q}_o), \operatorname{rank}(\mathbf{Q}_c)) \tag{2.335}$$

[4] Thomas Kailath, Linear Systems, Prentice-Hall, 1st Edition

- If the realization $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$ is not minimal then compute the singular value decomposition (*svd*) of the product $\mathbf{Q}_o\mathbf{Q}_c$:

$$\mathbf{Q}_o\mathbf{Q}_c = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{2.336}$$

  Matrix $\mathbf{\Sigma}$ is a rectangular diagonal matrix with non-negative real coefficients situated on its diagonal. The strictly positive coefficients of $\mathbf{\Sigma}$ are called the singular values of $\mathbf{Q}_o\mathbf{Q}_c$. The number of singular values of $\mathbf{Q}_o\mathbf{Q}_c$ (which are the strictly positive coefficients within the diagonal of matrix $\mathbf{\Sigma}$) is the dimension n of the system. Again note that $n \neq r$ if the realization $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$ is not minimal.

- Let $\mathbf{\Sigma}_n$ be the square diagonal matrix built from the n singular values of $\mathbf{Q}_o\mathbf{Q}_c$ (which are the non-zero coefficients within the diagonal matrix $\mathbf{\Sigma}$), $\mathbf{U}_n$ the matrix built from the n columns of $\mathbf{U}$ corresponding to the aux n singular values and $\mathbf{V}_n$ the matrix built from the n columns of $\mathbf{V}$ corresponding to the aux n singular values:

$$\mathbf{Q}_o\mathbf{Q}_c = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \begin{bmatrix} \mathbf{U}_n & \mathbf{U}_s \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_s \end{bmatrix} \begin{bmatrix} \mathbf{V}_n^T \\ \mathbf{V}_s^T \end{bmatrix} \tag{2.337}$$

- Matrices $\mathcal{O}_n$ and $\mathcal{C}_n$ are defined as follows:

$$\mathcal{O}_n\mathcal{C}_n = \mathbf{U}_n\mathbf{\Sigma}_n\mathbf{V}_n^T \text{ where } \begin{cases} \mathcal{O}_n = \mathbf{U}_n\mathbf{\Sigma}_n^{1/2} \\ \mathcal{C}_n = \mathbf{\Sigma}_n^{1/2}\mathbf{V}_n^T \end{cases} \tag{2.338}$$

- Then the state matrix $\mathbf{A}_m$ of a minimal realization is obtained as follows:

$$\mathbf{A}_m = \mathbf{\Sigma}_n^{-1/2}\mathbf{U}_n^T\left(\mathbf{Q}_o\mathbf{A}\mathbf{Q}_c\right)\mathbf{V}_n\mathbf{\Sigma}_n^{-1/2} \tag{2.339}$$

- Let m be the number of inputs of the system and p its number of outputs and $\mathbb{I}_m$ the identity matrix of size m. Matrix $\mathbf{B}_m$ and $\mathbf{C}_m$ of the minimal realization are obtained as follows:

$$\begin{cases} \mathbf{B}_m = \mathcal{C}_n \begin{bmatrix} \mathbb{I}_m \\ \mathbf{0} \\ \vdots \end{bmatrix} = \mathbf{\Sigma}_n^{1/2}\mathbf{V}_n^T \begin{bmatrix} \mathbb{I}_m \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \\ \mathbf{C}_m = \begin{bmatrix} \mathbb{I}_p & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \mathcal{O}_n = \begin{bmatrix} \mathbb{I}_p & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \mathbf{U}_n\mathbf{\Sigma}_n^{1/2} \end{cases} \tag{2.340}$$

- Matrix $\mathbf{D}$ is independent of the realization.

Chapter 3

Analysis of Linear Time Invariant systems

3.1 Introduction

This chapter is dedicated to the analysis of linear dynamical systems. More specifically we will concentrate on the solution of the state equation and we will present the notions of controllability, observability and stability. Those notions will enable the modal analysis of Linear Time Invariant (LTI) dynamical systems

3.2 Solving the time invariant state equation

We have seen that the state equation attached to a linear time invariant system is the following:

$$\dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \tag{3.1}$$

The purpose of this section is to obtain the general solution of this linear differential equation, which is actually a vector equation.

The solution of the non-homogeneous state equation $\dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t)$ can be obtained by the Laplace transform. Indeed the Laplace transform of this equation yields:

$$s\underline{X}(s) - \underline{x}(0) = \mathbf{A}\underline{X}(s) + \mathbf{B}\underline{U}(s) \tag{3.2}$$

That is:

$$(s\mathbb{I} - \mathbf{A})\,\underline{X}(s) = \underline{x}(0) + \mathbf{B}\underline{U}(s) \tag{3.3}$$

Pre-multiplying both sides of this equation by $(s\mathbb{I} - \mathbf{A})^{-1}$ leads to the following equation:

$$\underline{X}(s) = (s\mathbb{I} - \mathbf{A})^{-1}\,\underline{x}(0) + (s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B}\underline{U}(s) \tag{3.4}$$

By taking the inverse Laplace transform of this equation we get the

expression of the state vector $\underline{x}(t)$:

$$\begin{aligned}\underline{x}(t) &= \mathcal{L}^{-1}\left[\underline{X}(s)\right] \\ &= \mathcal{L}^{-1}\left[(s\mathbb{I}-\mathbf{A})^{-1}\underline{x}(0)+(s\mathbb{I}-\mathbf{A})^{-1}\mathbf{B}\underline{U}(s)\right] \\ &= \mathcal{L}^{-1}\left[(s\mathbb{I}-\mathbf{A})^{-1}\right]\underline{x}(0)+\mathcal{L}^{-1}\left[(s\mathbb{I}-\mathbf{A})^{-1}\mathbf{B}\underline{U}(s)\right]\end{aligned} \tag{3.5}$$

To inverse the preceding equation in the s domain and come back in the time domain we will use the following properties of the Laplace transform:

- Convolution theorem: let $x(t)$ and $y(t)$ be two causal scalar signals and denote by $X(s)$ and $Y(s)$ their Laplace transforms, respectively. Then the product $X(s)Y(s)$ is the Laplace transform of the convolution between $x(t)$ and $y(t)$ which is denoted by $x(t)*y(t)$:

$$X(s)Y(s)=\mathcal{L}\left[x(t)*y(t)\right]\Leftrightarrow\mathcal{L}^{-1}\left[X(s)Y(s)\right]=x(t)*y(t) \tag{3.6}$$

  Where:

$$x(t)*y(t)=\int_0^t x(t-\tau)y(\tau)d\tau \tag{3.7}$$

  This relation is readily extended to the vector case where $\mathbf{X}(t)$ is a matrix and $\underline{y}(t)$ a vector:

$$\mathcal{L}^{-1}\left[\mathbf{X}(s)\underline{Y}(s)\right]=\mathbf{X}(t)*\underline{y}(t)=\int_0^t \mathbf{X}(t-\tau)\underline{y}(\tau)d\tau \tag{3.8}$$

- Laplace transform of exponential matrix: in the scalar case we have seen that:

$$\mathcal{L}\left[e^{at}\right]=\frac{1}{s-a}=(s-a)^{-1}\Leftrightarrow\mathcal{L}^{-1}\left[(s-a)^{-1}\right]=e^{at} \tag{3.9}$$

  This relation is readily extended to the vector case as follows:

$$\mathcal{L}\left[e^{\mathbf{A}t}\right]=(s\mathbb{I}-\mathbf{A})^{-1}\Leftrightarrow\mathcal{L}^{-1}\left[(s\mathbb{I}-\mathbf{A})^{-1}\right]=e^{\mathbf{A}t} \tag{3.10}$$

Thus the inverse Laplace transform of Equation (3.5) leads to the expression of the state vector $\underline{x}(t)$ which solves the state equation (3.1):

$$\underline{x}(t)=e^{\mathbf{A}t}\underline{x}(0)+\int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau \tag{3.11}$$

The solution $\underline{x}(t)$ of Equation (3.5) is often referred to as the state trajectory or the system trajectory.

Exponential $e^{\mathbf{A}t}$ is defined as the transition matrix $\Phi(t)$:

$$\mathbf{\Phi}(t)=e^{\mathbf{A}t} \tag{3.12}$$

In the more general case of time dependent linear differential equation of the form $\dot{\underline{x}}(t)=\mathbf{A}(t)\underline{x}(t)+\mathbf{B}(t)\underline{u}(t)$ the expression of the state vector is $\underline{x}(t)=\mathbf{\Phi}(t,t_0)\underline{x}(0)+\int_{t_0}^t\mathbf{\Phi}(t,\tau)\mathbf{B}\underline{u}(\tau)d\tau$ where $\mathbf{\Phi}(t,t_0)$ is also named the transition

matrix (or Green's matrix). In this case the transition matrix is a solution of the homogeneous equation $\frac{\partial \mathbf{\Phi}(t,t_0)}{\partial t} = \mathbf{A}(t)\mathbf{\Phi}(t,t_0)$. In addition $\mathbf{\Phi}(t,t) = \mathbb{I}\ \forall t$ and $\mathbf{\Phi}(t,t_0) = \phi(t)\phi^{-1}(t_0)$ where $\phi(t)$ is the solution of $\dot{\underline{x}}(t) = \mathbf{A}(t)\underline{x}(t)$. For a linear time invariant system the transition matrix $\mathbf{\Phi}(t,t_0)$ is $\mathbf{\Phi}(t,t_0) = e^{\mathbf{A}(t-t_0)}$; as far as for the time invariant case the initial time t_0 is meaningless we can choose $t_0 = 0$ and we retrieve $\mathbf{\Phi}(t) = e^{\mathbf{A}t}$.

3.3 Output response

We have seen that the output vector $\underline{y}(t)$ of the state-space representation is obtained using the output equation. Using the expression of the state vector $\underline{x}(t)$ we get:

$$\begin{aligned} \underline{y}(t) &= \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \\ &= \mathbf{C}\left(e^{\mathbf{A}t}\underline{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau\right) + \mathbf{D}\underline{u}(t) \\ &= \mathbf{C}e^{\mathbf{A}t}\underline{x}(0) + \int_0^t \mathbf{C}e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau + \mathbf{D}\underline{u}(t) \end{aligned} \tag{3.13}$$

- The term $\mathbf{C}e^{\mathbf{A}t}\underline{x}(0)$ is called the zero-input response (or output) of the system; this is the response of the system when there is no input signal $\underline{u}(t)$ applied on the system;
- The term $\int_0^t \mathbf{C}e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau + \mathbf{D}\underline{u}(t)$ is called the zero-state output (or response) of the system; this is the response of the system when there is no initial condition $\underline{x}(0)$ applied on the system.

3.4 Impulse and unit step responses

The impulse response of a dynamical system is the zero-state output of the system when the input signal $\underline{u}(t)$ is the impulse $\delta(t)$ called the Dirac delta function.

Setting in (3.13) the input signal $\underline{u}(t)$ to the Dirac delta function $\underline{\delta}(t)$ and putting the initial conditions $\underline{x}(0)$ to zero leads to the following expression of the impulse response of the system:

$$\underline{y}(t) = \int_0^t \mathbf{C}e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{\delta}(\tau)d\tau + \mathbf{D}\underline{\delta}(t) \tag{3.14}$$

The term $\int_0^t \mathbf{C}e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{\delta}(\tau)d\tau$ can be expressed as the convolution between the matrix $\mathbf{C}e^{\mathbf{A}t}\mathbf{B}$ and the input vector $\underline{\delta}(\tau)$. We get:

$$\underline{y}(t) = \mathbf{C}e^{\mathbf{A}t}\mathbf{B} * \underline{\delta}(t) + \mathbf{D}\underline{\delta}(t) \tag{3.15}$$

Using the fact that the Dirac delta function $\underline{\delta}(t)$ is the neutral element for convolution we can write $\mathbf{C}e^{\mathbf{A}t}\mathbf{B} * \underline{\delta}(t) = \mathbf{C}e^{\mathbf{A}t}\mathbf{B}$. Consequently the output vector $\underline{y}(t)$, that is the impulse response of a linear time invariant system which will be denoted $\mathbf{h}(t)$, can be expressed as follows:

$$\underline{y}(t) := \mathbf{h}(t) = \mathbf{C}e^{\mathbf{A}t}\mathbf{B} + \mathbf{D}\delta(t) \tag{3.16}$$

The Laplace transform of the impulse response $\mathbf{h}(t)$ is defined to be the transfer function $\mathbf{F}(s)$ of the system. Using the fact that the Laplace transform of the Dirac delta function $\delta(t)$ is 1 we retrieve the following expression for the transfer function $\mathbf{F}(s)$ of the linear system:

$$\mathbf{F}(s) = \mathcal{L}\left[\mathbf{h}(t)\right] = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} \tag{3.17}$$

The unit step response is the response of the system to the unit step input. Setting in (3.13) the input signal $\underline{u}(t)$ to $\underline{u}(t) = \underline{1}\ \forall t > 0$ and putting the initial conditions $\underline{x}(0)$ to zero leads to the following expression of the unit step response of the system:

$$\begin{aligned}
\underline{y}(t) &= \int_0^t \mathbf{C}e^{\mathbf{A}(t-\tau)}\mathbf{B}d\tau + \mathbf{D}\underline{1} \\
&= \mathbf{C}e^{\mathbf{A}t}\left(\int_0^t e^{-\mathbf{A}\tau}d\tau\right)\mathbf{B} + \mathbf{D}\underline{1} \\
&= \mathbf{C}e^{\mathbf{A}t}\left(-\mathbf{A}^{-1}e^{-\mathbf{A}\tau}\Big|_{\tau=0}^{t}\right)\mathbf{B} + \mathbf{D}\underline{1} \\
&= \mathbf{C}e^{\mathbf{A}t}\left(\mathbf{A}^{-1} - \mathbf{A}^{-1}e^{-\mathbf{A}t}\right)\mathbf{B} + \mathbf{D}\underline{1}
\end{aligned} \tag{3.18}$$

Using the fact that $e^{\mathbf{A}t}\mathbf{A}^{-1} = \mathbf{A}^{-1}e^{\mathbf{A}t}$ (which is easy to show using the series expansion of $e^{\mathbf{A}t}$) and assuming that matrix $\mathbf{A}^{-1}$ exists, we finally get the following expression for the unit step response of the system:

$$\underline{y}(t) = \mathbf{C}\mathbf{A}^{-1}\left(e^{\mathbf{A}t} - \mathbb{I}\right)\mathbf{B} + \mathbf{D}\underline{1} \tag{3.19}$$

3.5 Matrix exponential

3.5.1 Definition

Let $\mathbf{A}$ be a $n \times n$ square matrix. The matrix exponential is a $n \times n$ matrix which is defined by analogy with the scalar exponential and its series as follows:

$$e^{\mathbf{A}t} = \sum_{k=0}^{\infty} \frac{(\mathbf{A}t)^k}{k!} = \mathbb{I} + \sum_{k=1}^{\infty} \frac{(\mathbf{A}t)^k}{k!} \tag{3.20}$$

This calculus involves an infinity of terms and it is in general impossible to compute it by hand except for some specific cases, for example if matrix $\mathbf{A}$ is nilpotent.

A matrix $\mathbf{A}$ is nilpotent if there exists an integer k such that $\mathbf{A}^k = \mathbf{0}$. The smallest value of k is called the index of nilpotency (of the nilpotent matrix). In this case the matrix exponential $e^{\mathbf{A}t}$ can be computed directly from the series expansion as the series terminates after a finite number of terms:

$$e^{\mathbf{A}t} = \mathbb{I} + \mathbf{A}t + \mathbf{A}^2\frac{t^2}{2!} + \cdots + \mathbf{A}^{k-1}\frac{t^{k-1}}{(k-1)!} \tag{3.21}$$

A necessary and sufficient condition for a $n \times n$ square matrix $\mathbf{A}$ to be nilpotent is that its characteristic polynomial $\det\left(s\mathbb{I} - \mathbf{A}\right)$ is equal to s^n:

$$\mathbf{A}^k = \mathbf{0} \Leftrightarrow \det\left(s\mathbb{I} - \mathbf{A}\right) = s^n \text{ where } k \leq n \tag{3.22}$$

We recall that the minimal polynomial $\pi_A(s)$ of a matrix $\mathbf{A}$ is the monic polynomial (a monic polynomial is a polynomial in which the leading coefficient (the nonzero coefficient of highest degree) is equal to 1) of least degree such that $\pi_A(\mathbf{A}) = \mathbf{0}$. The minimal polynomial divides the characteristic polynomial $\chi_A(s) := \det(s\mathbb{I} - \mathbf{A})$ of matrix $\mathbf{A}$. Consequently its degree is lower or equal to the order n of matrix $\mathbf{A}$.

A matrix $\mathbf{A}$ is said to be *cyclic* if and only if its characteristic polynomial is the same than its minimal polynomial.

Furthermore matrix $\mathbf{A}$ is diagonalizable if and only if its minimal polynomial's roots are all of multiplicity one.

The previous result to compute $e^{\mathbf{A}t}$ when $\mathbf{A}$ is nilpotent can be generalized to the case where the minimal polynomial $\pi_A(s)$ of $\mathbf{A}$ reads $(s-\lambda)^k$. Indeed we get in this case:

$$\pi_A(s) = (s-\lambda)^k \Rightarrow (\mathbf{A} - \lambda\mathbb{I})^k = \mathbf{0} \tag{3.23}$$

Thus matrix $\mathbf{A} - \lambda\mathbb{I}$ is nilpotent and we can write:

$$e^{(\mathbf{A}-\lambda\mathbb{I})t} = \mathbb{I} + (\mathbf{A} - \lambda\mathbb{I})\,t + (\mathbf{A} - \lambda\mathbb{I})^2 \frac{t^2}{2!} + \cdots + (\mathbf{A} - \lambda\mathbb{I})^{k-1} \frac{t^{k-1}}{(k-1)!} \tag{3.24}$$

As soon as matrices $\mathbf{A}$ and $\lambda\mathbb{I}$ commute the following relation holds:

$$e^{(\mathbf{A}-\lambda\mathbb{I})t} = e^{-\lambda t} e^{\mathbf{A}t} \Rightarrow e^{\mathbf{A}t} = e^{\lambda t}\, e^{(\mathbf{A}-\lambda\mathbb{I})t} \tag{3.25}$$

Thus as soon as $\pi_A(s) = (s-\lambda)^k$ we finally get the following result:

$$e^{\mathbf{A}t} = \\ e^{\lambda t}\left(\mathbb{I} + (\mathbf{A} - \lambda\mathbb{I})\,t + (\mathbf{A} - \lambda\mathbb{I})^2 \frac{t^2}{2!} + \cdots + (\mathbf{A} - \lambda\mathbb{I})^{k-1} \frac{t^{k-1}}{(k-1)!}\right) \tag{3.26}$$

Example 3.1. *Let* $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. *The characteristic polynomial of* $\mathbf{A}$ *is:*

$$\det(s\mathbb{I} - \mathbf{A}) = \det\left(\begin{bmatrix} s & -1 \\ 0 & s \end{bmatrix}\right) = s^2 \tag{3.27}$$

Consequently matrix $\mathbf{A}$ *is nilpotent and* $e^{\mathbf{A}t}$ *can be computed as follows:*

$$\begin{aligned} &\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \Rightarrow \mathbf{A}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ &\Rightarrow e^{\mathbf{A}t} = \mathbb{I} + \mathbf{A}t = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \end{aligned} \tag{3.28}$$

■

3.5.2 Properties

The following properties hold[1]:

- Value at $t = 0$:
$$e^{\mathbf{A}t}\big|_{t=0} = e^{\mathbf{0}} = \mathbb{I} \tag{3.29}$$
- Derivation:
$$\frac{d}{dt}e^{\mathbf{At}} = \mathbf{A}e^{\mathbf{At}} = e^{\mathbf{At}}\mathbf{A} \tag{3.30}$$
- Integration:
$$e^{\mathbf{At}} = \mathbb{I} + \mathbf{A}\int_0^t e^{\mathbf{A}\tau}d\tau \tag{3.31}$$
- In general:
$$e^{(\mathbf{A}+\mathbf{B})t} \neq e^{\mathbf{A}t}e^{\mathbf{B}t} \neq e^{\mathbf{B}t}e^{\mathbf{A}t} \tag{3.32}$$
  Nevertheless if matrices $\mathbf{A}$ and $\mathbf{B}$ commute (meaning that $\mathbf{AB} = \mathbf{BA}$) then:
$$e^{(\mathbf{A}+\mathbf{B})t} = e^{\mathbf{A}t}e^{\mathbf{B}t} = e^{\mathbf{B}t}e^{\mathbf{A}t} \tag{3.33}$$
  As far as the product $\mathbf{AA}$ commutes we have:
$$e^{\mathbf{A}t}e^{\mathbf{A}\tau} = e^{\mathbf{A}(t+\tau)} = e^{\mathbf{A}\tau}e^{\mathbf{A}t} \tag{3.34}$$
  And thus setting τ to $-t$ we get:
$$\left(e^{\mathbf{A}t}\right)^{-1} = e^{-\mathbf{A}t} \tag{3.35}$$
- Let $\lambda(\mathbf{A})$ be the eigenvalues of matrix $\mathbf{A}$. Then:
$$\lambda\left(e^{\mathbf{A}t}\right) = e^{\lambda(\mathbf{A})t} \tag{3.36}$$
- Let $\det(\mathbf{A})$ be the determinant of matrix $\mathbf{A}$ and $\mathrm{tr}(\mathbf{A})$ be the trace of matrix $\mathbf{A}$. Then:
$$\det\left(e^{\mathbf{A}t}\right) = e^{\mathrm{tr}(\mathbf{A})t} \tag{3.37}$$

Example 3.2. *Let's consider the following matrices* $\mathbf{A}$ *and* $\mathbf{B}$*:*

$$\begin{cases} \mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\ \\ \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \end{cases} \tag{3.38}$$

It is clear that $\mathbf{A}$ *et* $\mathbf{B}$ *do not commute. Indeed:*

$$\mathbf{AB} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \neq \mathbf{BA} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{3.39}$$

[1] https://en.wikipedia.org/wiki/Matrix_exponential

Consequently we expect that $e^{\mathbf{A}t}e^{\mathbf{B}t} \neq e^{\mathbf{B}t}e^{\mathbf{A}t}$. *We will check it by using the preceding definitions and properties:*

$$e^{\mathbf{A}t} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \tag{3.40}$$

And:

$$\begin{aligned} (\mathbf{B}t)^k = \begin{bmatrix} t^k & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow e^{\mathbf{B}t} = \mathbb{I} + \textstyle\sum_{k=1}^{\infty} \frac{(\mathbf{B}t)^k}{k!} = \begin{bmatrix} 1 + \sum_{k=1}^{\infty} \frac{t^k}{k!} & 0 \\ 0 & 1 \end{bmatrix} \\ \Leftrightarrow e^{\mathbf{B}t} = \begin{bmatrix} e^t & 0 \\ 0 & 1 \end{bmatrix} \end{aligned} \tag{3.41}$$

It is clear that:

$$e^{\mathbf{A}t}e^{\mathbf{B}t} = \begin{bmatrix} e^t & t \\ 0 & 1 \end{bmatrix} \neq e^{\mathbf{B}t}e^{\mathbf{A}t} = \begin{bmatrix} e^t & te^t \\ 0 & 1 \end{bmatrix} \tag{3.42}$$

We can also easily check the following properties:

$$\begin{cases} \frac{d}{dt}e^{\mathbf{At}} = \mathbf{A}e^{\mathbf{At}} = e^{\mathbf{At}}\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\ \frac{d}{dt}e^{\mathbf{Bt}} = \mathbf{B}e^{\mathbf{Bt}} = e^{\mathbf{Bt}}\mathbf{B} = \begin{bmatrix} e^t & 0 \\ 0 & 0 \end{bmatrix} \end{cases} \tag{3.43}$$

■

3.5.3 Computation of $e^{\mathbf{A}t}$ using the diagonal form of A

We will assume in that section that matrix $\mathbf{A}$ is diagonalizable or equivalently that matrix $\mathbf{A}$ has linearly independent eigenvectors; this means that for all eigenvalues λ_i of $\mathbf{A}$ the rank of matrix $\lambda_i\mathbb{I} - \mathbf{A}$ is equal to the size of $\mathbf{A}$ minus the multiplicity of λ_i.

Assuming that matrix $\mathbf{A}$ is diagonalizable then there exists a similarity transformation such that:

$$\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1} \text{ where } \mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \tag{3.44}$$

The change of basis matrix $\mathbf{P}$, as well as its inverse $\mathbf{P}^{-1}$, can be obtained as follows:

- Let $\underline{v}_i$ be the eigenvector of $\mathbf{A}$ corresponding to eigenvalue λ_i. As far as the $n \times n$ matrix $\mathbf{A}$ is assumed to have n linearly independent eigenvectors we can write:

$$\begin{cases} \mathbf{A}\underline{v}_1 = \lambda_1\underline{v}_1 \\ \vdots \\ \mathbf{A}\underline{v}_n = \lambda_n\underline{v}_n \end{cases} \tag{3.45}$$

– The preceding relation can be written in a vector form as follows:

$$\mathbf{A}\left[\begin{array}{ccc}\underline{v}_1 & \dots & \underline{v}_n\end{array}\right] = \left[\begin{array}{ccc}\underline{v}_1 & \dots & \underline{v}_n\end{array}\right]\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix} \tag{3.46}$$

– Identifying the preceding equation with $\mathbf{AP} = \mathbf{P\Lambda}$ we finally get

$$\mathbf{P} = \left[\begin{array}{cccc}\underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_n\end{array}\right] \tag{3.47}$$

– Furthermore let $\underline{w}_i$ be the eigenvectors of matrix $\mathbf{A}^T$:

$$\mathbf{A}^T\underline{w}_j = \lambda_j\underline{w}_j \Leftrightarrow \underline{w}_j^T\mathbf{A} = \lambda_j\underline{w}_j^T \tag{3.48}$$

It can be seen that vectors $\underline{w}_j$ et $\underline{v}_i$ are orthogonal. Indeed:

$$\begin{array}{l}\lambda_j\underline{w}_j^T\underline{v}_i = \underline{w}_j^T\mathbf{A}\underline{v}_i = \underline{w}_j^T\lambda_i\underline{v}_i \Leftrightarrow (\lambda_i - \lambda_j)\,\underline{w}_j^T\underline{v}_i = 0 \\ \Rightarrow \underline{w}_j^T\underline{v}_i = 0 \text{ si } j \neq i\end{array} \tag{3.49}$$

Thus imposing $\underline{w}_i^T\underline{v}_i = 1\ \forall i$, the inverse of matrix $\mathbf{P}$ is obtained as follows:

$$\underline{w}_i^T\underline{v}_i = 1\ \forall i \Rightarrow \mathbf{P}^{-1} = \begin{bmatrix}\underline{w}_1^T \\ \underline{w}_2^T \\ \vdots \\ \underline{w}_n^T\end{bmatrix} \tag{3.50}$$

Indeed using $\underline{w}_j^T\underline{v}_i = \begin{cases}0 \text{ if } j \neq i \\ 1 \text{ if } j = i\end{cases}$ we get:

$$\begin{aligned}\mathbf{P}^{-1}\mathbf{P} &= \begin{bmatrix}\underline{w}_1^T \\ \underline{w}_2^T \\ \vdots \\ \underline{w}_n^T\end{bmatrix}\left[\begin{array}{cccc}\underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_n\end{array}\right] \\ &= \begin{bmatrix}\underline{w}_1^T\underline{v}_1 & \underline{w}_1^T\underline{v}_2 & \cdots & \underline{w}_1^T\underline{v}_n \\ \underline{w}_2^T\underline{v}_1 & \underline{w}_2^T\underline{v}_2 & \cdots & \underline{w}_2^T\underline{v}_n \\ \vdots & \vdots & & \vdots \\ \underline{w}_n^T\underline{v}_1 & \underline{w}_n^T\underline{v}_2 & \cdots & \underline{w}_n^T\underline{v}_n\end{bmatrix} \\ &= \begin{bmatrix}1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & 1\end{bmatrix} = \mathbb{I}\end{aligned} \tag{3.51}$$

Then, as soon as matrix $\mathbf{A}$ diagonalizable, $e^{\mathbf{A}t}$ can be obtained using the following relation :

$$e^{\mathbf{A}t} = \mathbf{P}e^{\mathbf{\Lambda}t}\mathbf{P}^{-1} = \sum_{i=1}^{n}\underline{v}_i\underline{w}_i^T e^{\lambda_i t} \tag{3.52}$$

The preceding relation is easily obtained by using the series development of the exponential:

$$e^{\mathbf{P}\mathbf{\Lambda}t\mathbf{P}^{-1}} = \sum_{k=0}^{\infty} \frac{\left(\mathbf{P}\mathbf{\Lambda}t\mathbf{P}^{-1}\right)^k}{k!} = \sum_{k=0}^{\infty} \frac{\mathbf{P}\left(\mathbf{\Lambda}t\right)^k \mathbf{P}^{-1}}{k!} = \mathbf{P}\left(\sum_{k=0}^{\infty} \frac{\left(\mathbf{\Lambda}t\right)^k}{k!}\right)\mathbf{P}^{-1} \tag{3.53}$$

As far as the diagonal (or modal) matrix $\mathbf{\Lambda}$ is diagonal we get:

$$\begin{cases} \mathbf{\Lambda}^k = \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} \\ e^{\mathbf{\Lambda}t} = \sum_{k=0}^{\infty} \frac{(\mathbf{\Lambda}t)^k}{k!} \end{cases} \Rightarrow e^{\mathbf{\Lambda}t} = \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} \tag{3.54}$$

Thus using the expression of $\mathbf{P}$ and $\mathbf{P}^{-1}$:

$$\begin{aligned} e^{\mathbf{A}t} &= \mathbf{P}e^{\mathbf{\Lambda}t}\mathbf{P}^{-1} = \mathbf{P}\begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix}\mathbf{P}^{-1} \\ &= \begin{bmatrix} \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_n \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} \underline{w}_1^T \\ \underline{w}_2^T \\ \vdots \\ \underline{w}_n^T \end{bmatrix} \end{aligned} \tag{3.55}$$

We finally get:

$$e^{\mathbf{A}t} = \sum_{i=1}^{n} \underline{v}_i e^{\lambda_i t} \underline{w}_i^T = \sum_{i=1}^{n} \underline{v}_i \underline{w}_i^T e^{\lambda_i t} \tag{3.56}$$

Example 3.3. *Compute* $e^{\mathbf{A}t}$ *where* $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & -5 \end{bmatrix}$.

Le characteristic polynomial of $\mathbf{A}$ *reads:*

$$\det\left(s\mathbb{I} - \mathbf{A}\right) = \det\left(\begin{bmatrix} s-1 & -2 \\ 0 & s+5 \end{bmatrix}\right) = (s-1)(s+5) \tag{3.57}$$

The two eigenvalues $\lambda_1 = 1$ *and* $\lambda_2 = -5$ *of* $\mathbf{A}$ *are distinct. Since the size of* $\mathbf{A}$ *is equal to the number of the distinct eigenvalues we conclude that matrix* $\mathbf{A}$ *is diagonalizable.*

- *Let* $\underline{v}_1 = \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix}$ *be the eigenvector of* $\mathbf{A}$ *corresponding to* $\lambda_1 = 1$. *We have:*

$$\begin{aligned} &\begin{bmatrix} 1 & 2 \\ 0 & -5 \end{bmatrix}\begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = 1 \times \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} \\ &\Leftrightarrow \begin{cases} v_{11} + 2v_{12} = v_{11} \\ -5v_{12} = v_{12} \end{cases} \\ &\Rightarrow v_{12} = 0 \end{aligned} \tag{3.58}$$

Thus the expression of eigenvector $\underline{v}_1$ *is:*

$$\underline{v}_1 = \begin{bmatrix} v_{11} \\ 0 \end{bmatrix} \tag{3.59}$$

- *Let* $\underline{v}_2 = \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix}$ *be the eigenvector of* $\mathbf{A}$ *corresponding to* $\lambda_2 = -5$. *We have:*

$$\begin{aligned} &\begin{bmatrix} 1 & 2 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = -5 \times \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} \\ &\Leftrightarrow \begin{cases} v_{21} + 2v_{22} = -5v_{21} \\ -5v_{22} = -5v_{22} \end{cases} \\ &\Rightarrow 6v_{21} + 2v_{22} = 0 \\ &\Leftrightarrow v_{22} = -3v_{21} \end{aligned} \tag{3.60}$$

Thus the expression of eigenvector $\underline{v}_2$ *is:*

$$\underline{v}_2 = \begin{bmatrix} v_{21} \\ -3v_{21} \end{bmatrix} \tag{3.61}$$

- *Let* $\underline{w}_1 = \begin{bmatrix} w_{11} \\ w_{12} \end{bmatrix}$ *be the eigenvector of* $\mathbf{A}^T$ *corresponding to* $\lambda_1 = 1$. *We have:*

$$\begin{aligned} &\begin{bmatrix} 1 & 0 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} w_{11} \\ w_{12} \end{bmatrix} = 1 \times \begin{bmatrix} w_{11} \\ w_{12} \end{bmatrix} \\ &\Leftrightarrow \begin{cases} w_{11} = w_{11} \\ 2w_{11} - 5w_{12} = w_{12} \end{cases} \\ &\Rightarrow 2w_{11} - 6w_{12} = 0 \\ &\Leftrightarrow w_{11} = 3w_{12} \end{aligned} \tag{3.62}$$

Thus the expression of eigenvector $\underline{w}_1$ *is:*

$$\underline{w}_1 = \begin{bmatrix} 3w_{12} \\ w_{12} \end{bmatrix} \tag{3.63}$$

It is clear that $\underline{w}_1$ *and* $\underline{v}_2$ *are orthogonal:*

$$\underline{w}_1^T \underline{v}_2 = \begin{bmatrix} 3w_{12} & w_{12} \end{bmatrix} \begin{bmatrix} v_{21} \\ -3v_{21} \end{bmatrix} = 0 \tag{3.64}$$

- *Let* $\underline{w}_2 = \begin{bmatrix} w_{21} \\ w_{22} \end{bmatrix}$ *be the eigenvector of* $\mathbf{A}^T$ *corresponding to* $\lambda_2 = -5$. *We have:*

$$\begin{aligned} &\begin{bmatrix} 1 & 0 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} w_{21} \\ w_{22} \end{bmatrix} = -5 \times \begin{bmatrix} w_{21} \\ w_{22} \end{bmatrix} \\ &\Leftrightarrow \begin{cases} w_{21} = -5w_{21} \\ 2w_{21} - 5w_{22} = -5w_{22} \end{cases} \\ &\Rightarrow w_{21} = 0 \end{aligned} \tag{3.65}$$

Thus the expression of eigenvector $\underline{w}_2$ *is:*

$$\underline{w}_2 = \begin{bmatrix} 0 \\ w_{22} \end{bmatrix} \tag{3.66}$$

It is clear that $\underline{w}_2$ *and* $\underline{v}_1$ *are orthogonal:*

$$\underline{w}_2^T \underline{v}_1 = \begin{bmatrix} 0 & w_{22} \end{bmatrix} \begin{bmatrix} v_{11} \\ 0 \end{bmatrix} = 0 \tag{3.67}$$

Finally $\underline{v}_1$, $\underline{v}_2$, $\underline{w}_1$ *and* $\underline{w}_2$ *are chosen such that* $\underline{w}_1^T\underline{v}_1 = \underline{w}_2^T\underline{v}_2 = 1$. *We can chose for example:*

$$\underline{v}_1 = \begin{bmatrix} v_{11} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \underline{w}_1 = \begin{bmatrix} 3w_{12} \\ w_{12} \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{1}{3} \end{bmatrix} \tag{3.68}$$

And:

$$\underline{v}_2 = \begin{bmatrix} v_{21} \\ -3v_{21} \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} \\ 1 \end{bmatrix} \text{ and } \underline{w}_2 = \begin{bmatrix} 0 \\ w_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{3.69}$$

Then applying Equation (3.52) *we get:*

$$\begin{aligned} e^{\mathbf{A}t} &= \textstyle\sum_{i=1}^{n} \underline{v}_i\underline{w}_i^T e^{\lambda_i t} \\ &= \underline{v}_1\underline{w}_1^T e^{\lambda_1 t} + \underline{v}_2\underline{w}_2^T e^{\lambda_2 t} \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^t \begin{bmatrix} 1 & \frac{1}{3} \end{bmatrix} + \begin{bmatrix} -\frac{1}{3} \\ 1 \end{bmatrix} e^{-5t} \begin{bmatrix} 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & 0 \end{bmatrix} e^t + \begin{bmatrix} 0 & -\frac{1}{3} \\ 0 & 1 \end{bmatrix} e^{-5t} \\ &= \begin{bmatrix} e^t & \frac{1}{3}e^t - \frac{1}{3}e^{-5t} \\ 0 & e^{-5t} \end{bmatrix} \end{aligned} \tag{3.70}$$

We can check that $e^{\mathbf{A}t}\big|_{t=0} = \mathbb{I}$.

■

3.5.4 Computation of $e^{\mathbf{A}t}$ using the Jordan form of A

In that section we will assume that the *algebraic multiplicity* n_i of an eigenvalue λ_i of matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, that is the number of times that λ_i appears as a root of the characteristic polynomial $\det(s\mathbb{I} - \mathbf{A})$, is not equal (actually strictly greater) to its *geometric multiplicity* q_i, that is the dimension of the kernel of $\lambda_i\mathbb{I} - \mathbf{A}$. In that situation eigenvalue λ_i is said to be *degenerate* and matrix $\mathbf{A}$ is not diagonalizable. Nevertheless, a block-diagonal form $\mathbf{\Lambda}$ of matrix $\mathbf{A}$ can be achieved using its Jordan form:

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{J} \end{bmatrix} \tag{3.71}$$

As in the previous section, it can be shown that:

$$e^{\mathbf{A}t} = \mathbf{P}e^{\mathbf{\Lambda}t}\mathbf{P}^{-1} \tag{3.72}$$

where:

$$\begin{aligned} e^{\mathbf{\Lambda}t} &= e^{\begin{bmatrix} \mathbf{D}t & \mathbf{0} \\ \mathbf{0} & \mathbf{J}t \end{bmatrix}} \\ &= \textstyle\sum_{k=0}^{\infty} \frac{1}{k!} \begin{bmatrix} \mathbf{D}t & \mathbf{0} \\ \mathbf{0} & \mathbf{J}t \end{bmatrix}^k \\ &= \begin{bmatrix} \sum_{k=0}^{\infty} \frac{1}{k!} (\mathbf{D}t)^k & \mathbf{0} \\ \mathbf{0} & \sum_{k=0}^{\infty} \frac{1}{k!} (\mathbf{J}t)^k \end{bmatrix} \\ &= \begin{bmatrix} e^{\mathbf{D}t} & \mathbf{0} \\ \mathbf{0} & e^{\mathbf{J}t} \end{bmatrix} \end{aligned} \tag{3.73}$$

- Matrix $\mathbf{D}$ is a diagonal matrix which is build form the *non-degenerate* eigenvalues of $\mathbf{A}$. It has been seen in the previous section how to get $e^{\mathbf{D}t}$ in that case;

- Matrix $\mathbf{J}$ is a block-diagonal matrix build from Jordan blocks $\mathbf{J}_i$:

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & 0 & \\ 0 & \ddots & \ddots \\ & \ddots & \mathbf{J}_i \end{bmatrix} \Rightarrow e^{\mathbf{J}t} = \begin{bmatrix} e^{\mathbf{J}_1 t} & 0 & \\ 0 & \ddots & \ddots \\ & \ddots & e^{\mathbf{J}_i t} \end{bmatrix} \tag{3.74}$$

  Each Jordan blocks $\mathbf{J}_i$ corresponds to a *degenerate* eigenvalue λ_i.

We will see hereafter how to compute $e^{\mathbf{J}_i t}$.

- Assume that λ_i is a real eigenvalue, $\lambda_i \in \mathbb{R}$. Then the Jordan block $\mathbf{J}_i$ corresponding to λ_i is a square matrix of order n_i with the following structure (note the the number of 1 over the diagonal is equal to $n_i - q_i$):

$$\mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & & 0 \\ 0 & \ddots & \ddots & \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \lambda_i \end{bmatrix} = \lambda_i \mathbb{I}_{n_i} + \mathbf{N} \tag{3.75}$$

  Matrix $\mathbf{N}$ is a nilpotent matrix of order n_i which obviously commutes with $\lambda_i \mathbb{I}_{n_i}$. Consequently, $e^{\mathbf{J}_i t}$ reads:

$$\begin{aligned} & e^{\mathbf{J}_i t} = e^{\lambda_i \mathbb{I}_{n_i} t} e^{\mathbf{N}t} \\ & \Leftrightarrow e^{\mathbf{J}_i t} = e^{\lambda_i t} \left(\mathbb{I}_{n_i} + \mathbf{N}\frac{t}{1!} + \cdots + \mathbf{N}^{n_i - 1} \frac{t^{n_i - 1}}{(n_i - 1)!} \right) \end{aligned} \tag{3.76}$$

- Alternatively, if λ_i is a complex valued eigenvalue, then λ_i reads $a + jb$. In that case, the complex conjugate of λ_i is also an eigenvalue of $\mathbf{A}$. In that case, the Jordan block $\mathbf{J}_i$ corresponding to λ_i and its complex conjugate value is a square matrix of order $2n_i$ with the following structure:

$$\mathbf{J}_i = \begin{bmatrix} \mathbf{\Lambda}_i & \mathbb{I} & & 0 \\ 0 & \ddots & \ddots & \\ 0 & 0 & \ddots & \mathbb{I} \\ 0 & 0 & 0 & \mathbf{\Lambda}_i \end{bmatrix} \text{ où } \begin{cases} \mathbf{\Lambda}_i = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \\ \mathbb{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{cases} \tag{3.77}$$

  The Jordan block can be rewritten as follows:

$$\mathbf{J}_i = a\mathbb{I}_{2n_i} + \begin{bmatrix} \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix} & & \\ & \ddots & \\ & & \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix} \end{bmatrix} + \mathbf{N} \tag{3.78}$$

Matrix $\mathbf{N}$ is a nilpotent matrix of order $2\,n_i$ and the exponential of matrix $\mathbf{J}_i$ reads:

$$e^{\mathbf{J}_i t} = e^{at} \begin{bmatrix} \mathbf{R}(bt) & & \\ & \ddots & \\ & & \mathbf{R}(bt) \end{bmatrix} \left(\mathbb{I}_{2n_i} + \mathbf{N}\tfrac{t}{1!} + \cdots + \mathbf{N}^{n_i-1}\tfrac{t^{n_i-1}}{(n_i-1)!} \right) \tag{3.79}$$

where matrix $\mathbf{R}(bt)$ is the following rotation matrix:

$$\mathbf{R}(bt) = e^{\begin{bmatrix} 0 & bt \\ -bt & 0 \end{bmatrix}} = \begin{bmatrix} \cos(bt) & \sin(bt) \\ -\sin(bt) & \cos(bt) \end{bmatrix} \tag{3.80}$$

Example 3.4. *Compute $e^{\mathbf{J}t}$ where matrix $\mathbf{J}$ reads as follows:*

$$\mathbf{J} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \tag{3.81}$$

Matrix $\mathbf{J}$ is a Jordan matrix. Then $e^{\mathbf{J}t}$ is computed as follows:

$$\mathbf{J} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \mathbb{I}_2 + \mathbf{N} \tag{3.82}$$

As far as nilpotent $\mathbf{N}$ commute with identity matrix and that the index of nilpotency of matrix $\mathbf{N}$ is 2 we can write:

$$\begin{aligned} e^{\mathbf{J}t} &= e^{\begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix}} e^{\mathbf{N}t} \\ &= \left(\sum_{k=0}^{\infty} \frac{\begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix}^k}{k!} \right) \left(\sum_{k=0}^{2-1} \frac{\begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix}^k}{k!} \right) \\ &= \begin{bmatrix} e^t & 0 \\ 0 & e^t \end{bmatrix} (\mathbb{I}_2 + \mathbf{N}t) = \begin{bmatrix} e^t & 0 \\ 0 & e^t \end{bmatrix} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} e^t & te^t \\ 0 & e^t \end{bmatrix} \end{aligned} \tag{3.83}$$

■

Example 3.5. *Compute $e^{\mathbf{\Lambda}t}$ where matrix $\mathbf{\Lambda}$ reads as follows:*

$$\mathbf{\Lambda} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.84}$$

Matrix $\mathbf{\Lambda}$ is block diagonal:

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{J} \end{bmatrix} \tag{3.85}$$

where:

$$\begin{cases} \mathbf{D} = 2 \\ \mathbf{J} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \end{cases} \tag{3.86}$$

As far as $\mathbf{\Lambda}$ *is block diagonal,* $e^{\mathbf{\Lambda}t}$ *is computed as follows:*

$$e^{\mathbf{\Lambda}t} = \begin{bmatrix} e^{\mathbf{D}t} & \mathbf{0} \\ \mathbf{0} & e^{\mathbf{J}t} \end{bmatrix} \tag{3.87}$$

Using the result of the preceding example, we get:

$$\begin{cases} e^{\mathbf{D}t} = e^{2t} \\ e^{\mathbf{J}t} = \begin{bmatrix} e^t & te^t \\ 0 & e^t \end{bmatrix} \end{cases} \tag{3.88}$$

Thus:

$$e^{\mathbf{\Lambda}t} = \begin{bmatrix} e^{2t} & 0 & 0 \\ 0 & e^t & te^t \\ 0 & 0 & e^t \end{bmatrix} \tag{3.89}$$

■

3.5.5 Dunford (or Jordan–Chevalley) decomposition

Dunford (or Jordan–Chevalley) decomposition (Nelson Dunford was an American mathematician (1906 - 1986)) states that for any matrix $\mathbf{A}$ with real coefficients there exists a unique pair of a diagonalizable matrix $\mathbf{D}$ and a nilpotent matrix $\mathbf{N}$ which commutes with $\mathbf{D}$ such that:

$$\mathbf{A} = \mathbf{D} + \mathbf{N} \text{ and } \mathbf{DN} = \mathbf{ND} \tag{3.90}$$

Note that a necessary and sufficient condition for an endomorphism $\mathbf{A}$ in a vector space of dimension n to be nilpotent is that its characteristic polynomial $\det(\lambda\mathbb{I} - \mathbf{A})$ be equal to λ^n.

Taking the exponential of the Dunford decomposition we get the following result since matrices $\mathbf{N}$ and $\mathbf{D}$ commute:

$$e^{\mathbf{A}t} = e^{(\mathbf{D}+\mathbf{N})t} = e^{\mathbf{D}t}e^{\mathbf{N}t} \tag{3.91}$$

Denoting k the index of nilpotency of matrix $\mathbf{N}$ the previous expression becomes:

$$e^{\mathbf{A}t} = e^{\mathbf{D}t}\left(\mathbb{I} + \sum_{i=1}^{k-1} \frac{(\mathbf{N}t)^i}{i!}\right) \tag{3.92}$$

Nilpotent matrix $\mathbf{N}$ is null when the dimension of the eigenspace generated by each eigenvalue of $\mathbf{A}$ is equal to the multiplicity of the eigenvalue or equivalently when all the roots of the minimal polynomial of $\mathbf{A}$ have multiplicity 1.

Note that we can write the following relation:

$$e^{\mathbf{D}t}e^{\mathbf{N}t} = e^{\mathbf{D}t} + e^{\mathbf{D}t}\left(e^{\mathbf{N}t} - \mathbb{I}\right) \tag{3.93}$$

As far as matrix $\mathbf{N}$ is nilpotent, its spectrum is $\{0\}$. Therefore the spectrum of $e^{\mathbf{N}t}$ is $\{1\}$. We therefore deduce that matrix $e^{\mathbf{D}t}\left(e^{\mathbf{N}t}-\mathbb{I}\right)$ is also a nilpotent matrix.

Let $\chi_A(\lambda) := \det(\lambda\mathbb{I}-\mathbf{A})$ be the characteristic polynomial of matrix $\mathbf{A}$. Then Dunford decomposition is obtained from the partial fraction expansion of the inverse of its characteristic polynomial $\chi_A(\lambda)$. Let λ_i be the eigenvalues of $\mathbf{A}$, n_i the multiplicity of λ_i and r the number of distinct eigenvalues of $\mathbf{A}$. Then[2]:

$$\begin{aligned} \chi_A(\lambda) &= \textstyle\prod_{i=1}^{r}(\lambda-\lambda_i)^{n_i} \\ \Rightarrow \tfrac{1}{\chi_A(\lambda)} &= \textstyle\sum_{i=1}^{r}\left(\sum_{k=1}^{n_i}\frac{a_{i,k}}{(\lambda-\lambda_i)^k}\right) = \sum_{i=1}^{r}\left(\sum_{k=1}^{n_i}\frac{a_{i,k}(\lambda-\lambda_i)^{n_i-k}}{(\lambda-\lambda_i)^{n_i}}\right) \end{aligned} \tag{3.94}$$

Setting $U_i(\lambda) := \sum_{k=1}^{n_i} a_{i,k}(\lambda-\lambda_i)^{n_i-k}$ where $U_i(\lambda)$ is a polynomial of degree n_i-1, we get:

$$\frac{1}{\chi_A(\lambda)} = \sum_{i=1}^{r}\frac{U_i(\lambda)}{(\lambda-\lambda_i)^{n_i}} \tag{3.95}$$

Then multiplying by $\chi_A(\lambda) = \prod_{i=1}^{r}(\lambda-\lambda_i)^{n_i}$ it comes:

$$1 = \sum_{i=1}^{r} U_i(\lambda)\prod_{i\neq j}(\lambda-\lambda_i)^{n_i} = \sum_{i=1}^{r} P_i(\lambda) \tag{3.96}$$

Then it can be shown that:

$$\mathbf{D} = \sum_{i=1}^{r}\lambda_i P_i(\mathbf{A}) \text{ and } \mathbf{N} = \mathbf{A}-\mathbf{D} \tag{3.97}$$

Example 3.6. *Let* $\mathbf{A}$ *be the following matrix:*

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix} \tag{3.98}$$

The characteristic polynomial of $\mathbf{A}$ *reads as follows:*

$$\chi_A(\lambda) := \det(\lambda\mathbb{I}-\mathbf{A}) = (\lambda-1)^2(\lambda-2) \tag{3.99}$$

The eigenvalues of $\mathbf{A}$ *are* $\lambda_1 = 1$ *with multiplicity* $n_1 = 2$ *and* $\lambda_2 = 2$ *with multiplicity* $n_2 = 1$. *Let's perform the product:*

$$(\mathbf{A}-\lambda_1\mathbb{I}_3)(\mathbf{A}-\lambda_2\mathbb{I}_3) = \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \neq \mathbf{0} \tag{3.100}$$

Therefore $(\lambda-1)(\lambda-2)$ *is not an annihilating polynomial of* $\mathbf{A}$. *The minimal polynomial of* $\mathbf{A}$ *is therefore the characteristic polynomial* $(\lambda-1)^2(\lambda-2)$. *As the minimal polynomial does not have all its simple roots the matrix* $\mathbf{A}$ *is not*

[2] Alaeddine Ben Rhouma. Autour de la décomposition de Dunford réelle ou complexe. Théorie spectrale et méthodes effectives. 2013. hal-00844141

diagonalizable. We will use the Dunford decomposition to write $\mathbf{A}$ *in the form* $\mathbf{A} = \mathbf{D} + \mathbf{N}$ *where* $\mathbf{D}$ *is diagonalizable and where* $\mathbf{N}$ *is nilpotent and commutes with* $\mathbf{D}$.

First we compute the partial fraction expansion of $\frac{1}{\chi_A(\lambda)}$:

$$\frac{1}{\chi_A(\lambda)} = -\frac{1}{\lambda - 1} - \frac{1}{(\lambda-1)^2} + \frac{1}{(\lambda-2)} = \frac{U_1(\lambda)}{(\lambda-1)^2} + \frac{U_2(\lambda)}{\lambda - 2} \tag{3.101}$$

where:

$$\begin{cases} U_1(\lambda) = -\lambda \\ U_2(\lambda) = 1 \end{cases} \tag{3.102}$$

Then multiplying by $\chi_A(\lambda) = (\lambda-1)^2(\lambda-2)$ *we get:*

$$1 = \sum_{i=1}^{2} U_i(\lambda) \prod_{i \neq j} (\lambda - \lambda_i)^{n_i} = P_1(\lambda) + P_2(\lambda) \tag{3.103}$$

where:

$$\begin{cases} P_1(\lambda) = -\lambda(\lambda - 2) \\ P_2(\lambda) = (\lambda - 1)^2 \end{cases} \tag{3.104}$$

We deduce that:

$$\begin{aligned} \mathbf{D} &= \lambda_1 P_1(\mathbf{A}) + \lambda_2 P_2(\mathbf{A}) \\ &= -\mathbf{A}(\mathbf{A} - 2\mathbb{I}_3) + 2(\mathbf{A} - \mathbb{I}_3)^2 \\ &= \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \end{aligned} \tag{3.105}$$

Let's compute the product $(\mathbf{D} - \lambda_1 \mathbb{I}_3)(\mathbf{D} - \lambda_2 \mathbb{I}_3)$. *We get:*

$$(\mathbf{D} - \lambda_1 \mathbb{I}_3)(\mathbf{D} - \lambda_2 \mathbb{I}_3) = \mathbf{0} \tag{3.106}$$

Consequently $(\lambda - 1)(\lambda - 2)$ *is an annihilating polynomial of* $\mathbf{D}$. *As all the roots of the annihilating polynomial of* $\mathbf{D}$ *have multiplicity* 1, *matrix* $\mathbf{D}$ *is diagonalizable. Then nilpotent matrix* $\mathbf{N}$ *is obtained as follows:*

$$\mathbf{N} = \mathbf{A} - \mathbf{D} = \begin{bmatrix} -1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.107}$$

We can check that $\mathbf{N}$ *commutes with* $\mathbf{D}$:

$$\mathbf{ND} = \mathbf{DN} = \begin{bmatrix} -1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.108}$$

Consequently taking into account that $\mathbf{N}^2 = \mathbf{0}$ *we finally get:*

$$e^{\mathbf{A}t} = e^{(\mathbf{D}+\mathbf{N})t} = e^{\mathbf{D}t} e^{\mathbf{N}t} = e^{\mathbf{D}t} (\mathbb{I}_3 + \mathbf{N}t) \tag{3.109}$$

■

3.5.6 Computation of $e^{\mathbf{A}t}$ using the Laplace transform

Computation of $e^{\mathbf{A}t}$ can be done using the Laplace transform. Denoting by $\mathcal{L}$ the Laplace transform and by $\mathcal{L}^{-1}$ the inverse Laplace transform we have the following property which extends to the vector case a well-known result in the scalar case:

$$\mathcal{L}\left[e^{\mathbf{A}t}\right] = (s\mathbb{I} - \mathbf{A})^{-1} \Leftrightarrow e^{\mathbf{A}t} = \mathcal{L}^{-1}\left[(s\mathbb{I} - \mathbf{A})^{-1}\right] \ \forall\, t \geq 0 \tag{3.110}$$

Matrix $(s\mathbb{I} - \mathbf{A})^{-1}$ is called the resolvent of matrix $\mathbf{A}$. It can be computed using the Faddeev-Leverrier's method for example (see section 1.6).

The inverse Laplace transform is taken for each term of the resolvent of matrix $\mathbf{A}$. We recall that the inverse Laplace transform of a strictly proper rational function $F(s)$ (i.e. the degree of the denominator is strictly greater than the degree of the denominator) can be obtained using the Mellin-Fourier integral.

The Mellin-Fourier integral reads:

$$g(t) = \mathcal{L}^{-1}\left[F(s)\right] = \sum_{\text{poles of } F(s)} \text{Res}\left[F(s)e^{st}\right] \ \forall t \geq 0 \tag{3.111}$$

The residue $\text{Res}\left[F(s)e^{st}\right]$ shall be computed around each pole of $F(s)$. Assuming that λ_k is a pole of multiplicity n_k then the residue of $F(s)$ around pole λ_k is given by:

$$\text{Res}_{s=\lambda_k}\left[F(s)e^{st}\right] = \frac{1}{(n_k-1)!}\frac{d^{n_k-1}}{ds^{n_k-1}}(s-\lambda_k)^{n_k}F(s)e^{st}\Big|_{s=\lambda_k} \tag{3.112}$$

Alternatively if resolvent of matrix $\mathbf{A}$ is decomposed as $\frac{\mathbf{N}(s)}{\Psi(s)}$ where degree of polynomial matrix $\mathbf{N}(s)$ is strictly lower than degree of polynomial $\Psi(s) = \prod_k (s-\lambda_k)^{n_k}$ then the use of Mellin-Fourier integral leads to the following expression of $e^{\mathbf{A}t}$:

$$\begin{aligned} &(s\mathbb{I} - \mathbf{A})^{-1} = \tfrac{\mathbf{N}(s)}{\Psi(s)} = \tfrac{\mathbf{N}(s)}{\prod_k (s-\lambda_k)^{n_k}} \\ &\Rightarrow e^{\mathbf{A}t} = \textstyle\sum_k \tfrac{1}{(n_k-1)!}\tfrac{d^{n_k-1}}{ds^{n_k-1}}(s-\lambda_k)^{n_k}\tfrac{\mathbf{N}(s)}{\Psi(s)}e^{st}\Big|_{s=\lambda_k} \end{aligned} \tag{3.113}$$

Example 3.7. *Compute $e^{\mathbf{A}t}$ where:*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{3.114}$$

Here $n = 2$ and the Faddeev-Leverrier's method (see section 1.6) reads:

$$\begin{cases} \mathbf{F}_0 = \mathbb{I} \\ d_1 = \text{tr}\left(\mathbf{A}\mathbf{F}_0\right) = \text{tr}\left(\mathbf{A}\right) = 0 \ \textit{and}\ \mathbf{F}_1 = \mathbf{A}\mathbf{F}_0 - d_1\mathbb{I} = \mathbf{A} \\ d_2 = \frac{1}{2}\text{tr}\left(\mathbf{A}\mathbf{F}_1\right) = \frac{1}{2}\text{tr}\left(\mathbf{A}^2\right) = 0 \end{cases} \tag{3.115}$$
$$\textit{and}\ \det\left(s\mathbb{I} - \mathbf{A}\right) = s^2 - d_1 s - d_2 = s^2$$

Thus:

$$(s\mathbb{I}-\mathbf{A})^{-1}=\frac{1}{\det(s\mathbb{I}-\mathbf{A})}(\mathbf{F}_0 s+\mathbf{F}_1)=\frac{1}{s^2}\begin{bmatrix} s & 1 \\ 0 & s \end{bmatrix}=\begin{bmatrix} \frac{1}{s} & \frac{1}{s^2} \\ 0 & \frac{1}{s} \end{bmatrix} \tag{3.116}$$

Then we have to apply the Mellin-Fourier integral as well as the residue theorem on each term of the resolvent matrix of $\mathbf{A}$*:*

$$\begin{aligned} \mathcal{L}^{-1}\left[\tfrac{1}{s}\right] &= \operatorname{Res}\left[\tfrac{1}{s}e^{st}\right]_{s=0}=\tfrac{1}{(1-1)!}\tfrac{d^{1-1}}{ds^{1-1}}s\tfrac{1}{s}e^{st}\Big|_{s=0}=e^{st}\big|_{s=0}=1 \\ \mathcal{L}^{-1}\left[\tfrac{1}{s^2}\right] &= \operatorname{Res}\left[\tfrac{1}{s^2}e^{st}\right]_{s=0}=\tfrac{1}{(2-1)!}\tfrac{d^{2-1}}{ds^{2-1}}s^2\tfrac{1}{s^2}e^{st}\Big|_{s=0}=te^{st}\big|_{s=0}=t \end{aligned} \tag{3.117}$$

We finally get:

$$\exp\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}t\right)=\mathcal{L}^{-1}\begin{bmatrix} \frac{1}{s} & \frac{1}{s^2} \\ 0 & \frac{1}{s} \end{bmatrix}=\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \tag{3.118}$$

Alternatively resolvent of matrix $\mathbf{A}$ *can be decomposed as* $\frac{\mathbf{N}(s)}{\Psi(s)}$*. Indeed we have seen that:*

$$(s\mathbb{I}-\mathbf{A})^{-1}=\frac{1}{s^2}\begin{bmatrix} s & 1 \\ 0 & s \end{bmatrix} \tag{3.119}$$

The use of Mellin-Fourier integral leads to the following expression of $e^{\mathbf{A}t}$*:*

$$\begin{aligned} (s\mathbb{I}-\mathbf{A})^{-1} &= \tfrac{\mathbf{N}(s)}{\Psi(s)} \\ \Rightarrow e^{\mathbf{A}t} &= \textstyle\sum_k \tfrac{1}{(n_k-1)!}\tfrac{d^{n_k-1}}{ds^{n_k-1}}(s-\lambda_k)^{n_k}\tfrac{\mathbf{N}(s)}{\Psi(s)}e^{st}\Big|_{s=\lambda_k} \\ &= \tfrac{1}{(2-1)!}\tfrac{d^{2-1}}{ds^{2-1}}s^2\tfrac{1}{s^2}\begin{bmatrix} s & 1 \\ 0 & s \end{bmatrix}e^{st}\Big|_{s=0} \\ &= \tfrac{d}{ds}\begin{bmatrix} s & 1 \\ 0 & s \end{bmatrix}e^{st}\Big|_{s=0} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}e^{st}\Big|_{s=0}+\begin{bmatrix} s & 1 \\ 0 & s \end{bmatrix}t\,e^{st}\Big|_{s=0} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}+\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}t \\ &= \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \end{aligned} \tag{3.120}$$

■

Example 3.8. *Compute* $e^{\mathbf{A}t}$ *where:*

$$\mathbf{A}=\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \tag{3.121}$$

We have:

$$(s\mathbb{I}-\mathbf{A})^{-1}=\left(\begin{bmatrix} s-1 & -1 \\ 0 & s \end{bmatrix}\right)^{-1}=\frac{1}{s(s-1)}\begin{bmatrix} s & 1 \\ 0 & s-1 \end{bmatrix} \tag{3.122}$$

Thus $(s\mathbb{I}-\mathbf{A})^{-1}=\frac{\mathbf{N}(s)}{\Psi(s)}$ *where* $\Psi(s)=s(s-1)$ *has two roots,* $\lambda_1=0$ *and* $\lambda_2=1$*, each of multiplicity* 1*:* $n_1=n_2=1$*.*

The use of Mellin-Fourier integral leads to the following expression of $e^{\mathbf{A}t}$:

$$\begin{aligned}
(s\mathbb{I}-\mathbf{A})^{-1} &= \tfrac{\mathbf{N}(s)}{\Psi(s)} \\
\Rightarrow e^{\mathbf{A}t} &= \textstyle\sum_k \frac{1}{(n_k-1)!}\frac{d^{n_k-1}}{ds^{n_k-1}}(s-\lambda_k)^{n_k}\frac{\mathbf{N}(s)}{\Psi(s)}e^{st}\Big|_{s=\lambda_k} \\
&= \tfrac{1}{(1-1)!}\tfrac{d^{1-1}}{ds^{1-1}}s\tfrac{1}{s(s-1)}\begin{bmatrix} s & 1 \\ 0 & s-1 \end{bmatrix}e^{st}\Big|_{s=0} \\
&\quad + \tfrac{1}{(1-1)!}\tfrac{d^{1-1}}{ds^{1-1}}(s-1)\tfrac{1}{s(s-1)}\begin{bmatrix} s & 1 \\ 0 & s-1 \end{bmatrix}e^{st}\Big|_{s=1} \\
&= \tfrac{1}{s-1}\begin{bmatrix} s & 1 \\ 0 & s-1 \end{bmatrix}e^{st}\Big|_{s=0} + \tfrac{1}{s}\begin{bmatrix} s & 1 \\ 0 & s-1 \end{bmatrix}e^{st}\Big|_{s=1} \\
&= \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}e^{t} \\
&= \begin{bmatrix} e^t & e^t-1 \\ 0 & 1 \end{bmatrix}
\end{aligned} \tag{3.123}$$

■

Example 3.9. *Compute $e^{\mathbf{A}t}$ where:*

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & -5 \end{bmatrix} \tag{3.124}$$

Here $n = 2$ and the Faddeev-Leverrier's method (see section 1.6) reads:

$$\begin{cases} \mathbf{F}_0 = \mathbb{I} \\ d_1 = \operatorname{tr}(\mathbf{A}\mathbf{F}_0) = -4 \text{ and } \mathbf{F}_1 = \mathbf{A}\mathbf{F}_0 - d_1\mathbb{I} = \begin{bmatrix} 5 & 2 \\ 0 & -1 \end{bmatrix} \\ d_2 = \frac{1}{2}\operatorname{tr}(\mathbf{A}\mathbf{F}_1) = \frac{1}{2}\operatorname{tr}\left(\begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}\right) = 5 \end{cases} \tag{3.125}$$

and $\det(s\mathbb{I}-\mathbf{A}) = s^2 - d_1 s - d_2 = s^2 + 4s - 5 = (s-1)(s+5)$

Thus:

$$\begin{aligned}
&(s\mathbb{I}-\mathbf{A})^{-1} = \tfrac{1}{\det(s\mathbb{I}-\mathbf{A})}(\mathbf{F}_0 s + \mathbf{F}_1) = \tfrac{1}{(s-1)(s+5)}\begin{bmatrix} s+5 & 2 \\ 0 & s-1 \end{bmatrix} \\
&\Leftrightarrow (s\mathbb{I}-\mathbf{A})^{-1} = \begin{bmatrix} \frac{1}{s-1} & \frac{2}{(s-1)(s+5)} \\ 0 & \frac{1}{s+5} \end{bmatrix}
\end{aligned} \tag{3.126}$$

Then we have to apply the Mellin-Fourier integral as well as the residue theorem on each term of the resolvent matrix of $\mathbf{A}$:

$$\begin{cases} \mathcal{L}^{-1}\left[\frac{1}{s-1}\right] = e^t \\ \mathcal{L}^{-1}\left[\frac{1}{s+5}\right] = e^{-5t} \\ \mathcal{L}^{-1}\left[\frac{2}{(s-1)(s+5)}\right] = 2\left(\frac{1}{6}e^t - \frac{1}{6}e^{-5t}\right) = \frac{1}{3}e^t - \frac{1}{3}e^{-5t} \end{cases} \tag{3.127}$$

We finally get:

$$\begin{aligned}
\exp\left(\begin{bmatrix} 1 & 2 \\ 0 & -5 \end{bmatrix}t\right) &= \mathcal{L}^{-1}\begin{bmatrix} \frac{1}{s-1} & \frac{2}{(s-1)(s+5)} \\ 0 & \frac{1}{s+5} \end{bmatrix} \\
&= \begin{bmatrix} e^t & \frac{1}{3}e^t - \frac{1}{3}e^{-5t} \\ 0 & e^{-5t} \end{bmatrix}
\end{aligned} \tag{3.128}$$

Alternatively resolvent of matrix $\mathbf{A}$ *can be decomposed as* $\frac{\mathbf{N}(s)}{\Psi(s)}$. *Indeed we have seen that:*

$$(s\mathbb{I}-\mathbf{A})^{-1}=\frac{1}{(s-1)(s+5)}\begin{bmatrix} s+5 & 2 \\ 0 & s-1 \end{bmatrix} \tag{3.129}$$

The use of Mellin-Fourier integral leads to the following expression of $e^{\mathbf{A}t}$*:*

$$\begin{aligned}
(s\mathbb{I}-\mathbf{A})^{-1} &= \tfrac{\mathbf{N}(s)}{\Psi(s)} \\
\Rightarrow e^{\mathbf{A}t} &= \textstyle\sum_k \frac{1}{(n_k-1)!}\frac{d^{n_k-1}}{ds^{n_k-1}}(s-\lambda_k)^{n_k}\frac{\mathbf{N}(s)}{\Psi(s)}e^{st}\Big|_{s=\lambda_k} \\
&= (s-1)\tfrac{1}{(s-1)(s+5)}\begin{bmatrix} s+5 & 2 \\ 0 & s-1 \end{bmatrix}e^{st}\Big|_{s=1} \\
&\quad + (s+5)\tfrac{1}{(s-1)(s+5)}\begin{bmatrix} s+5 & 2 \\ 0 & s-1 \end{bmatrix}e^{st}\Big|_{s=-5} \\
&= \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & 0 \end{bmatrix}e^{st}\Big|_{s=1} + \begin{bmatrix} 0 & -\frac{1}{3} \\ 0 & 1 \end{bmatrix}e^{st}\Big|_{s=-5} \\
&= \begin{bmatrix} e^t & \frac{1}{3}e^t \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -\frac{1}{3}e^{-5t} \\ 0 & e^{-5t} \end{bmatrix} \\
&= \begin{bmatrix} e^t & \frac{1}{3}e^t-\frac{1}{3}e^{-5t} \\ 0 & e^{-5t} \end{bmatrix}
\end{aligned} \tag{3.130}$$

■

Example 3.10. *Compute* $e^{\mathbf{A}t}$ *where:*

$$\mathbf{A}=\begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \tag{3.131}$$

From the Faddeev-Leverrier's method (see section 1.6) we get:

$$(s\mathbb{I}-\mathbf{A})^{-1}=\begin{bmatrix} \frac{1}{s-2} & \frac{-1}{(s-1)(s-2)} & 0 \\ 0 & \frac{1}{s-1} & 0 \\ \frac{1}{(s-1)(s-2)} & \frac{-1}{(s-1)(s-2)} & \frac{1}{s-1} \end{bmatrix} \tag{3.132}$$

Then we have to apply the Mellin-Fourier integral as well as the residue theorem on each term of the resolvent matrix of $\mathbf{A}$*:*

$$\begin{cases} \mathcal{L}^{-1}\left[\frac{1}{s-2}\right]=e^{2t} \\ \mathcal{L}^{-1}\left[\frac{1}{s-1}\right]=e^{t} \\ \mathcal{L}^{-1}\left[\frac{1}{(s-1)(s-2)}\right]=e^{2t}-e^{t} \end{cases} \tag{3.133}$$

We finally get:

$$e^{\mathbf{A}t}=\begin{bmatrix} e^{2t} & e^t-e^{2t} & 0 \\ 0 & e^t & 0 \\ e^{2t}-e^t & e^t-e^{2t} & e^t \end{bmatrix} \tag{3.134}$$

■

3.6 Stability

There are two different definitions of stability: internal stability and input-output stability:

- A linear time-invariant system is internally stable if its the zero-input state $e^{\mathbf{A}t}\underline{x}_0$ moves towards zero for any initial state $\underline{x}_0$. This is a basic requirement for feedback system. Indeed interconnected systems may be subject to some nonzero initial conditions and some (possibly small) errors and it is not acceptable from a practical point of view that such nonzero initial conditions lead to unbounded signals in the closed-loop system.

- A linear time-invariant system is input-output stable if its zero-state output is bounded for all bounded inputs. This type of stability is also called Bounded-Input Bounded-Output (BIBO) stability. Bounded-Input Bounded-Output (BIBO) stability guarantees that all signals in a system are bounded provided that the input signals (at any locations) are bounded.

We have seen in (3.13) that the output response $\underline{y}(t)$ of a linear time-invariant system is the following:

$$\underline{y}(t) = \mathbf{C}e^{\mathbf{A}t}\underline{x}_0 + \int_0^t \mathbf{C}e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau + \mathbf{D}\underline{u}(t) \tag{3.135}$$

Assuming that matrix $\mathbf{A}$ is diagonalizable, we have seen in (3.52) that $e^{\mathbf{A}t}$ can be obtained using the following relation :

$$e^{\mathbf{A}t} = \mathbf{P}e^{\mathbf{\Lambda}t}\mathbf{P}^{-1} = \sum_{i=1}^{n} \underline{v}_i\underline{w}_i^T e^{\lambda_i t} \tag{3.136}$$

Thus;

- The zero-input state, which is obtained when $\underline{u}(t) = \underline{0}$, has the following expression:

$$\underline{x}(t) = e^{\mathbf{A}t}\underline{x}_0 = \sum_{i=1}^{n} \underline{v}_i\underline{w}_i^T e^{\lambda_i t}\underline{x}_0 \tag{3.137}$$

  Consequently the zero-input state moves towards zero for any initial state $\underline{x}_0$ as soon as all the eigenvalues λ_i of matrix $\mathbf{A}$ are situated in the open left-half plane (they have strictly negative real part). This means that a linear time-invariant system is internally stable when all the eigenvalues λ_i of matrix $\mathbf{A}$ are situated in the open left-half plane (i.e. they have strictly negative real part).

  The result which have been shown assuming that matrix $\mathbf{A}$ is diagonalizable can be extended to the general case where matrix $\mathbf{A}$ is not diagonalizable; in that situation this is the Jordan form of $\mathbf{A}$ which leads to the same result concerning internal stability.

– The zero-state output, which is obtained when $\underline{x}_0 = \underline{0}$, has the following expression:

$$\begin{aligned} \underline{y}(t) &= \int_0^t \mathbf{C}e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau + \mathbf{D}\underline{u}(t) \\ &= \left(\mathbf{C}e^{\mathbf{A}t}\mathbf{B} + \mathbf{D}\delta(t)\right) * \underline{u}(t) \end{aligned} \tag{3.138}$$

It can be shown that the zero-state output is bounded if and only all the poles of each term of the transfer function $\mathbf{F}(s)$ are situated in the open left-half plane (i.e. they have strictly negative real part):

$$\mathbf{F}(s) = \mathcal{L}\left[\mathbf{C}e^{\mathbf{A}t}\mathbf{B} + \mathbf{D}\delta(t)\right] = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} \tag{3.139}$$

The two types of stability are related. Indeed:

– If a linear time invariant system is internally stable it is also input-output (or BIBO) stable because all the poles of the transfer function $\mathbf{F}(s)$ are eigenvalues of matrix $\mathbf{A}$;

– Nevertheless the converse is not true since matrix $\mathbf{A}$ could have unstable hidden modes which do not appear in the poles of $\mathbf{F}(s)$. Indeed there may be pole-zero cancellation while computing $\mathbf{F}(s)$. Thus a system may be BIBO stable even when some eigenvalues of $\mathbf{A}$ do not have negative real part.

Example 3.11. *Let's consider the following realization:*

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) = \left(\begin{array}{cc|c} -1 & 10 & -2 \\ 0 & 1 & 0 \\ \hline -2 & 3 & -2 \end{array}\right) \tag{3.140}$$

Matrix $\mathbf{A}$ *has a stable mode, which is* -1*, and an unstable mode, which is* 1*. Thus the system is not internally stable.*

When computing the transfer function of the system we can observe a pole / zero cancellation of the unstable mode:

$$\begin{aligned} F(s) &= \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} \\ &= \begin{bmatrix} -2 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{s+1} & \frac{10}{s^2-1} \\ 0 & \frac{1}{s-1} \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix} - 2 \\ &= \frac{4}{s+1} - 2 \\ &= \frac{-2s+2}{s+1} \end{aligned} \tag{3.141}$$

The pole of the transfer function $F(s)$ *is* -1*. Thus the system is BIBO stable but not internally stable.*

■

3.7 Controllability

3.7.1 Definition

Let's consider the state trajectory $\underline{x}(t)$ of a linear time-invariant system:

$$\underline{x}(t) = e^{\mathbf{A}t}\underline{x}_0 + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau \tag{3.142}$$

Where $\mathbf{A}$ is a $n \times n$ real matrix and $\mathbf{B}$ is a $n \times m$ real matrix.

Controllability answers the question whether it is possible to control the state vector $\underline{x}(t)$ through an appropriate choice of the input signal $\underline{u}(t)$.

More precisely an initial state $\underline{x}_0$ is said controllable if and only if there exists an input signal $\underline{u}(t)$ which is able to move the state vector $\underline{x}(t)$ from an initial state $\underline{x}(0) = \underline{x}_0$ at $t = 0$ to the origin $\underline{x}(t_f) = \underline{0}$ in a finite time t_f. We said that a system is controllable when any arbitrary initial state $\underline{x}_0 \in \mathbb{R}^n$ is controllable [3].

If the system is controllable then the input signal $\underline{u}(t)$ which is able to move the state vector $\underline{x}(t)$ from an initial state $\underline{x}(0) = \underline{x}_0$ at $t = 0$ to the origin $\underline{x}(t_f) = \underline{0}$ in a finite time t_f reads[4]:

$$\underline{u}(t) = -\mathbf{B}^T e^{\mathbf{A}^T(t_f - t)}\, \mathbf{W}_c^{-1}(t_f) e^{\mathbf{A}t_f}\, \underline{x}_0 \tag{3.143}$$

Where $\mathbf{W}_c(t_f)$ is a symmetric matrix defined as follows:

$$\mathbf{W}_c(t_f) = \int_0^{t_f} e^{\mathbf{A}\tau} \mathbf{B}\mathbf{B}^T e^{\mathbf{A}^T\tau} d\tau \tag{3.144}$$

Indeed when $\underline{x}(t_f)$ is computed with this control we get $\underline{x}(t_f) = \underline{0}$:

$$\begin{aligned}
\underline{x}(t_f) &= e^{\mathbf{A}t_f}\underline{x}_0 + \textstyle\int_0^{t_f} e^{\mathbf{A}(t_f-\tau)}\mathbf{B}\underline{u}(\tau)d\tau \\
&= e^{\mathbf{A}t_f}\underline{x}_0 - \textstyle\int_0^{t_f} e^{\mathbf{A}(t_f-\tau)}\mathbf{B}\mathbf{B}^T e^{\mathbf{A}^T(t_f-\tau)}\, \mathbf{W}_c^{-1}(t_f) e^{\mathbf{A}t_f}\, \underline{x}_0\, d\tau \\
&= e^{\mathbf{A}t_f}\underline{x}_0 + \left(\textstyle\int_{t_f}^{0} e^{\mathbf{A}v}\mathbf{B}\mathbf{B}^T e^{\mathbf{A}^T v}\, dv\right) \mathbf{W}_c^{-1}(t_f) e^{\mathbf{A}t_f}\, \underline{x}_0 \\
&= e^{\mathbf{A}t_f}\underline{x}_0 - \left(\textstyle\int_0^{t_f} e^{\mathbf{A}\tau}\mathbf{B}\mathbf{B}^T e^{\mathbf{A}^T\tau}\, d\tau\right) \mathbf{W}_c^{-1}(t_f) e^{\mathbf{A}t_f}\, \underline{x}_0 \\
&= e^{\mathbf{A}t_f}\underline{x}_0 - \mathbf{W}_c(t_f)\mathbf{W}_c^{-1}(t_f) e^{\mathbf{A}t_f}\, \underline{x}_0 \\
&= \underline{0}
\end{aligned} \tag{3.145}$$

More generally one can verify that a particular input which achieves $\underline{x}(t_f) = \underline{x}_f$ is given by[4]:

$$\underline{u}(t) = -\mathbf{B}^T e^{\mathbf{A}^T(t_f - t)}\, \mathbf{W}_c^{-1}(t_f) \left(e^{\mathbf{A}t_f}\, \underline{x}_0 - \underline{x}_f\right) \tag{3.146}$$

Consequently a system is controllable if and only if symmetric matrix $\mathbf{W}_c(t_f)$ is nonsingular for any $t_f > 0$. Furthermore $\mathbf{W}_c(t)$ is the solution of the following differential equation:

$$\mathbf{A}\mathbf{W}_c(t) + \mathbf{W}_c(t)\mathbf{A}^T + \mathbf{B}\mathbf{B}^T = \frac{d}{dt}\mathbf{W}_c(t) \tag{3.147}$$

If all eigenvalues of matrix $\mathbf{A}$ have negative real parts ($\mathbf{A}$ is said to be *Hurwitz*) then $\mathbf{W}_c(t)$ tends towards a constant matrix as $t \to \infty$. The controllability Gramian $\mathbf{W}_c$ is the following positive definite symmetric matrix:

$$\mathbf{W}_c = \int_0^{\infty} e^{\mathbf{A}\tau} \mathbf{B}\mathbf{B}^T e^{\mathbf{A}^T\tau} d\tau \tag{3.148}$$

It can be shown that $\mathbf{W}_c$ is the unique solution of the following Lyapunov equation:

$$\mathbf{A}\mathbf{W}_c + \mathbf{W}_c\mathbf{A}^T + \mathbf{B}\mathbf{B}^T = \mathbf{0} \tag{3.149}$$

[3] https://en.wikibooks.org/wiki/Control_Systems/Controllability_and_Observability

[4] S. Skogestad and I. Postlethwaite: Multivariable Feedback Control Analysis and design, Wiley, 1996; 2005

3.7.2 Kalman's controllability rank condition

Let $\mathbf{Q}_c$ be the controllability matrix. Matrix $\mathbf{Q}_c$ is defined as follows:

$$\mathbf{Q}_c = \left[\begin{array}{cccc} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{array} \right] \tag{3.150}$$

It can be shown that a linear system is controllable if and only if the rank of the controllability matrix $\mathbf{Q}_c$ is equal to n. This is the Kalman's controllability rank condition.

The sketch of the demonstration is the following:

- First we recall that the expression of the state vector $\underline{x}(t)$ at time $t = t_f$ which solves the state equation (3.1) is:

$$\underline{x}(t_f) = e^{\mathbf{A}t_f}\underline{x}_0 + \int_0^{t_f} e^{\mathbf{A}(t_f-\tau)}\mathbf{B}\underline{u}(\tau)d\tau \tag{3.151}$$

  As far as $\underline{x}_0$, t_f and $\underline{x}(t_f)$ are assumed to be known we rewrite the preceding equation as follows:

$$e^{-\mathbf{A}t_f}\underline{x}(t_f) - \underline{x}_0 = \int_0^{t_f} e^{-\mathbf{A}\tau}\mathbf{B}\underline{u}(\tau)d\tau \tag{3.152}$$

- To continue the sketch of the proof we need the Cayley–Hamilton theorem. Let $\chi_A(s)$ be the characteristic polynomial of the $n \times n$ matrix $\mathbf{A}$. We write the characteristic polynomial $\chi_A(s)$ of matrix $\mathbf{A}$ as follows:

$$\chi_A(s) := \det(s\mathbb{I} - \mathbf{A}) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \tag{3.153}$$

  The Cayley–Hamilton theorem states that substituting matrix $\mathbf{A}$ for s in the characteristic polynomial $\chi_A(s)$ of matrix $\mathbf{A}$ results in the zero matrix [5]:

$$\chi_A(\mathbf{A}) = \mathbf{0} = \mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \cdots + a_1\mathbf{A} + a_0\mathbb{I} \tag{3.154}$$

  From the preceding equation it is clear that we can express $\mathbf{A}^n$ as a function of $\mathbf{A}^k$ where $0 \leq k \leq n-1$:

$$\mathbf{A}^n = -a_{n-1}\mathbf{A}^{n-1} - \cdots - a_1\mathbf{A} - a_0\mathbb{I} \tag{3.155}$$

  More generally this relation allows to replace a term of the form $\mathbf{A}^m$ where $m \geq n$ by a linear combination of $\mathbf{A}^k$ where $0 \leq k \leq n-1$. When we use this property to replace the terms $\mathbf{A}^m$ where $m \geq n$ in the series expansion of $e^{\mathbf{A}t}$ we get the following relation:

$$\left\{ \begin{array}{l} e^{\mathbf{A}t} = \sum_{k=0}^{\infty} \frac{(\mathbf{A}t)^k}{k!} = \sum_{k=0}^{n-1} \frac{\mathbf{A}^k t^k}{k!} + \sum_{k=n}^{\infty} \frac{\mathbf{A}^k t^k}{k!} \\ \mathbf{A}^n = -a_{n-1}\mathbf{A}^{n-1} - \cdots - a_1\mathbf{A} - a_0\mathbb{I} \\ \mathbf{A}^m = \sum_{i=0}^{n-1} \alpha_i \mathbf{A}^i \quad \forall m \geq n \end{array} \right. \tag{3.156}$$

[5] https://en.wikipedia.org/wiki/Cayley–Hamilton_theorem

Consequently the series expansion of $e^{\mathbf{A}t}$ can be cut so that no power of matrix $\mathbf{A}$ greater that n appears in the series expansion:

$$e^{\mathbf{A}t} = \sum_{k=0}^{n-1} \gamma_k(t)\mathbf{A}^k \tag{3.157}$$

where $\gamma_k(t)$ are functions of time t. As far as $\det(s\mathbb{I} - \mathbf{A})$ is equal to zero when $s = \lambda_i$ is an eigenvalue of matrix $\mathbf{A}$ the preceding matrix relation is also be obtained for all the eigenvalues of matrix $\mathbf{A}$. We obtain the following relation which is satisfied by the functions $\gamma_k(t)$:

$$e^{\lambda_i t} = \sum_{k=0}^{n-1} \gamma_k(t)\lambda_i^k \tag{3.158}$$

– Using (3.157) and the fact that functions $\gamma_k(t)$ are scalar functions (3.152) is rewritten as follows:

$$\begin{aligned} e^{-\mathbf{A}t_f}\underline{x}(t_f) - \underline{x}_0 &= \textstyle\int_0^{t_f} e^{-\mathbf{A}\tau}\mathbf{B}\underline{u}(\tau)d\tau \\ &= \textstyle\int_0^{t_f} \sum_{k=0}^{n-1} \gamma_k(-\tau)\mathbf{A}^k\mathbf{B}\underline{u}(\tau)d\tau \\ &= \textstyle\sum_{k=0}^{n-1} \int_0^{t_f} \gamma_k(-\tau)\mathbf{A}^k\mathbf{B}\underline{u}(\tau)d\tau \\ &= \textstyle\sum_{k=0}^{n-1} \mathbf{A}^k\mathbf{B} \int_0^{t_f} \gamma_k(-\tau)\underline{u}(\tau)d\tau \end{aligned} \tag{3.159}$$

Now let's introduce vector $\underline{w}(t_f)$ whose n components $w_0(t_f), w_1(t_f), \cdots, w_{n-1}(t_f)$ are defined as follows:

$$w_k(t_f) = \int_0^{t_f} \gamma_k(-\tau)\underline{u}(\tau)d\tau \quad \forall\ 0 \le k \le n-1 \tag{3.160}$$

Thus Equation (3.159) reads:

$$\begin{aligned} e^{-\mathbf{A}t_f}\underline{x}(t_f) - \underline{x}_0 &= \textstyle\sum_{k=0}^{n-1} \mathbf{A}^k\mathbf{B} \int_0^{t_f} w_k(t_f) \\ &= \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} w_0(t_f) \\ w_1(t_f) \\ \vdots \\ w_{n-1}(t_f) \end{bmatrix} \\ &= \mathbf{Q}_c\underline{w}(t_f) \end{aligned} \tag{3.161}$$

In order to be able to compute the expression of vector $\underline{w}(t_f)$ and then solving the integral equation in the input signal $\underline{u}(t)$, the controllability matrix $\mathbf{Q}_o$ shall be invertible; consequently the rank of the controllability matrix $\mathbf{Q}_c$ shall be equal to n. Thus we retrieve the Kalman's controllability rank condition.

Example 3.12. *Let's consider the following realization:*

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) = \left(\begin{array}{cc|c} -1 & 10 & -2 \\ 0 & 1 & 0 \\ \hline -2 & 3 & -2 \end{array}\right) \tag{3.162}$$

The controllability matrix $\mathbf{Q}_c$ *reads:*

$$\begin{aligned}\mathbf{Q}_c &= \left[\ \mathbf{B} \mid \mathbf{AB}\ \right] \\ &= \left[\begin{array}{c|c} -2 & \left[\begin{array}{cc} -1 & 10 \\ 0 & 1 \end{array}\right]\left[\begin{array}{c} -2 \\ 0 \end{array}\right] \\ 0 & \end{array}\right] \\ &= \left[\begin{array}{cc} -2 & 2 \\ 0 & 0 \end{array}\right]\end{aligned} \tag{3.163}$$

Consequently $\operatorname{rank}(\mathbf{Q}_c) = 1 \neq 2$. *We conclude that the system in not controllable.*

■

3.7.3 Use of the diagonal form: Gilbert's criteria

Assuming that all eigenvalues are distinct, controllability property can be readily analyzed by inspecting the null rows of the input matrix $\mathbf{B}_m$ as soon as we get the modal (or diagonal) form of the state-space representation.

Indeed we have seen in the the chapter dedicated to the state-space representation that for a single input system matrices $\mathbf{A}_m$ and $\mathbf{B}_m$ of the state equation $\underline{\dot{x}}(t) = \mathbf{A}_m\underline{x}(t) + \mathbf{B}_m\underline{u}(t)$ read as follows assuming that matrix $\mathbf{A}_m$ has n independent eigenvectors :

$$\left\{\begin{array}{l} \mathbf{A}_m = \left[\begin{array}{cccc} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{array}\right] \\ \\ \mathbf{B}_m = \left[\begin{array}{c} b_1 \\ b_2 \\ \vdots \\ b_n \end{array}\right] \end{array}\right. \tag{3.164}$$

Thus in the time domain the diagonal form of the state-space representation $\underline{\dot{x}}(t) = \mathbf{A}_m\underline{x}(t) + \mathbf{B}_m\underline{u}(t)$ reads:

$$\left\{\begin{array}{l} \dot{x}_1(t) = \lambda_1 x_1(t) + b_1 u(t) \\ \dot{x}_2(t) = \lambda_2 x_2(t) + b_2 u(t) \\ \vdots \\ \dot{x}_n(t) = \lambda_n x_n(t) + b_n u(t) \end{array}\right. \tag{3.165}$$

Gilbert's controllability criteria (1963) states that a multi inputs system with distinct eigenvalues is controllable if and only if each row of control matrix $\mathbf{B}_m$ of the diagonal realization (all eigenvalues are distinct) has at least one non zero element.

Indeed if at least one of the b_i's coefficients is zero then the state component $x_i(t)$ is independent of the input signal $u(t)$ and the state is uncontrollable.

Alternatively, if we apply Kalman's controllability rank condition in the modal basis we get:

$$\begin{aligned}
\mathbf{A}_m &= \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \\
\Rightarrow \mathbf{Q}_{cm} &= \begin{bmatrix} \mathbf{B}_m & \mathbf{A}_m\mathbf{B}_m & \cdots & \mathbf{A}_m^{n-1}\mathbf{B}_m \end{bmatrix} = \begin{bmatrix} b_1 & b_1\lambda_1 & \cdots & b_1\lambda_1^{n-1} \\ b_2 & b_2\lambda_2 & \cdots & b_2\lambda_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ b_n & b_n\lambda_n & \cdots & b_n\lambda_n^{n-1} \end{bmatrix}
\end{aligned} \tag{3.166}$$

Then it can be seen that $\mathbf{Q}_{cm}$ is full rank as soon as $b_i \neq 0 \ \forall i = 1, \cdots, n$.

For multi inputs system with m inputs then matrix $\mathbf{B}$ has m columns and the preceding analysis is readily extended to each column of matrix $\mathbf{B}$ assuming that the state-space representation is the diagonal form.

3.7.4 Popov-Belevitch-Hautus (PBH) test

There exists another test for controllability which is called the Popov-Belevitch-Hautus (PBH) test.

Popov-Belevitch-Hautus (PBH) test indicates that a linear system is controllable when the rank of matrix $\begin{bmatrix} \mathbf{A} - \lambda_i\mathbb{I} & \mathbf{B} \end{bmatrix}$ is equal to n for all eigenvalues $\{\lambda_i\}$ of matrix $\mathbf{A}$.

A linear system is *stabilizable* when the rank of matrix $\begin{bmatrix} \mathbf{A} - \lambda_i\mathbb{I} & \mathbf{B} \end{bmatrix}$ is equal to n for all unstable eigenvalues $\{\lambda_i\}$ of matrix $\mathbf{A}$.

Eigenvalues λ_i for which rank of matrix $\begin{bmatrix} \mathbf{A} - \lambda_i\mathbb{I} & \mathbf{B} \end{bmatrix}$ is not equal to n are said uncontrollable.

Equivalently, an eigenvalue λ_i is controllable[6] if all its corresponding left eigenvectors $\underline{w}_i$ (i.e. $\underline{w}_i \neq \underline{0}$ such that $\underline{w}_i^T\mathbf{A} = \lambda_i\underline{w}_i^T$) satisfy $\underline{w}_i^T\mathbf{B} \neq \underline{0}$.

Example 3.13. *Let's consider the following realization:*

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) = \left(\begin{array}{cc|c} -1 & 10 & -2 \\ 0 & 1 & 0 \\ \hline -2 & 3 & -2 \end{array}\right) \tag{3.167}$$

Matrix $\mathbf{A}$ *has two modes,* $\lambda_1 = -1$ *and* $\lambda_2 = 1$.
Let's apply the PBH test for $\lambda_1 = -1$*:*

$$\operatorname{rank}\begin{bmatrix} \mathbf{A} - \lambda_1\mathbb{I} \mid \mathbf{B} \end{bmatrix} = \operatorname{rank}\left[\begin{array}{cc|c} 0 & 10 & -2 \\ 0 & 2 & 0 \end{array}\right] = 2 \tag{3.168}$$

We conclude that the mode $\lambda_1 = -1$ *is controllable.*
Let's apply the PBH test for $\lambda_2 = 1$*:*

$$\operatorname{rank}\begin{bmatrix} \mathbf{A} - \lambda_2\mathbb{I} \mid \mathbf{B} \end{bmatrix} = \operatorname{rank}\left[\begin{array}{cc|c} -2 & 10 & -2 \\ 0 & 0 & 0 \end{array}\right] = 1 \neq 2 \tag{3.169}$$

[6] Antsaklis P. J. and Michel A. N., A Linear Systems Primer, Birkhäuser Boston, Boston,MA, 2007

We conclude that the mode $\lambda_2 = 1$ is not controllable.

■

3.7.5 Uncontrollable mode

Following Bélanger[7] a state $\underline{x}_{\bar{c}} \neq \underline{0}$ is uncontrollable if the zero-state output of the system (i.e. the system's response to the input signal $\underline{u}(t) \neq \underline{0}$ when the initial state $\underline{x}_0$ is zero) is orthogonal to $\underline{x}_{\bar{c}}$ for all final time $t_f > 0$ and all input signal $\underline{u}(t)$. An uncontrollable state $\underline{x}_{\bar{c}} \neq \underline{0}$ satisfies the following equation:

$$\underline{x}_{\bar{c}}^T \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} = \underline{0} \tag{3.170}$$

Equivalently the pair $(\mathbf{A}, \mathbf{B})$ is uncontrollable if and only if there exists an eigenvector $\underline{w}_i$ of $\mathbf{A}^T$ such that $\underline{w}_i^T\mathbf{B} = \underline{0}$. If $\underline{w}_i^T\mathbf{B} = \underline{0}$ then the mode λ_i (i.e. the eigenvalue) corresponding to $\underline{w}_i$ is called an uncontrollable mode.

Indeed if $\underline{w}_i$ is an eigenvector of matrix $\mathbf{A}^T$ corresponding to the mode (i.e. the eigenvalue) λ_i then $\mathbf{A}^T\underline{w}_i = \lambda_i\underline{w}_i \Leftrightarrow \underline{w}_i^T\mathbf{A} = \lambda_i\underline{w}_i^T$. Specializing $\underline{x}_{\bar{c}}$ to $\underline{w}_i$ Equation (3.170) reads:

$$\begin{aligned} \underline{0} &= \underline{w}_i^T \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \\ &= \underline{w}_i^T \begin{bmatrix} \mathbf{B} & \lambda_i\mathbf{B} & \cdots & \lambda_i^{n-1}\mathbf{B} \end{bmatrix} \\ &= \underline{w}_i^T\mathbf{B} \begin{bmatrix} 1 & \lambda_i & \cdots & \lambda_i^{n-1} \end{bmatrix} \end{aligned} \tag{3.171}$$

Coupling $\underline{w}_i^T\mathbf{A} = \lambda_i\underline{w}_i^T$ and $\underline{w}_i^T\mathbf{B} = \underline{0}$ leads to the Popov-Belevitch-Hautus (PBH) test for controllability:

$$\begin{cases} \underline{w}_i^T\mathbf{A} = \lambda_i\underline{w}_i^T \\ \underline{w}_i^T\mathbf{B} = \underline{0} \end{cases} \Leftrightarrow \underline{w}_i^T \begin{bmatrix} \mathbf{A} - \lambda_i\mathbb{I} & \mathbf{B} \end{bmatrix} = \underline{0} \tag{3.172}$$

Example 3.14. *Let's consider the following realization:*

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) = \left(\begin{array}{cc|c} -1 & 10 & -2 \\ 0 & 1 & 0 \\ \hline -2 & 3 & -2 \end{array}\right) \tag{3.173}$$

Matrix $\mathbf{A}$ has two modes, $\lambda_1 = -1$ and $\lambda_2 = 1$. We have seen that the mode $\lambda_2 = 1$ is not controllable. We will check that there no input signal $\underline{u}(t)$ which is able to move towards zero an initial state $\underline{x}_0$ which is set to the value of an eigenvector of $\mathbf{A}^T$ corresponding to the uncontrollable mode $\lambda_2 = 1$.

Let $\underline{w}_2$ be an eigenvector of $\mathbf{A}^T$ corresponding to the uncontrollable mode $\lambda_2 = 1$:

$$\mathbf{A}^T\underline{w}_2 = \lambda_2\underline{w}_2 \Leftrightarrow \begin{bmatrix} -1 & 0 \\ 10 & 1 \end{bmatrix} \underline{w}_2 = \underline{w}_2 \tag{3.174}$$

We expand $\underline{w}_2$ as $\begin{bmatrix} w_{21} \\ w_{22} \end{bmatrix}$ to get:

$$\begin{bmatrix} -1 & 0 \\ 10 & 1 \end{bmatrix} \begin{bmatrix} w_{21} \\ w_{22} \end{bmatrix} = \begin{bmatrix} w_{21} \\ w_{22} \end{bmatrix} \Rightarrow \begin{cases} -w_{21} = w_{21} \\ 10w_{21} + w_{22} = w_{22} \end{cases} \tag{3.175}$$

[7] P. Bélanger, Control EngineeControl Engineering: A Modern Approach, P. Bélanger, Oxford University Press, 2005

We finally get:

$$w_{21} = 0 \Rightarrow \underline{w}_2 = \begin{bmatrix} 0 \\ w_{22} \end{bmatrix} \tag{3.176}$$

Now let's express the state vector $\underline{x}(t)$ assuming that the initial state $\underline{x}_0$ is set to $\underline{w}_2$. We have:

$$\begin{aligned} \underline{x}(t) &= e^{\mathbf{A}t}\underline{x}_0 + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau \\ &= e^{\mathbf{A}t}\begin{bmatrix} 0 \\ w_{22} \end{bmatrix} + \int_0^t e^{\mathbf{A}(t-\tau)}\begin{bmatrix} -2 \\ 0 \end{bmatrix}\underline{u}(\tau)d\tau \end{aligned} \tag{3.177}$$

Where:

$$\begin{aligned} e^{\mathbf{A}t} &= \mathcal{L}^{-1}\left[(s\mathbb{I} - \mathbf{A})^{-1}\right] \\ &= \mathcal{L}^{-1}\left(\begin{bmatrix} s+1 & -10 \\ 0 & s-1 \end{bmatrix}^{-1}\right) \\ &= \mathcal{L}^{-1}\left(\frac{1}{(s+1)(s-1)}\begin{bmatrix} s-1 & 10 \\ 0 & s+1 \end{bmatrix}\right) \\ &= \mathcal{L}^{-1}\left(\begin{bmatrix} \frac{1}{s+1} & \frac{10}{(s+1)(s-1)} \\ 0 & \frac{1}{s-1} \end{bmatrix}\right) \\ &= \begin{bmatrix} e^{-t} & 5e^t - 5e^{-t} \\ 0 & e^t \end{bmatrix} \end{aligned} \tag{3.178}$$

Consequently state vector $\underline{x}(t)$ reads:

$$\begin{aligned} \underline{x}(t) = &\begin{bmatrix} e^{-t} & 5e^t - 5e^{-t} \\ 0 & e^t \end{bmatrix}\begin{bmatrix} 0 \\ w_{22} \end{bmatrix} \\ &+ \int_0^t \begin{bmatrix} e^{-(t-\tau)} & 5e^{(t-\tau)} - 5e^{-(t-\tau)} \\ 0 & e^{(t-\tau)} \end{bmatrix}\begin{bmatrix} -2 \\ 0 \end{bmatrix}\underline{u}(\tau)d\tau \end{aligned} \tag{3.179}$$

That is:

$$\underline{x}(t) = \begin{bmatrix} 5e^t - 5e^{-t} \\ e^t \end{bmatrix} w_{22} + \int_0^t \begin{bmatrix} -2e^{-(t-\tau)} \\ 0 \end{bmatrix}\underline{u}(\tau)d\tau \tag{3.180}$$

It is clear that for this specific initial state the input vector $\underline{u}(t)$ will not act on the second component of $\underline{x}(t)$ whatever its expression. Consequently it will not be possible to find a control $\underline{u}(t)$ which moves towards zero the initial state vector $\underline{x}_0 = \underline{w}_2$: this state is uncontrollable and the system is said uncontrollable.

■

3.7.6 Stabilizability

A linear system is stabilizable if all unstable modes are controllable or equivalently if all uncontrollable modes are stable.

3.8 Observability

3.8.1 Definition

Let's consider the output response $\underline{y}(t)$ of a linear time-invariant system:

$$\underline{y}(t) = \mathbf{C}e^{\mathbf{A}t}\underline{x}_0 + \int_0^t \mathbf{C}e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau + \mathbf{D}\underline{u}(t) \tag{3.181}$$

Let's define vector $\underline{\tilde{y}}(t)$ as follows:

$$\underline{\tilde{y}}(t) = \underline{y}(t) - \int_0^t \mathbf{C}e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau - \mathbf{D}\underline{u}(t) \tag{3.182}$$

Thus we get:

$$\mathbf{C}e^{\mathbf{A}t}\underline{x}_0 = \underline{\tilde{y}}(t) \tag{3.183}$$

Observability answers the question whether it is possible to determine the initial state $\underline{x}_0$ through the observation of $\underline{\tilde{y}}(t)$, that is from the output signal $\underline{y}(t)$ and the knowledge of the input signal $\underline{u}(t)$.

More precisely an initial state $\underline{x}_0$ is observable if and only if the initial state can be determined from $\underline{\tilde{y}}(t)$ which is observed through the time interval $0 \leq t \leq t_f$, that is from the knowledge of the output signal $\underline{y}(t)$ and the input signal $\underline{u}(t)$ that are observed through the time interval $0 \leq t \leq t_f$. A system is said to be observable when any arbitrary initial state $\underline{x}_0 \in \mathbb{R}^n$ is observable.

If the system is observable then the value $\underline{x}_0$ of the initial state can be determined from signal $\underline{\tilde{y}}(t)$ that has been observed through the time interval $0 \leq t \leq t_f$ as follows:

$$\underline{x}_0 = \mathbf{W}_o^{-1}(t_f)\int_0^{t_f} e^{\mathbf{A}^T\tau}\mathbf{C}^T\underline{\tilde{y}}(\tau)d\tau \tag{3.184}$$

Where $\mathbf{W}_o(t_f)$ is a symmetric matrix defined as follows:

$$\mathbf{W}_o(t_f) = \int_0^{t_f} e^{\mathbf{A}^T\tau}\mathbf{C}^T\mathbf{C}e^{\mathbf{A}\tau}d\tau \tag{3.185}$$

Indeed from $\mathbf{C}e^{\mathbf{A}t}\underline{x}_0 = \underline{\tilde{y}}(t)$ we get:

$$\begin{aligned}
&\mathbf{C}e^{\mathbf{A}t}\underline{x}_0 = \underline{\tilde{y}}(t)\\
&\Rightarrow e^{\mathbf{A}^Tt}\mathbf{C}^T\mathbf{C}e^{\mathbf{A}t}\underline{x}_0 = e^{\mathbf{A}^Tt}\mathbf{C}^T\underline{\tilde{y}}(t)\\
&\Rightarrow \textstyle\int_0^{t_f} e^{\mathbf{A}^T\tau}\mathbf{C}^T\mathbf{C}e^{\mathbf{A}\tau}d\tau\underline{x}_0 = \int_0^{t_f} e^{\mathbf{A}^T\tau}\mathbf{C}^T\underline{\tilde{y}}(\tau)d\tau\\
&\Leftrightarrow \mathbf{W}_o(t_f)\underline{x}_0 = \textstyle\int_0^{t_f} e^{\mathbf{A}^T\tau}\mathbf{C}^T\underline{\tilde{y}}(\tau)d\tau\\
&\Leftrightarrow \underline{x}_0 = \mathbf{W}_o^{-1}(t_f)\textstyle\int_0^{t_f} e^{\mathbf{A}^T\tau}\mathbf{C}^T\underline{\tilde{y}}(\tau)d\tau
\end{aligned} \tag{3.186}$$

Consequently a system is observable if and only if symmetric matrix $\mathbf{W}_o(t_f)$ is nonsingular for any $t_f > 0$. Furthermore $\mathbf{W}_o(t)$ is the solution of the following differential equation:

$$\mathbf{A}^T\mathbf{W}_o(t) + \mathbf{W}_o(t)\mathbf{A} + \mathbf{C}^T\mathbf{C} = \frac{d}{dt}\mathbf{W}_o(t) \tag{3.187}$$

If all eigenvalues of matrix $\mathbf{A}$ have negative real parts ($\mathbf{A}$ is said to be *Hurwitz*) then $\mathbf{W}_o(t)$ tends towards a constant matrix as $t \to \infty$. The observability Gramian $\mathbf{W}_o$ is the following positive definite symmetric matrix:

$$\mathbf{W}_o = \int_0^\infty e^{\mathbf{A}^T\tau}\mathbf{C}^T\mathbf{C}e^{\mathbf{A}\tau}d\tau \tag{3.188}$$

It can be shown that $\mathbf{W}_o$ is the unique solution of the following Lyapunov equation:

$$\mathbf{A}^T\mathbf{W}_o + \mathbf{W}_o\mathbf{A} + \mathbf{C}^T\mathbf{C} = \mathbf{0} \tag{3.189}$$

3.8.2 Kalman's observability rank condition

Let $\mathbf{Q}_o$ be the observability matrix. Matrix $\mathbf{Q}_o$ is defined as follows:

$$\mathbf{Q}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \tag{3.190}$$

It can be shown that a linear system is observable if and only if the rank of the observability matrix $\mathbf{Q}_o$ is equal to n. This is the Kalman's observability rank condition.

The sketch of the demonstration is the following:

- First we recall that the expression of the output vector $\underline{y}(t)$ at time t with respect to the state vector $\underline{x}(t)$ is:

$$\underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \tag{3.191}$$

  where $\underline{x}(t)$ solves the state equation (3.1):

$$\underline{x}(t) = e^{\mathbf{A}t}\underline{x}_0 + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau \tag{3.192}$$

  Thus:

$$\underline{y}(t) = \mathbf{C}e^{\mathbf{A}t}\underline{x}_0 + \mathbf{C}\int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau + \mathbf{D}\underline{u}(t) \tag{3.193}$$

  As far as $\underline{y}(t)$, t and $\underline{u}(t)$ are assumed to be known we rewrite the preceding equation as follows:

$$\underline{y}(t) - \mathbf{C}\int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau - \mathbf{D}\underline{u}(t) = \mathbf{C}e^{\mathbf{A}t}\underline{x}_0 \tag{3.194}$$

- To continue the sketch of the proof we need the Cayley–Hamilton theorem. As shown in (3.157) this theorem indicates that $e^{\mathbf{A}t}$ can be written as follows:

$$e^{\mathbf{A}t} = \sum_{k=0}^{n-1} \gamma_k(t)\mathbf{A}^k \tag{3.195}$$

  where $\gamma_k(t)$ are functions of time t.

– Using (3.157) and the fact that functions $\gamma_k(t)$ are scalar functions $\mathbf{C}e^{\mathbf{A}t}\underline{x}_0$ reads:

$$\begin{aligned}\mathbf{C}e^{\mathbf{A}t}\underline{x}_0 &= \mathbf{C}\textstyle\sum_{k=0}^{n-1}\gamma_k(t)\mathbf{A}^k\underline{x}_0\\ &= \textstyle\sum_{k=0}^{n-1}\mathbf{C}\gamma_k(t)\mathbf{A}^k\underline{x}_0\\ &= \left(\textstyle\sum_{k=0}^{n-1}\gamma_k(t)\mathbf{C}\mathbf{A}^k\right)\underline{x}_0\end{aligned} \tag{3.196}$$

Now let's sample the time interval $0 \le t \le t_f$ into n values of time, $t_1 = 0, t_2, \cdots, t_{n-1}, t_n = t_f$ which are situated inside the time interval $0 \le t \le t_f$. Using (3.194) for each value t_i of the time we get:

$$\begin{aligned}&\begin{bmatrix}\underline{y}(t_1) - \mathbf{C}\int_0^{t_1} e^{\mathbf{A}(t_1-\tau)}\mathbf{B}\underline{u}(\tau)d\tau - \mathbf{D}\underline{u}(t_1)\\ \underline{y}(t_2) - \mathbf{C}\int_0^{t_2} e^{\mathbf{A}(t_2-\tau)}\mathbf{B}\underline{u}(\tau)d\tau - \mathbf{D}\underline{u}(t_2)\\ \vdots\\ \underline{y}(t_n) - \mathbf{C}\int_0^{t_n} e^{\mathbf{A}(t_n-\tau)}\mathbf{B}\underline{u}(\tau)d\tau - \mathbf{D}\underline{u}(t_n)\end{bmatrix}\\ &= \begin{bmatrix}\gamma_0(t_1) & \gamma_1(t_1) & \cdots & \gamma_{n-1}(t_1)\\ \gamma_0(t_2) & \gamma_1(t_2) & \cdots & \gamma_{n-1}(t_2)\\ \vdots & \vdots & \vdots & \vdots\\ \gamma_0(t_n) & \gamma_1(t_n) & \cdots & \gamma_{n-1}(t_n)\end{bmatrix}\begin{bmatrix}\mathbf{C}\\ \mathbf{CA}\\ \vdots\\ \mathbf{CA}^{n-1}\end{bmatrix}\underline{x}_0\end{aligned} \tag{3.197}$$

That is:

$$\begin{bmatrix}\underline{y}(t_1) - \mathbf{C}\int_0^{t_1} e^{\mathbf{A}(t_1-\tau)}\mathbf{B}\underline{u}(\tau)d\tau - \mathbf{D}\underline{u}(t_1)\\ \underline{y}(t_2) - \mathbf{C}\int_0^{t_2} e^{\mathbf{A}(t_2-\tau)}\mathbf{B}\underline{u}(\tau)d\tau - \mathbf{D}\underline{u}(t_2)\\ \vdots\\ \underline{y}(t_n) - \mathbf{C}\int_0^{t_n} e^{\mathbf{A}(t_n-\tau)}\mathbf{B}\underline{u}(\tau)d\tau - \mathbf{D}\underline{u}(t_n)\end{bmatrix} = \mathbf{V}\mathbf{Q}_o\underline{x}_0 \tag{3.198}$$

Where:

$$\mathbf{V} = \begin{bmatrix}\gamma_0(t_1) & \gamma_1(t_1) & \cdots & \gamma_{n-1}(t_1)\\ \gamma_0(t_2) & \gamma_1(t_2) & \cdots & \gamma_{n-1}(t_2)\\ \vdots & \vdots & \vdots & \vdots\\ \gamma_0(t_n) & \gamma_1(t_n) & \cdots & \gamma_{n-1}(t_n)\end{bmatrix} \tag{3.199}$$

In order to be able to compute the expression of vector $\underline{x}_0$ from (3.198) and assuming that matrix $\mathbf{V}$ is invertible (which is always the case when all the eigenvalues of matrix $\mathbf{A}$ are distinct), the observability matrix $\mathbf{Q}_o$ shall be invertible; consequently the rank of the observability matrix $\mathbf{Q}_o$ shall be equal to n. Thus we retrieve the Kalman's observability rank condition.

Example 3.15. *Let's consider the following realization:*

$$\left(\begin{array}{c|c}\mathbf{A} & \mathbf{B}\\ \hline \mathbf{C} & \mathbf{D}\end{array}\right) = \left(\begin{array}{cc|c}-1 & 0 & -2\\ 10 & 1 & 3\\ \hline -2 & 0 & -2\end{array}\right) \tag{3.200}$$

The observability matrix $\mathbf{Q}_o$ *reads:*

$$\begin{aligned}\mathbf{Q}_o &= \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} -2 & 0 \end{bmatrix} \\ \begin{bmatrix} -2 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 10 & 1 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} -2 & 0 \\ 2 & 0 \end{bmatrix}\end{aligned} \tag{3.201}$$

Consequently $\operatorname{rank}(\mathbf{Q}_o) = 1 \neq 2$. *We conclude that the system in not observable.*

■

3.8.3 Use of the diagonal form: Gilbert's criteria

Assuming that all eigenvalues are distinct, observability property can be readily analyzed by inspecting the null columns of the output matrix $\mathbf{C}$ as soon as we get the modal (or diagonal) form of the state-space representation.

Indeed we have seen in the the chapter dedicated to the state-space representation that for a single output system matrix $\mathbf{C}_m$ of the output equation $y(t) = \mathbf{C}_m\,\underline{x}(t) + \mathbf{D}\underline{u}(t)$ read as follows assuming that state matrix $\mathbf{A}$ has n independent eigenvectors :

$$\mathbf{C}_m = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix} \tag{3.202}$$

Thus in the time domain the diagonal form of the state-space representation $y(t) = \mathbf{C}_m\,\underline{x}(t) + \mathbf{D}\underline{u}(t)$ reads:

$$y(t) = c_1x_1(t) + c_2x_2(t) + \cdots + c_nx_n(t) + \mathbf{D}\underline{u}(t) \tag{3.203}$$

Gilbert's observability criteria (1963) states that a system with distinct eigenvalues is observable if and only if each column of output matrix $\mathbf{C}$ of the diagonal realization (all eigenvalues are distinct) has at least one non zero element.

Indeed if at least one of the c_i's coefficients is zero then the output signal $y(t)$ is independent of the state component $x_i(t)$ and the state is unobservable.

Alternatively, if we apply Kalman's observability rank condition in the modal basis we get:

$$\begin{aligned}\mathbf{A}_m &= \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \\ \Rightarrow \mathbf{Q}_{om} &= \begin{bmatrix} \mathbf{C}_m \\ \mathbf{C}_m\mathbf{A}_m \\ \vdots \\ \mathbf{C}_m\mathbf{A}_m^{n-1} \end{bmatrix} = \begin{bmatrix} c_1 & \cdots & c_n \\ c_1\lambda_1 & \cdots & c_n\lambda_n \\ \vdots & \vdots & \vdots \\ c_1\lambda_1^{n-1} & \cdots & c_n\lambda_n n-1 \end{bmatrix}\end{aligned} \tag{3.204}$$

Then it can be seen that $\mathbf{Q}_{om}$ is full rank as soon as $c_i \neq 0\ \forall i = 1, \cdots, n$.

For multi outputs system with p outputs then matrix $\mathbf{C}_m$ has p rows and the preceding analysis is readily extended to each row of matrix $\mathbf{C}_m$ assuming that the state-space representation is the diagonal form.

3.8.4 Input-output differential equation and observable modes

Assume that the modal basis is used and let $\mathbf{C}_m$ be as follows:

$$\mathbf{C}_m = \begin{bmatrix} \underline{c}_1 & \cdots & \underline{c}_q & \underline{0} & \cdots & \underline{0} \end{bmatrix} \tag{3.205}$$

It is clear from the Gilbert's criteria that $\{\lambda_1, \cdots, \lambda_q\}$ forms the set of observable modes.

Furthermore, the following relation holds:

$$\mathbf{C}_m \left(\mathbf{A}_m - \lambda_1 \mathbb{I}\right) \cdots \left(\mathbf{A}_m - \lambda_q \mathbb{I}\right) = \mathbf{0} \tag{3.206}$$

Indeed:

$$\begin{aligned}
\mathbf{C}_m \quad & \left(\mathbf{A}_m - \lambda_1 \mathbb{I}\right) \cdots \left(\mathbf{A}_m - \lambda_q \mathbb{I}\right) \\
& = \mathbf{C}_m \begin{bmatrix} 0 & & & & \\ & \lambda_2 - \lambda_1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \lambda_n - \lambda_1 \end{bmatrix} \cdots \begin{bmatrix} \ddots & & & & \\ & \lambda_{q-1} - \lambda_q & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & \lambda_n - \lambda_q \end{bmatrix} \\
& = \begin{bmatrix} \underline{c}_1 & \cdots & \underline{c}_q & \underline{0} & \cdots & \underline{0} \end{bmatrix} \begin{bmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & \alpha_{q+1} & & \\ & & & & \ddots & \\ & & & & & \alpha_n \end{bmatrix} \\
& = \mathbf{0}
\end{aligned} \tag{3.207}$$

This completes the proof. ■

Coming back to any state space representation, (3.206) reads:

$$\begin{cases} \mathbf{A}_m = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \\ \mathbf{C}_m = \mathbf{C}\mathbf{P} \end{cases} \Rightarrow \mathbf{C}\left(\mathbf{A} - \lambda_1 \mathbb{I}\right) \cdots \left(\mathbf{A} - \lambda_q \mathbb{I}\right) = \mathbf{0} \tag{3.208}$$

Furthemore, $\left(\mathbf{A} - \lambda_1 \mathbb{I}\right) \cdots \left(\mathbf{A} - \lambda_q \mathbb{I}\right)$ can be developed as follows:

$$\left(\mathbf{A} - \lambda_1 \mathbb{I}\right) \cdots \left(\mathbf{A} - \lambda_q \mathbb{I}\right) = \mathbf{A}^q + a_{q-1}\,\mathbf{A}^{q-1} + \cdots + a_1\ \mathbf{A} + a_0\,\mathbb{I} \tag{3.209}$$

where coefficients a_i are the coefficients of the polynomial obtained from the observable modes:

$$\left(s - \lambda_1\right) \cdots \left(s - \lambda_q\right) = s^q + a_{q-1}\,s^{q-1} + \cdots + a_1\ s + a_0 \tag{3.210}$$

Thus (3.208) reads as follows, where $\{\lambda_1, \cdots, \lambda_q\}$ is the set of observable modes:

$$\begin{aligned}&(s-\lambda_1)\cdots(s-\lambda_q) = s^q + a_{q-1}\, s^{q-1} + \cdots + a_1\, s + a_0 \\ &\Rightarrow \mathbf{CA}^q + a_{q-1}\,\mathbf{CA}^{q-1} + \cdots + a_1\,\mathbf{CA} + a_0\,\mathbf{C} = \mathbf{0}\end{aligned} \tag{3.211}$$

In other words, $(s-\lambda_1)\cdots(s-\lambda_q)$ is an annihilating polynomial of $\mathbf{CA}^i$.

Finally we can use relation $\underline{y} = \mathbf{C}\underline{x} + \mathbf{D}\underline{u}$ to compute $\underline{y}^{(q)} + a_{q-1}\,\underline{y}^{(q-1)} + \cdots + a_1\,\dot{\underline{y}} + a_0\,\underline{y}$ and finally obtain the input-output relation between output vector $\underline{y}$ and input vector $\underline{u}$. Indeed, this expression will exhibit the term $\left(\mathbf{CA}^q + a_{q-1}\,\mathbf{CA}^{q-1} + \cdots + a_1\,\mathbf{CA} + a_0\,\mathbf{C}\right)\underline{x}$, which is zero from (3.211).

3.8.5 Popov-Belevitch-Hautus (PBH) test

There exists another test for observability which is called the Popov-Belevitch-Hautus (PBH) test.

Popov-Belevitch-Hautus (PBH) test indicates that a linear system is observable when the rank of matrix $\begin{bmatrix} \mathbf{A}-\lambda_i\mathbb{I} \\ \mathbf{C} \end{bmatrix}$ is equal to n for all eigenvalues $\{\lambda_i\}$ of matrix $\mathbf{A}$.

A linear system is *detectable* when the rank of matrix $\begin{bmatrix} \mathbf{A}-\lambda_i\mathbb{I} \\ \mathbf{C} \end{bmatrix}$ is equal to n for all unstable eigenvalues $\{\lambda_i\}$ of matrix $\mathbf{A}$.

Eigenvalues λ_i for which rank of matrix $\begin{bmatrix} \mathbf{A}-\lambda_i\mathbb{I} \\ \mathbf{C} \end{bmatrix}$ is not equal to n are said unobservable.

Equivalently, an eigenvalue λ_i is observable[8] if all its corresponding right eigenvectors $\underline{v}_i$ (i.e. $\underline{v}_i \neq \underline{0}$ such that $\mathbf{A}\underline{v}_i = \lambda_i\underline{v}_i$) satisfy $\mathbf{C}\underline{v}_i \neq \underline{0}$.

Example 3.16. *Let's consider the following realization:*

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) = \left(\begin{array}{cc|c} -1 & 0 & -2 \\ 10 & 1 & 3 \\ \hline -2 & 0 & -2 \end{array}\right) \tag{3.212}$$

Matrix $\mathbf{A}$ *has two modes,* $\lambda_1 = -1$ *and* $\lambda_2 = 1$.
Let's apply the PBH test for $\lambda_1 = -1$*:*

$$\operatorname{rank}\begin{bmatrix} \mathbf{A}-\lambda_1\mathbb{I} \\ \mathbf{C} \end{bmatrix} = \operatorname{rank}\begin{bmatrix} 0 & 0 \\ 10 & 2 \\ -2 & 0 \end{bmatrix} = 2 \tag{3.213}$$

We conclude that the mode $\lambda_1 = -1$ *is observable.*
Let's apply the PBH test for $\lambda_2 = 1$*:*

$$\operatorname{rank}\begin{bmatrix} \mathbf{A}-\lambda_2\mathbb{I} \\ \mathbf{C} \end{bmatrix} = \operatorname{rank}\begin{bmatrix} -2 & 0 \\ 10 & 0 \\ -2 & 0 \end{bmatrix} = 1 \neq 2 \tag{3.214}$$

[8] Antsaklis P. J. and Michel A. N., A Linear Systems Primer, Birkhäuser Boston, Boston,MA, 2007

We conclude that the mode $\lambda_2 = 1$ is not observable.

■

3.8.6 Unobservable mode

Following Bélanger[7] a state $\underline{x}_{\bar{o}} \neq \underline{0}$ is said to be unobservable if the zero-input response of the system (i.e. the system's response to the initial condition $\underline{x}(0) = \underline{x}_0 \neq \underline{0}$ when the input signal $\underline{u}(t)$ is zero) with $\underline{x}(0) = \underline{x}_{\bar{o}}$ is zero $\forall\ t \geq 0$. An unobservable state $\underline{x}_{\bar{o}} \neq \underline{0}$ satisfies the following equation:

$$\begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \underline{x}_{\bar{o}} = \underline{0} \tag{3.215}$$

Equivalently the pair $(\mathbf{A}, \mathbf{C})$ is unobservable if and only if there exists an eigenvector $\underline{v}_i$ of matrix $\mathbf{A}$ such that $\mathbf{C}\underline{v}_i = \underline{0}$. If $\mathbf{C}\underline{v}_i = \underline{0}$ then the mode λ_i (i.e. the eigenvalue) corresponding to $\underline{v}_i$ is called an unobservable mode.

Indeed if $\underline{v}_i$ is an eigenvector of matrix $\mathbf{A}$ corresponding to the mode (i.e. the eigenvalue) λ_i then $\mathbf{A}\underline{v}_i = \lambda_i \underline{v}_i$. Specializing $\underline{x}_{\bar{o}}$ to $\underline{v}_i$ Equation (3.215) reads:

$$\begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \underline{v}_i = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\lambda_i \\ \vdots \\ \mathbf{C}\lambda_i^{n-1} \end{bmatrix} \underline{v}_i = \begin{bmatrix} 1 \\ \lambda_i \\ \vdots \\ \lambda_i^{n-1} \end{bmatrix} \mathbf{C}\underline{v}_i = \underline{0} \tag{3.216}$$

Coupling $\mathbf{A}\underline{v}_i = \lambda_i \underline{v}_i$ and $\mathbf{C}\underline{v}_i = \underline{0}$ leads to the Popov-Belevitch-Hautus (PBH) test for observability:

$$\begin{cases} \mathbf{A}\underline{v}_i = \lambda_i \underline{v}_i \\ \mathbf{C}\underline{v}_i = \underline{0} \end{cases} \Leftrightarrow \begin{bmatrix} \mathbf{A} - \lambda_i \mathbb{I} \\ \mathbf{C} \end{bmatrix} \underline{v}_i = \underline{0} \tag{3.217}$$

3.8.7 Detectability

A linear system is detectable if all unstable modes are observable or equivalently if all unobservable modes are stable.

3.8.8 Additional observability topic: relation between linear differential equation and state-space representation

We consider hereafter a SISO system which is described by the following linear differential equation where $m < n_o$:

$$\begin{aligned} a_0\, y(t) + a_1 \frac{dy(t)}{dt} + \cdots + a_{n_o-1} \frac{d^{n_o-1}y(t)}{dt^{n_o-1}} + a_{n_o} \frac{d^{n_o}y(t)}{dt^{n_o}} = \\ b_0\, u(t) + b_1 \frac{du(t)}{dt} + \cdots + b_{m-1} \frac{d^{m-1}u(t)}{dt^{m-1}} + b_m \frac{d^m u(t)}{dt^m} \end{aligned} \tag{3.218}$$

As far as $m < n_o$, the state-space representation of (3.218) is the following:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}u(t) \\ y(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{3.219}$$

Now, we will see what is the constraint on state matrix $\mathbf{A}$ to retrieve the linear differential equation from the state-space representation. First, we will compute the derivatives of output $y(t)$:

$$\begin{cases} \dot{y}(t) &= \mathbf{C}\dot{\underline{x}}(t) = \mathbf{CA}\underline{x}(t) + \mathbf{CB}u(t) \\ \ddot{y}(t) &= \mathbf{CA}^2\underline{x}(t) + \mathbf{CAB}u(t) + \mathbf{CB}\dot{u}(t) = \mathbf{CA}^2\underline{x}(t) + \mathbf{C}\left(\mathbf{AB}u(t) + \mathbf{B}\dot{u}(t)\right) \\ & \vdots \\ \frac{d^k y(t)}{dt^k} &= \mathbf{CA}^k\underline{x}(t) + \mathbf{C}\left(\sum_{i=0}^{k-1} \mathbf{A}^i\mathbf{B}\frac{d^{k-1-i}u(t)}{dt^{k-1-i}}\right) \end{cases} \tag{3.220}$$

Thus, the linear combination $\sum_{k=0}^{n_o} a_k \frac{d^k y(t)}{dt^k}$ which appears in the linear differential equation (3.218) reads:

$$\sum_{k=0}^{n_o} a_k \frac{d^k y(t)}{dt^k} = \sum_{k=0}^{n_o} a_k \left(\mathbf{CA}^k\underline{x}(t) + \mathbf{C}\left(\sum_{i=0}^{k-1} \mathbf{A}^i\mathbf{B}\frac{d^{k-1-i}u(t)}{dt^{k-1-i}}\right)\right) \tag{3.221}$$

We conclude that the linear combination $\sum_{k=0}^{n_o} a_k \frac{d^k y(t)}{dt^k}$ can be written as a linear combination of input $u(t)$ and its derivatives as in the linear differential equation (3.218) as soon as state vector $\underline{x}(t)$ no more appears in this the linear combination. Thus the following relation must hold:

$$\sum_{k=0}^{n_o} a_k \mathbf{CA}^k = 0 \Leftrightarrow \mathbf{C}\left(\sum_{k=0}^{n_o} a_k \mathbf{A}^k\right) = 0 \tag{3.222}$$

It can be shown that polynomial $\sum_{k=0}^{n_o} a_k\, s^k$ is the polynomial corresponding to the observable modes of the system:

$$\sum_{k=0}^{n_o} a_k\, s^k = \prod_{i=1}^{n_o} (s - \lambda_i) \quad \text{where } \{\lambda_1, \cdots, \lambda_{n_o}\} \text{ observable modes} \tag{3.223}$$

3.9 Interpretation of the diagonal (or modal) decomposition

When the state matrix $\mathbf{A}$ is diagonalizable we have seen in (3.56) that $e^{\mathbf{A}t}$ reads as follows where $\underline{v}_i$ is a right eigenvector corresponding to eigenvalue λ_i and $\underline{w}_i$ is a left eigenvector corresponding to the same eigenvalue λ_i:

$$e^{\mathbf{A}t} = \sum_{i=1}^{n} \underline{v}_i \underline{w}_i^T e^{\lambda_i t} \tag{3.224}$$

On the other hand we know from (3.13) that the output response of the system can be expressed as follows:

$$\underline{y}(t) = \mathbf{C}e^{\mathbf{A}t}\underline{x}(0) + \int_0^t \mathbf{C}e^{\mathbf{A}(t-\tau)}\mathbf{B}\underline{u}(\tau)d\tau + \mathbf{D}\underline{u}(t) \tag{3.225}$$

Gathering the two previous results leads to the following expression of the output vector $\underline{y}(t)$ where it is worth noticing that $\underline{w}_i^T \underline{x}(0)$ is a scalar:

$$\underline{y}(t) = \sum_{i=1}^{n} \mathbf{C}\underline{v}_i e^{\lambda_i t} \left(\underline{w}_i^T \underline{x}(0)\right) + \sum_{i=1}^{n} \mathbf{C}\underline{v}_i \int_0^t e^{\lambda_i (t-\tau)} \underline{w}_i^T \mathbf{B}\underline{u}(\tau) d\tau + \mathbf{D}\underline{u}(t) \quad (3.226)$$

The product $\mathbf{C}\underline{v}_i$ is called the direction in the output space associated with eigenvalue λ_i. From the preceding equation it is clear that if $\mathbf{C}\underline{v}_i = \underline{0}$ then any motion in the direction $\underline{v}_i$ cannot be observed in the output $\underline{y}(t)$ and we say that eigenvalue λ_i is unobservable.

The product $\underline{w}_i^T \mathbf{B}$ is called the direction in the input space associated with eigenvalue λ_i. From the preceding equation we cannotice that if $\underline{w}_i^T \mathbf{B} = \underline{0}$ the control input $\underline{u}(t)$ cannot participate to the motion in the direction $\underline{v}_i$ and we say that eigenvalue λ_i is uncontrollable.

As a consequence the coupling between inputs, states and outputs is set by the eigenvectors $\underline{v}_i$ and $\underline{w}_i^T$. It can be seen that those vectors also influence the numerator of the transfer function $\mathbf{F}(s)$ which reads:

$$\mathbf{F}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1} \mathbf{B} + \mathbf{D} = \sum_{i=1}^{n} \frac{\mathbf{C}\underline{v}_i \underline{w}_i^T \mathbf{B}}{s - \lambda_i} + \mathbf{D} \quad (3.227)$$

Indeed let $\mathbf{\Lambda}$ be the diagonal form of the diagonalizable matrix $\mathbf{A}$:

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \quad (3.228)$$

We have seen that the change of basis matrix $\mathbf{P}$ as well as its inverse $\mathbf{P}^{-1}$ have the following expression:

$$\mathbf{\Lambda} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \text{ where } \begin{cases} \mathbf{P} = \begin{bmatrix} \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_n \end{bmatrix} \\ \mathbf{P}^{-1} = \begin{bmatrix} \underline{w}_1^T \\ \underline{w}_2^T \\ \vdots \\ \underline{w}_n^T \end{bmatrix} \end{cases} \quad (3.229)$$

Using the fact that $(\mathbf{XY})^{-1} = \mathbf{Y}^{-1}\mathbf{X}^{-1}$ for any two inversible square

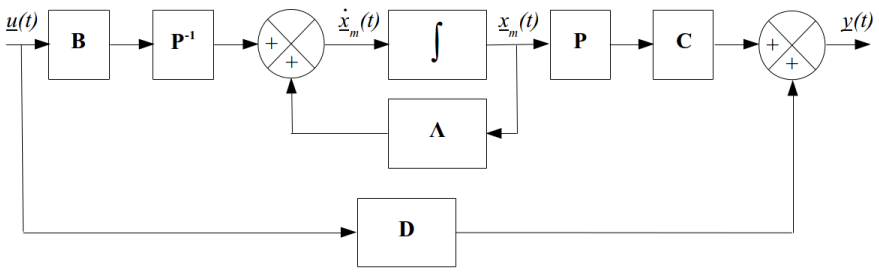

Figure 3.1: Modal decomposition of a transfer function

matrices the transfer function $\mathbf{F}(s)$ reads:

$$
\begin{aligned}
\mathbf{F}(s) &= \mathbf{C}\left(s\mathbb{I}-\mathbf{A}\right)^{-1}\mathbf{B}+\mathbf{D} \\
&= \mathbf{C}\left(\mathbf{P}\left(s\mathbb{I}-\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right)\mathbf{P}^{-1}\right)^{-1}\mathbf{B}+\mathbf{D} \\
&= \mathbf{C}\mathbf{P}\left(s\mathbb{I}-\mathbf{\Lambda}\right)^{-1}\mathbf{P}^{-1}\mathbf{B}+\mathbf{D} \\
&= \mathbf{C}\left(\mathbf{P}\left(s\mathbb{I}-\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right)\mathbf{P}^{-1}\right)^{-1}\mathbf{B}+\mathbf{D} \\
&= \mathbf{C}\mathbf{P}\left(s\mathbb{I}-\mathbf{\Lambda}\right)^{-1}\mathbf{P}^{-1}\mathbf{B}+\mathbf{D} \\
&= \begin{bmatrix} \mathbf{C}\underline{v}_1 & \mathbf{C}\underline{v}_2 & \cdots & \mathbf{C}\underline{v}_n \end{bmatrix}
\begin{bmatrix} \frac{1}{s-\lambda_1} & & \\ & \ddots & \\ & & \frac{1}{s-\lambda_n} \end{bmatrix}
\begin{bmatrix} \underline{w}_1^T\mathbf{B} \\ \underline{w}_2^T\mathbf{B} \\ \vdots \\ \underline{w}_n^T\mathbf{B} \end{bmatrix}+\mathbf{D} \\
&= \begin{bmatrix} \mathbf{C}\underline{v}_1 & \mathbf{C}\underline{v}_2 & \cdots & \mathbf{C}\underline{v}_n \end{bmatrix}
\begin{bmatrix} \frac{\underline{w}_1^T\mathbf{B}}{s-\lambda_1} \\ \frac{\underline{w}_2^T\mathbf{B}}{s-\lambda_2} \\ \vdots \\ \frac{\underline{w}_n^T\mathbf{B}}{s-\lambda_n} \end{bmatrix}+\mathbf{D}
\end{aligned}
\tag{3.230}
$$

We finally get:

$$
\mathbf{F}(s)=\sum_{i=1}^{n}\frac{\mathbf{C}\underline{v}_i\underline{w}_i^T\mathbf{B}}{s-\lambda_i}+\mathbf{D} \tag{3.231}
$$

Figure 3.1 presents the diagonal (or modal) decomposition of the transfer function $\mathbf{F}(s)$ where $\underline{x}_m(t)$ is the state vector expressed in the diagonal (or modal) basis and matrices $\mathbf{\Lambda}$, $\mathbf{P}$ and $\mathbf{P}^{-1}$ are defined as follows:

$$
\begin{cases}
\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \\
\mathbf{P} = \begin{bmatrix} \underline{v}_1 & \cdots & \underline{v}_n \end{bmatrix} \\
\mathbf{P}^{-1} = \begin{bmatrix} \underline{w}_1^T \\ \vdots \\ \underline{w}_n^T \end{bmatrix}
\end{cases}
\tag{3.232}
$$

3.10 Duality principle

The concept of controllability and observability was introduced by American-Hungarian engineer Rudolf E. Kalman for linear dynamic systems [9]. Let's consider a system which is denoted Σ. Then system Σ_D, which is the dual of Σ, as defined as follows:

$$\Sigma = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) \Rightarrow \Sigma_D = dual\,(\Sigma) = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)^T = \left(\begin{array}{c|c} \mathbf{A}^T & \mathbf{C}^T \\ \hline \mathbf{B}^T & \mathbf{D}^T \end{array}\right) \tag{3.233}$$

The duality principle indicates that:

- System Σ is observable if and only if system Σ_D is controllable.
- System Σ is controllable if and only if system Σ_D is observable.

Furthermore we cannotice that the observable canonical form is the dual of the controllable canonical form.

3.11 Kalman decomposition

3.11.1 Controllable / uncontrollable decomposition

We recall that the controllability matrix $\mathbf{Q}_c$ is defined as follows:

$$\mathbf{Q}_c = \left[\begin{array}{cccc} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{array}\right] \tag{3.234}$$

Suppose that the system is *not* controllable, meaning that:

$$\text{rank}\,(\mathbf{Q}_c) = n_c < n \tag{3.235}$$

Let $\mathbf{P}_{c\bar{c}}$ be the following change of basis matrix which defines a new state vector $\underline{x}_{c\bar{c}}(t)$ as follows:

$$\underline{x}(t) = \mathbf{P}_{c\bar{c}}\,\underline{x}_{c\bar{c}}(t) \Leftrightarrow \underline{x}_{c\bar{c}}(t) = \mathbf{P}_{c\bar{c}}^{-1}\underline{x}(t) \tag{3.236}$$

The first n_c columns of $\mathbf{P}_{c\bar{c}}$ are chosen to be n_c independent columns of $\mathbf{Q}_c$ whereas the remaining $n - n_c$ columns are arbitrarily chosen such that $\mathbf{P}_{c\bar{c}}$ is invertible:

$$\mathbf{P}_{c\bar{c}} = \left[\begin{array}{ccc|ccc} \underline{q}_1 & \cdots & \underline{q}_{n_c} & \underline{q}_{n_c+1} & \cdots & \underline{q}_n \end{array}\right] \tag{3.237}$$

Then, according to the results in section 2.2, the state-space representation involving the state vector $\underline{x}_{c\bar{c}}(t)$ reads:

$$\left\{\begin{array}{l} \dot{\underline{x}}_{c\bar{c}}(t) = \mathbf{A}_{c\bar{c}}\underline{x}_{c\bar{c}}(t) + \mathbf{B}_{c\bar{c}}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}_{c\bar{c}}\underline{x}_{c\bar{c}}(t) + \mathbf{D}\underline{u}(t) \end{array}\right. \tag{3.238}$$

[9] R. E. Kalman, On the General Theory of Control Systems, Proceeding of the 1st IFAC congress, Moscow 1960

where:

$$\begin{cases} \mathbf{A}_{c\bar{c}} = \mathbf{P}_{c\bar{c}}^{-1}\mathbf{A}\mathbf{P}_{c\bar{c}} := \begin{bmatrix} \mathbf{A}_c & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{\bar{c}} \end{bmatrix} \\ \mathbf{B}_{c\bar{c}} = \mathbf{P}_{c\bar{c}}^{-1}\mathbf{B} := \begin{bmatrix} \mathbf{B}_c \\ \mathbf{0} \end{bmatrix} \\ \mathbf{C}_{c\bar{c}} = \mathbf{C}\mathbf{P}_{c\bar{c}} := \begin{bmatrix} \mathbf{C}_c & \mathbf{C}_{\bar{c}} \end{bmatrix} \end{cases} \tag{3.239}$$

It is worth noticing that the feedforward matrix $\mathbf{D}$ is independent of the choice of the state vector.

The state vector $\underline{x}_{c\bar{c}}(t)$ can be split into vector $\underline{x}_c(t)$, which represents the controllable states, and vector $\underline{x}_{\bar{c}}(t)$ which represents the uncontrollable states:

$$\underline{x}_{c\bar{c}}(t) := \begin{bmatrix} \underline{x}_c(t) \\ \underline{x}_{\bar{c}}(t) \end{bmatrix} \tag{3.240}$$

Furthermore, the reduced-order state equation of the controllable state vector $\underline{x}_c(t)$ is controllable and has the same transfer function than the full state equation:

$$\begin{cases} \dot{\underline{x}}_c(t) = \mathbf{A}_c\underline{x}_c(t) + \mathbf{B}_c\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}_c\underline{x}_c(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{3.241}$$

3.11.2 Observable / unobservable decomposition

We recall that the observability matrix $\mathbf{Q}_o$ is defined as follows:

$$\mathbf{Q}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \tag{3.242}$$

Suppose that the system is *not* observable, meaning that:

$$\operatorname{rank}(\mathbf{Q}_o) = n_o < n \tag{3.243}$$

Let $\mathbf{P}_{o\bar{o}}$ be the following change of basis matrix which defines a new state vector $\underline{x}_{o\bar{o}}(t)$ as follows:

$$\underline{x}(t) = \mathbf{P}_{o\bar{o}}\,\underline{x}_{o\bar{o}}(t) \Leftrightarrow \underline{x}_{o\bar{o}}(t) = \mathbf{P}_{o\bar{o}}^{-1}\underline{x}(t) \tag{3.244}$$

The first n_o rows of $\mathbf{P}_{o\bar{o}}^{-1}$ are chosen to be n_o independent rows of $\mathbf{Q}_o$ whereas the remaining $n - n_o$ rows are arbitrarily chosen such that $\mathbf{P}_{o\bar{o}}^{-1}$ is invertible:

$$\mathbf{P}_{o\bar{o}}^{-1} = \begin{bmatrix} \underline{q}_1 \\ \vdots \\ \underline{q}_{n_o} \\ \hline \underline{q}_{n_o+1} \\ \vdots \\ \underline{q}_n \end{bmatrix} \tag{3.245}$$

Then, according to the results in section 2.2, the state-space representation involving the state vector $\underline{x}_{o\bar{o}}(t)$ reads:

$$\begin{cases} \dot{\underline{x}}_{o\bar{o}}(t) = \mathbf{A}_{o\bar{o}}\underline{x}_{o\bar{o}}(t) + \mathbf{B}_{o\bar{o}}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}_{o\bar{o}}\underline{x}_{o\bar{o}}(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{3.246}$$

where:

$$\begin{cases} \mathbf{A}_{o\bar{o}} = \mathbf{P}_{o\bar{o}}^{-1}\mathbf{A}\mathbf{P}_{o\bar{o}} := \begin{bmatrix} \mathbf{A}_o & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{\bar{o}} \end{bmatrix} \\ \mathbf{B}_{o\bar{o}} = \mathbf{P}_{o\bar{o}}^{-1}\mathbf{B} := \begin{bmatrix} \mathbf{B}_o \\ \mathbf{B}_{\bar{o}} \end{bmatrix} \\ \mathbf{C}_{o\bar{o}} = \mathbf{C}\mathbf{P}_{o\bar{o}} := \begin{bmatrix} \mathbf{C}_o & \mathbf{0} \end{bmatrix} \end{cases} \tag{3.247}$$

It is worth noticing that the feedforward matrix $\mathbf{D}$ is independent of the choice of the state vector.

The state vector $\underline{x}_{o\bar{o}}(t)$ can be split into vector $\underline{x}_o(t)$, which represents the observable states, and vector $\underline{x}_{\bar{o}}(t)$ which represents the unobservable states:

$$\underline{x}_{o\bar{o}}(t) := \begin{bmatrix} \underline{x}_o(t) \\ \underline{x}_{\bar{o}}(t) \end{bmatrix} \tag{3.248}$$

Furthermore, the reduced-order state equation of the observable state vector $\underline{x}_o(t)$ is observable and has the same transfer function than the full state equation:

$$\begin{cases} \dot{\underline{x}}_o(t) = \mathbf{A}_o\underline{x}_o(t) + \mathbf{B}_o\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}_o\underline{x}_o(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{3.249}$$

3.11.3 Canonical decomposition

Kalman decomposition is a state-space representation which makes clear the observable and controllable components of the system. More precisely any linear system $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$ can be transformed by a similarity transformation as follows:

$$\begin{cases} \dot{\underline{x}}_K(t) &= \mathbf{A}_K\underline{x}_K(t) + \mathbf{B}_K\underline{u}(t) \\ \underline{y}(t) &= \mathbf{C}_K\underline{x}_K(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{3.250}$$

The Kalman decomposition expands as follows:

$$\left(\begin{array}{c|c} \mathbf{A}_K & \mathbf{B}_K \\ \hline \mathbf{C}_K & \mathbf{D} \end{array}\right) = \left(\begin{array}{cccc|c} \mathbf{A}_{c\bar{o}} & \mathbf{A}_{12} & \mathbf{A}_{13} & \mathbf{A}_{14} & \mathbf{B}_{c\bar{o}} \\ \mathbf{0} & \mathbf{A}_{co} & \mathbf{0} & \mathbf{A}_{24} & \mathbf{B}_{co} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_{\bar{c}\bar{o}} & \mathbf{A}_{34} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{\bar{c}o} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{C}_{co} & \mathbf{0} & \mathbf{C}_{\bar{c}o} & \mathbf{D} \end{array}\right) \tag{3.251}$$

This leads to the conclusion that[10]:

- Subsystem $\left(\begin{array}{c|c} \mathbf{A}_{co} & \mathbf{B}_{co} \\ \hline \mathbf{C}_{co} & \mathbf{D} \end{array}\right)$ is both controllable and observable;

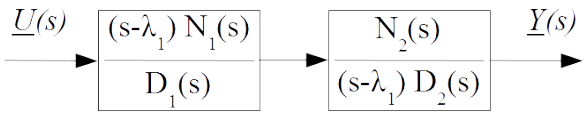

Figure 3.2: Example of uncontrollable mode through pole / zero cancellation in series interconnection

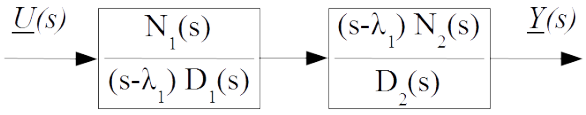

Figure 3.3: Example of unobservable mode through pole / zero cancellation in series interconnection

– Subsystem $\left(\begin{array}{cc|c} \mathbf{A}_{c\bar{o}} & \mathbf{A}_{12} & \mathbf{B}_{c\bar{o}} \\ \mathbf{0} & \mathbf{A}_{co} & \mathbf{B}_{co} \\ \hline \mathbf{0} & \mathbf{C}_{co} & \mathbf{D} \end{array}\right)$ is controllable;

– Subsystem $\left(\begin{array}{cc|c} \mathbf{A}_{co} & \mathbf{A}_{24} & \mathbf{B}_{co} \\ \mathbf{0} & \mathbf{A}_{\bar{c}o} & \mathbf{0} \\ \hline \mathbf{C}_{co} & \mathbf{C}_{\bar{c}o} & \mathbf{D} \end{array}\right)$ is observable.

Practical cases of uncontrollability and unobservability may appear in pole(s) / zero(s) cancellation in series interconnection as represented in Figure 3.2 and Figure 3.3.

In the special case where matrix $\mathbf{A}$ has distinct eigenvalues then matrices $\mathbf{A}_{ij} = \mathbf{0}\ \forall\ i \neq j$ and Kalman decomposition reduces as follows:

$$\left(\begin{array}{c|c} \mathbf{A}_K & \mathbf{B}_K \\ \hline \mathbf{C}_K & \mathbf{D} \end{array}\right) = \left(\begin{array}{cccc|c} \mathbf{A}_{c\bar{o}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{c\bar{o}} \\ \mathbf{0} & \mathbf{A}_{co} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{co} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_{\bar{c}\bar{o}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{\bar{c}o} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{C}_{co} & \mathbf{0} & \mathbf{C}_{\bar{c}o} & \mathbf{D} \end{array}\right) \tag{3.252}$$

Figure 3.4 represents the Kalman decomposition: there is no path, direct or through a block, from the input to either of the uncontrollable blocks. Similarly the unobservable blocks have no path to the output.

The new state representation $\left(\begin{array}{c|c} \mathbf{A}_K & \mathbf{B}_K \\ \hline \mathbf{C}_K & \mathbf{D} \end{array}\right)$ is obtained using the change of basis matrix $\mathbf{P}_K$:

$$\begin{cases} \mathbf{A}_K = \mathbf{P}_K^{-1}\mathbf{A}\mathbf{P}_K \\ \mathbf{B}_K = \mathbf{P}_K^{-1}\mathbf{B} \\ \mathbf{C}_K = \mathbf{C}\mathbf{P}_K \end{cases} \tag{3.253}$$

Let $\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n$ be the eigenvectors of matrix $\mathbf{A}$ and $\underline{w}_1, \underline{w}_2, \cdots, \underline{w}_n$ be the eigenvectors of matrix $\mathbf{A}^T$. The change of basis matrix $\mathbf{P}_K$ is an invertible matrix defined as follows:

$$\mathbf{P}_K = \left[\begin{array}{cccc} \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_n \end{array}\right] = \left[\begin{array}{cccc} \mathbf{P}_{c\bar{o}} & \mathbf{P}_{co} & \mathbf{P}_{\bar{c}\bar{o}} & \mathbf{P}_{\bar{c}o} \end{array}\right] \tag{3.254}$$

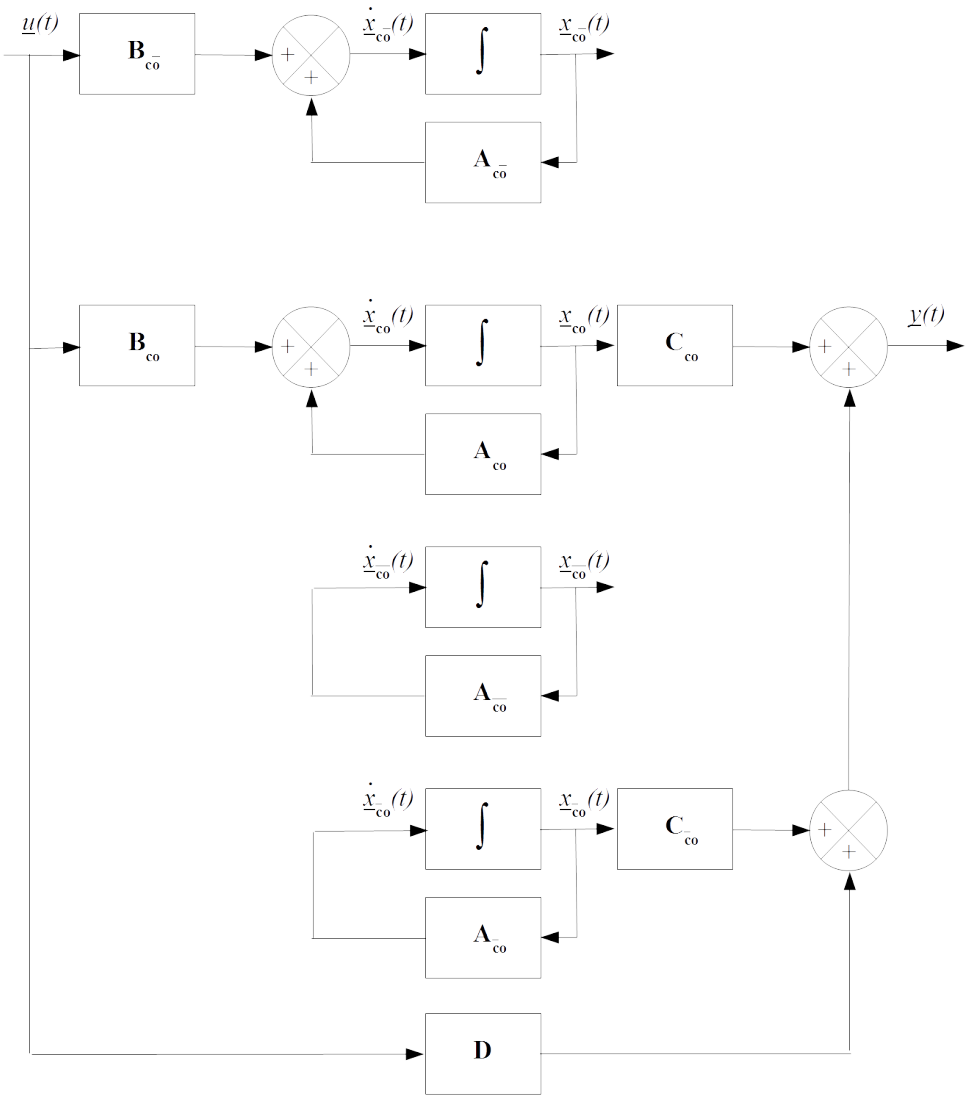

Figure 3.4: Kalman decomposition in the special case where matrix **A** has distinct eigenvalues

Where[10]:

- $\mathbf{P}_{c\bar{o}}$ is a matrix whose columns span the subspace of states which are both controllable and unobservable: $\underline{w}_i^T\mathbf{B} \neq \underline{0}$ and $\mathbf{C}\underline{v}_i = \underline{0}$;
- $\mathbf{P}_{co}$ is chosen so that the columns of $\begin{bmatrix} \mathbf{P}_{c\bar{o}} & \mathbf{P}_{co} \end{bmatrix}$ are a basis for the controllable subspace: $\underline{w}_i^T\mathbf{B} \neq \underline{0}$ and $\mathbf{C}\underline{v}_i \neq \underline{0}$;
- $\mathbf{P}_{\bar{c}\bar{o}}$ is chosen so that the columns of $\begin{bmatrix} \mathbf{P}_{c\bar{o}} & \mathbf{P}_{\bar{c}\bar{o}} \end{bmatrix}$ are a basis for the unobservable subspace: $\underline{w}_i^T\mathbf{B} = \underline{0}$ and $\mathbf{C}\underline{v}_i = \underline{0}$;
- $\mathbf{P}_{\bar{c}o}$ is chosen so that $\mathbf{P}_K$ is invertible: $\underline{w}_i^T\mathbf{B} = \underline{0}$ and $\mathbf{C}\underline{v}_i \neq \underline{0}$.

It is worth noticing that some of those matrices may not exist. For example if the system is both controllable and observable then $\mathbf{P}_K = \mathbf{P}_{co}$; thus other matrices do not exist. In addition Kalman decomposition is more than getting a diagonal form for the state matrix $\mathbf{A}$. When state matrix $\mathbf{A}$ is diagonal observability and controllability have still to be checked using the rank condition test. Finally all realizations obtained from a transfer function are both controllable and observable.

Example 3.17. *Let's consider the following realization:*

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) = \left(\begin{array}{cc|c} -1 & 10 & -2 \\ 0 & 1 & 0 \\ \hline -2 & 3 & -2 \end{array}\right) \tag{3.255}$$

Matrix $\mathbf{A}$ *has a stable mode, which is* -1*, and an unstable mode, which is* 1*. When computing the transfer function of the system we can observe a pole / zero cancellation of the unstable mode:*

$$\begin{aligned} F(s) &= \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} \\ &= \begin{bmatrix} -2 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{s+1} & \frac{10}{s^2-1} \\ 0 & \frac{1}{s-1} \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix} - 2 \\ &= \frac{4}{s+1} - 2 \\ &= \frac{-2s+2}{s+1} \end{aligned} \tag{3.256}$$

From PBH test it can be checked that mode -1 *is both controllable and observable whereas mode* 1 *is observable but not controllable. Thus the system is not stabilizable.*

Internal stability (which implies input-output stability, or BIBO stability) is required in practice. This cannot be achieved unless the plant is both detectable and stabilizable.

■

[10] https://en.wikipedia.org/wiki/Kalman_decomposition

3.12 Minimal realization (again!)

We have seen that a realization of a transfer function $\mathbf{F}(s)$ is *minimal* if there exists no realization of lesser order whose transfer function is $\mathbf{F}(s)$. The order of a realization is the size of matrix $\mathbf{A}$.

A realization is said to be minimal if and only if the realization is both controllable and observable. Consequently a minimal realization of the transfer function $\mathbf{F}(s) = \mathbf{C}\,(s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B} + \mathbf{D}$ is $\mathbf{C}_{co}\,(s\mathbb{I} - \mathbf{A}_{co})^{-1}\,\mathbf{B}_{co} + \mathbf{D}$:

$$\mathbf{F}(s) = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{A}_{co} & \mathbf{B}_{co} \\ \hline \mathbf{C}_{co} & \mathbf{D} \end{array}\right) \tag{3.257}$$

Indeed:

$$\begin{aligned}
(s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B} &= \begin{bmatrix} s\mathbb{I} - \mathbf{A}_{c\bar{o}} & -\mathbf{A}_{12} & -\mathbf{A}_{13} & -\mathbf{A}_{14} \\ \mathbf{0} & s\mathbb{I} - \mathbf{A}_{co} & \mathbf{0} & -\mathbf{A}_{24} \\ \mathbf{0} & \mathbf{0} & s\mathbb{I} - \mathbf{A}_{\bar{c}\bar{o}} & -\mathbf{A}_{34} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & s\mathbb{I} - \mathbf{A}_{\bar{c}o} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{B}_{c\bar{o}} \\ \mathbf{B}_{co} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \\
&= \begin{bmatrix} (s\mathbb{I} - \mathbf{A}_{c\bar{o}})^{-1} & * & * & * \\ \mathbf{0} & (s\mathbb{I} - \mathbf{A}_{co})^{-1} & * & * \\ \mathbf{0} & \mathbf{0} & (s\mathbb{I} - \mathbf{A}_{\bar{c}\bar{o}})^{-1} & * \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & (s\mathbb{I} - \mathbf{A}_{\bar{c}o})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{c\bar{o}} \\ \mathbf{B}_{co} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \\
&= \begin{bmatrix} * \\ (s\mathbb{I} - \mathbf{A}_{co})^{-1}\,\mathbf{B}_{co} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}
\end{aligned} \tag{3.258}$$

And:

$$\begin{aligned}
\mathbf{F}(s) &= \mathbf{C}\,(s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B} + \mathbf{D} \\
&= \begin{bmatrix} \mathbf{0} & \mathbf{C}_{co} & \mathbf{0} & \mathbf{C}_{\bar{c}o} \end{bmatrix} \begin{bmatrix} * \\ (s\mathbb{I} - \mathbf{A}_{co})^{-1}\,\mathbf{B}_{co} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} + \mathbf{D} \\
&= \mathbf{C}_{co}\,(s\mathbb{I} - \mathbf{A}_{co})^{-1}\,\mathbf{B}_{co} + \mathbf{D} \\
&= \left(\begin{array}{c|c} \mathbf{A}_{co} & \mathbf{B}_{co} \\ \hline \mathbf{C}_{co} & \mathbf{D} \end{array}\right)
\end{aligned} \tag{3.259}$$

The number of states of a minimal realization can be evaluated by the rank of the product of the observability and the controllability matrix [11].

[11] Albertos P., Sala A., Multivariable Control Systems, Springer, p78

Chapter 4

Observer design

4.1 Introduction

The components of the state vector $\underline{x}$ may not be fully available as measurements. Observers are used in order to estimate state variables of a dynamical system, which will be denoted $\hat{\underline{x}}$ in the following, from the output signal $\underline{y}(t)$ and the input signal $\underline{u}(t)$ as depicted on Figure 4.1.

Several methods may be envisioned to reconstruct the state vector $\underline{x}(t)$ of a system from the observation of its output signal $\underline{y}(t)$ and the knowledge of the input signal $\underline{u}(t)$:

- From the output equation $\underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t)$ we can imagine to build $\underline{x}(t)$ from the relation $\underline{x}(t) = \mathbf{C}^{-1}\left(\underline{y}(t) - \mathbf{D}\underline{u}(t)\right)$. Unfortunately this relation holds as soon as matrix $\mathbf{C}$ is square and invertible, which is seldom the case;

- Assuming that the size of the state vector is n we may also imagine to take the derivative of the output signal $n-1$ times and use the state equation $\dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t)$ to get n equations where the state vector $\underline{x}(t)$ is the unknown. Unfortunately this not possible in practice because each derivative of an unsmoothed signal increases its noise ;

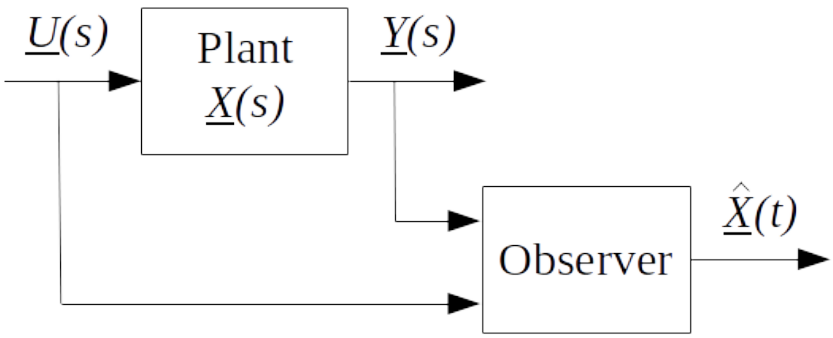

Figure 4.1: Observer principle

– We can also use a Luenberger observer which will be developed in the next section. David Gilbert Luenberger, born in 1937 in Los Angeles, is an American mathematical scientist[1]. The theory of full order observer originated in the work of Luenberger in 1964.

This chapter focuses on observers design. More specifically Luenberger observer, state observer for SISO systems in observable canonical form, state observer for SISO systems in arbitrary state-space representation and state observer for MIMO systems will be presented. We will also present reduced-order observer design.

4.2 Luenberger observer

4.2.1 State observer

Consider the following state space representation where $\underline{x}(t)$ denotes the state vector, $\underline{y}(t)$ the measured output and $\underline{u}(t)$ the control input:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{4.1}$$

We assume that state vector $\underline{x}(t)$ cannot be measured. The purpose of the state observer is to estimate $\underline{x}(t)$ based on the measured output $\underline{y}(t)$. Luenberger observer provides an estimation $\hat{\underline{x}}(t)$ of the state vector $\underline{x}(t)$ through the following differential equation where output signal $\underline{y}(t)$ and input signal $\underline{u}(t)$ are known and where matrices $\mathbf{F}$, $\mathbf{J}$ and $\mathbf{L}$ have to be determined:

$$\dot{\hat{\underline{x}}}(t) = \mathbf{F}\hat{\underline{x}}(t) + \mathbf{J}\underline{u}(t) + \mathbf{L}\underline{y}(t) \tag{4.2}$$

The estimation error $\underline{e}(t)$ is defined as follows:

$$\underline{e}(t) = \underline{x}(t) - \hat{\underline{x}}(t) \tag{4.3}$$

The time derivative of the estimation error reads:

$$\begin{aligned} \dot{\underline{e}}(t) &= \dot{\underline{x}}(t) - \dot{\hat{\underline{x}}}(t) \\ &= \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) - \mathbf{F}\hat{\underline{x}}(t) - \mathbf{J}\underline{u}(t) - \mathbf{L}\underline{y}(t) \end{aligned} \tag{4.4}$$

Thanks to the output equation $\underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t)$ and the relation $\underline{x}(t) = \underline{e}(t) + \hat{\underline{x}}(t)$ we get:

$$\begin{aligned} \dot{\underline{e}}(t) &= \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) - \mathbf{F}\hat{\underline{x}}(t) - \mathbf{J}\underline{u}(t) - \mathbf{L}\left(\mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t)\right) \\ &= \left(\mathbf{A} - \mathbf{LC}\right)\underline{x}(t) + \left(\mathbf{B} - \mathbf{J} - \mathbf{LD}\right)\underline{u}(t) - \mathbf{F}\hat{\underline{x}}(t) \\ &= \left(\mathbf{A} - \mathbf{LC}\right)\underline{e}(t) + \left(\mathbf{B} - \mathbf{J} - \mathbf{LD}\right)\underline{u}(t) + \left(\mathbf{A} - \mathbf{LC} - \mathbf{F}\right)\hat{\underline{x}}(t) \end{aligned} \tag{4.5}$$

As soon as the purpose of the observer is to move the estimation error $\underline{e}(t)$ towards zero independently of control $\underline{u}(t)$ and true state vector $\underline{x}(t)$ we choose matrices $\mathbf{J}$ and $\mathbf{F}$ as follows:

[1] https://en.wikipedia.org/wiki/David_Luenberger

$$\begin{cases} \mathbf{J} = \mathbf{B} - \mathbf{LD} \\ \mathbf{F} = \mathbf{A} - \mathbf{LC} \end{cases} \tag{4.6}$$

Thus the dynamics of the estimation error $\underline{e}(t)$ reduces to be:

$$\dot{\underline{e}}(t) = (\mathbf{A} - \mathbf{LC})\,\underline{e}(t) \tag{4.7}$$

In order that the estimation error $\underline{e}(t)$ moves towards zero, meaning that the estimated state vector $\hat{\underline{x}}$ becomes equal to the actual state vector $\underline{x}(t)$, matrix $\mathbf{L}$ shall be chosen such that all the eigenvalues of $\mathbf{A} - \mathbf{LC}$ are situated in the left half plane.

With the expression of matrices $\mathbf{J}$ and $\mathbf{F}$ the dynamics of the Luenberger observer can be written as follows:

$$\begin{aligned} \dot{\hat{\underline{x}}}(t) &= \mathbf{F}\hat{\underline{x}}(t) + \mathbf{J}\underline{u}(t) + \mathbf{L}\underline{y}(t) \\ &= (\mathbf{A} - \mathbf{LC})\,\hat{\underline{x}}(t) + (\mathbf{B} - \mathbf{LD})\,\underline{u}(t) + \mathbf{L}\underline{y}(t) \end{aligned} \tag{4.8}$$

That is:

$$\dot{\hat{\underline{x}}}(t) = \mathbf{A}\hat{\underline{x}}(t) + \mathbf{B}\underline{u}(t) + \mathbf{L}\left(\underline{y}(t) - \hat{\underline{y}}(t)\right) \tag{4.9}$$

Where:

$$\hat{\underline{y}}(t) = \mathbf{C}\hat{\underline{x}}(t) + \mathbf{D}\underline{u}(t) \tag{4.10}$$

Thus the dynamics of the Luenberger observer is the same than the dynamics of the original system with the additional term $\mathbf{L}\left(\underline{y}(t) - \hat{\underline{y}}(t)\right)$ where $\mathbf{L}$ is a gain to be set. This additional term is proportional to the error $\underline{y}(t) - \hat{\underline{y}}(t)$. It enables to drive the estimated state $\hat{\underline{x}}(t)$ towards its actual value $\underline{x}(t)$ when the measured output $\underline{y}(t)$ deviates from the estimated output $\hat{\underline{y}}(t)$.

In order to compute a state space representation and the transfer function of the observer we first identify its input and output.

- As discussed previously the input vector $\underline{u}_o(t)$ of the observer is composed of the output $\underline{y}(t)$ of the plant whose state is estimated and its input $\underline{u}(t)$:

$$\underline{u}_o(t) = \begin{bmatrix} \underline{y}(t) \\ \underline{u}(t) \end{bmatrix} \tag{4.11}$$

- The output $\underline{y}_o(t)$ of the observer is the estimated state vector $\hat{\underline{x}}(t)$ of the plant:

$$\underline{y}_o(t) = \hat{\underline{x}}(t) \tag{4.12}$$

Consequently (4.9) and (4.10) can be organized to obtain a state space representation of the observer:

$$\begin{cases} \dot{\hat{\underline{x}}}(t) &= \mathbf{A}\hat{\underline{x}}(t) + \mathbf{B}\underline{u}(t) + \mathbf{L}\left(\underline{y}(t) - \hat{\underline{y}}(t)\right) \\ &= \mathbf{A}\hat{\underline{x}}(t) + \mathbf{B}\underline{u}(t) + \mathbf{L}\left(\underline{y}(t) - \mathbf{C}\hat{\underline{x}}(t) - \mathbf{D}\underline{u}(t)\right) \\ &= (\mathbf{A} - \mathbf{LC})\,\hat{\underline{x}}(t) + \begin{bmatrix} \mathbf{L} & \mathbf{B} - \mathbf{LD} \end{bmatrix} \begin{bmatrix} \underline{y}(t) \\ \underline{u}(t) \end{bmatrix} \\ &= (\mathbf{A} - \mathbf{LC})\,\hat{\underline{x}}(t) + \begin{bmatrix} \mathbf{L} & \mathbf{B} - \mathbf{LD} \end{bmatrix} \underline{u}_o(t) \\ &:= \mathbf{A}_{obs}\hat{\underline{x}}(t) + \mathbf{B}_{obs}\underline{u}_o(t) \\ \underline{y}_o(t) &= \hat{\underline{x}}(t) \\ &= \mathbb{I}\hat{\underline{x}}(t) + \mathbf{0}\underline{u}_o(t) \\ &:= \mathbf{C}_{obs}\hat{\underline{x}}(t) + \mathbf{D}_{obs}\underline{u}_o(t) \end{cases} \tag{4.13}$$

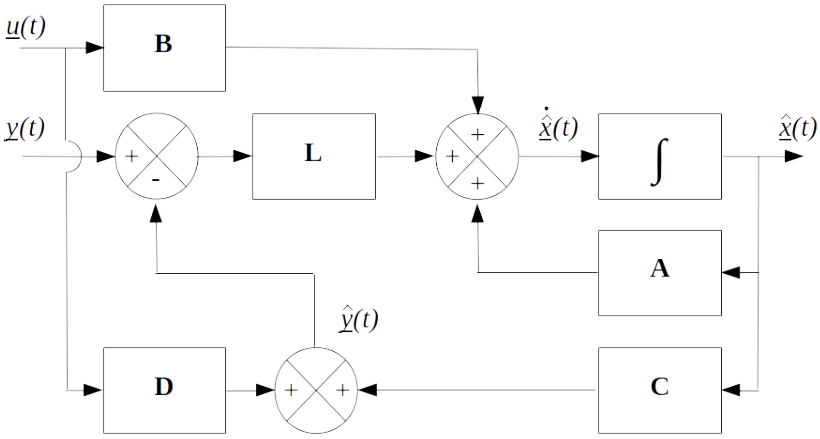

Figure 4.2: Block diagram of the state-space representation of an observer

It is worth noticing that the state matrix of the observer reads $(\mathbf{A} - \mathbf{LC})$. Setting the eigenvalues of $(\mathbf{A} - \mathbf{LC})$ through the observer gain $\mathbf{L}$ is equivalent by duality to set the eigenvalues of matrix $(\mathbf{A} - \mathbf{LC})^T = \mathbf{A}^T - \mathbf{C}^T\mathbf{K}$ through the state feedback gain $\mathbf{K}$. Observer gain $\mathbf{L}$ is related to state feedback $\mathbf{K}$ as follows:

$$\mathbf{L} := \mathbf{K}^T \tag{4.14}$$

Finally the transfer function of the observer, which is obviously a Multi-Input Multi-Output (MIMO) system, reads:

$$\begin{aligned} \mathbf{G}_{obs}(s) &= \mathbf{C}_{obs}\left(s\mathbb{I} - \mathbf{A}_{obs}\right)^{-1}\mathbf{B}_{obs} + \mathbf{D}_{obs} \\ &= \left(s\mathbb{I} - (\mathbf{A} - \mathbf{LC})\right)^{-1}\left[\begin{array}{cc} \mathbf{L} & \mathbf{B} - \mathbf{LD} \end{array}\right] \end{aligned} \tag{4.15}$$

The block diagram corresponding to state-space representation (4.9) is shown in Figure 4.2.

In the following we will assume that the system is observable, or at least detectable, such that it is possible to design a state observer.

4.2.2 Estimating output derivative

Assume that the plant has no feedforward matrix ($\mathbf{D} = \mathbf{0}$) and that the following state observer has been designed:

$$\mathbf{D} = \mathbf{0} \Rightarrow \dot{\underline{\hat{x}}}(t) = (\mathbf{A} - \mathbf{LC})\,\underline{\hat{x}}(t) + \left[\begin{array}{cc} \mathbf{L} & \mathbf{B} \end{array}\right]\left[\begin{array}{c} \underline{y}(t) \\ \underline{u}(t) \end{array}\right] \tag{4.16}$$

We wish to estimate the derivative $\dot{\underline{\hat{y}}}(t)$ of the actual output $\underline{y}(t)$ of the plant from the output $\underline{\hat{x}}(t)$ of the observer. We get:

$$\begin{aligned} \underline{y}(t) = \mathbf{C}\underline{x}(t) \Rightarrow \dot{\underline{y}}(t) &= \mathbf{C}\dot{\underline{x}}(t) \\ &= \mathbf{C}\left(\mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t)\right) \\ &= \mathbf{CA}\underline{x}(t) + \mathbf{CB}\underline{u}(t) \end{aligned} \tag{4.17}$$

We finally get:

$$\dot{\underline{\hat{y}}}(t) = \mathbf{CA}\underline{\hat{x}}(t) + \mathbf{CB}\underline{u}(t) \tag{4.18}$$

4.2.3 Output observer

Consider the following state space representation where $\underline{x}(t)$ denotes the state vector, $\underline{y}(t)$ the measured output, $\underline{z}(t)$ the performance output and $\underline{u}(t)$ the control input:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \\ \underline{z}(t) = \mathbf{T}\underline{x}(t) \end{cases} \tag{4.19}$$

We assume that the performance output $\underline{z}(t)$ cannot be measured. The purpose of the output observer is to estimate the performance output $\underline{z}(t)$ based on the measured output $\underline{y}(t)$. Luenberger observer provides an estimation $\hat{\underline{z}}(t)$ of the performance output $\underline{z}(t)$ through the following differential equation where output signal $\underline{y}(t)$ and input signal $\underline{u}(t)$ are known and where matrices $\mathbf{F}$, $\mathbf{J}$ and $\mathbf{L}$ have to be determined:

$$\dot{\hat{\underline{z}}}(t) = \mathbf{F}\hat{\underline{z}}(t) + \mathbf{TB}\underline{u}(t) + \mathbf{L}\underline{y}(t) \tag{4.20}$$

The estimation error $\underline{e}(t)$ is defined as follows:

$$\underline{e}(t) = \underline{z}(t) - \hat{\underline{z}}(t) \tag{4.21}$$

Thanks to equation $\underline{z}(t) = \mathbf{T}\underline{x}(t)$ and the relation $\underline{z}(t) = \underline{e}(t) + \hat{\underline{z}}(t)$ the time derivative of the estimation error reads:

$$\begin{aligned} \dot{\underline{e}}(t) &= \dot{\underline{z}}(t) - \dot{\hat{\underline{z}}}(t) \\ &= \mathbf{T}\left(\mathbf{A}\underline{x}(t) + \cancel{\mathbf{B}\underline{u}(t)}\right) - \mathbf{F}\hat{\underline{z}}(t) - \cancel{\mathbf{TB}\underline{u}(t)} - \mathbf{L}\underline{y}(t) \\ &= \mathbf{TA}\underline{x}(t) - \mathbf{F}\hat{\underline{z}}(t) - \mathbf{L}\underline{y}(t) \\ &= \mathbf{TA}\underline{x}(t) - \mathbf{F}\left(\underline{z}(t) - \underline{e}(t)\right) - \mathbf{L}\underline{y}(t) \\ &= \left(\mathbf{TA} - \mathbf{FT} - \mathbf{LC}\right)\underline{x}(t) + \mathbf{F}\underline{e}(t) \end{aligned} \tag{4.22}$$

Thus, as soon as the following relation holds:

$$\mathbf{LC} = \mathbf{TA} - \mathbf{FT} \tag{4.23}$$

We get:

$$\dot{\underline{e}}(t) = \mathbf{F}\underline{e}(t) \Rightarrow \underline{e}(t) = e^{\mathbf{F}t}\,\underline{e}(0) \tag{4.24}$$

In addition, if $\mathbf{F}$ is stable, we get:

$$\lim_{t\to\infty} \underline{e}(t) = \underline{0} \tag{4.25}$$

Moreover, if we wish to estimate the state vector $\underline{x}(t)$ through $\hat{\underline{z}}(t)$ and the actual output $\underline{y}(t)$ we can use matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ such that:

$$\begin{aligned} \hat{\underline{x}}(t) &= \mathbf{M}_1\hat{\underline{z}}(t) + \mathbf{M}_2\underline{y}(t) \\ &= \mathbf{M}_1\left(\underline{z}(t) - \underline{e}(t)\right) + \mathbf{M}_2\underline{y}(t) \\ &= \left(\mathbf{M}_1\mathbf{T} + \mathbf{M}_2\mathbf{C}\right)\underline{x}(t) - \mathbf{M}_1\underline{e}(t) \end{aligned} \tag{4.26}$$

Then choosing $\mathbf{M}_1$ and $\mathbf{M}_2$ such that:

$$\mathbf{M}_1\mathbf{T} + \mathbf{M}_2\mathbf{C} = \mathbb{I} \tag{4.27}$$

We get:

$$\underline{\hat{x}}(t) = \underline{x}(t) - \mathbf{M}_1 \underline{e}(t) \tag{4.28}$$

And, thanks to (4.25):

$$\lim_{t \to \infty} \underline{\hat{x}}(t) = \underline{x}(t) \tag{4.29}$$

Matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ can be computed as follows assuming that matrix $\begin{bmatrix} \mathbf{T} \\ \mathbf{C} \end{bmatrix}$ is full rank[2]:

$$\begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{T} \\ \mathbf{C} \end{bmatrix}^{-1} \tag{4.30}$$

If matrix $\mathbf{T}$ is full rank, then we can choose $\mathbf{M}_2 = \mathbf{0}$. Then we get for the dynamics of the estimator $\underline{\hat{x}}(t)$ of state $\underline{x}(t)$:

$$\begin{aligned} &\begin{cases} \dot{\underline{\hat{z}}}(t) = \mathbf{F}\underline{\hat{z}}(t) + \mathbf{TB}\underline{u}(t) + \mathbf{L}\underline{y}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \\ \underline{\hat{z}}(t) = \mathbf{T}\underline{\hat{x}}(t) \end{cases} \\ &\Rightarrow \dot{\underline{\hat{x}}}(t) = \mathbf{T}^{-1}\dot{\underline{\hat{z}}}(t) = \mathbf{T}^{-1}\left(\mathbf{FT}\underline{\hat{x}}(t) + \mathbf{TB}\underline{u}(t) + \mathbf{L}\underline{y}(t)\right) \end{aligned} \tag{4.31}$$

We finally retrieve the expression of the state observer by using (4.23). Indeed, for the general case where $\mathbf{T} \neq \mathbb{I}$ we get:

$$\begin{aligned} \mathbf{FT} = \mathbf{TA} - \mathbf{LC} \Rightarrow \dot{\underline{\hat{x}}}(t) &= \mathbf{T}^{-1}\left((\mathbf{TA} - \mathbf{LC})\,\underline{\hat{x}}(t) + \mathbf{TB}\underline{u}(t) + \mathbf{L}\underline{y}(t)\right) \\ &= \mathbf{A}\underline{\hat{x}}(t) + \mathbf{B}\underline{u}(t) + \mathbf{T}^{-1}\mathbf{L}\left(\underline{y}(t) - \mathbf{C}\underline{\hat{x}}(t)\right) \end{aligned} \tag{4.32}$$

4.3 State observer for SISO systems in observable canonical form

Let $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$ be an observable Single-Input Single-Output (SISO) linear time-invariant system of order n and let $\chi_{A-LC}(s)$ be an imposed n^{th} order monic polynomial (a monic polynomial is a polynomial in which the leading coefficient, that is the nonzero coefficient of highest degree, is equal to 1). Polynomial $\chi_{A-LC}(s)$ corresponds to the characteristic polynomial of matrix $\mathbf{A} - \mathbf{LC}$. It is formed thanks to the predefined eigenvalues $\lambda_{L1}, \cdots, \lambda_{Ln}$ assigned for the dynamics of the observer:

$$\chi_{A-LC}(s) = \det\left(s\mathbb{I} - (\mathbf{A} - \mathbf{LC})\right) = (s - \lambda_{L1}) \cdots (s - \lambda_{Ln}) \tag{4.33}$$

When expanding the preceding product we get:

$$\chi_{A-LC}(s) = s^n + p_{n-1}s^{n-1} + \cdots + p_1 s + p_0 \tag{4.34}$$

[2]Daniel Alazard, Pierre Apkarian, Christelle Cumer, Gilles Ferreres, Michel Gauvrit, Robustesse et Commande Optimale, Cepadues, 1999

We begin with the case where the system realization $\left(\begin{array}{c|c}\mathbf{A} & \mathbf{B}\\ \hline \mathbf{C} & \mathbf{D}\end{array}\right)$ is the observable canonical form. Then matrices $\mathbf{A}_o$ and $\mathbf{C}_o$ are the following:

$$\left\{\begin{array}{l} \mathbf{A}_o = \begin{bmatrix} 0 & 0 & 0 & & -a_0 \\ 1 & 0 & 0 & \ddots & -a_1 \\ 0 & 1 & 0 & \ddots & -a_2 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & & 0 & 1 & -a_{n-1} \end{bmatrix} \\ \\ \mathbf{C}_o = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \end{array}\right. \tag{4.35}$$

Matrices $\mathbf{B}_o$ and $\mathbf{D}$ will not be used.

Let $\mathbf{L}_o$ be the observer gain matrix when the observable canonical form of the system is used. For a SISO system this is a $n \times 1$ matrix. Let L_i be the (scalar) component in the ith row of matrix $\mathbf{L}_o$:

$$\mathbf{L}_o = \begin{bmatrix} L_1 \\ \vdots \\ L_n \end{bmatrix} \tag{4.36}$$

Then matrix $\mathbf{A}_o - \mathbf{L}_o\mathbf{C}_o$ reads:

$$\begin{aligned} \mathbf{A}_o - \mathbf{L}_o\mathbf{C}_o &= \begin{bmatrix} 0 & 0 & 0 & & -a_0 \\ 1 & 0 & 0 & \ddots & -a_1 \\ 0 & 1 & 0 & \ddots & -a_2 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & & 0 & 1 & -a_{n-1} \end{bmatrix} - \begin{bmatrix} L_1 \\ \vdots \\ L_n \end{bmatrix} \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & & -a_0 - L_1 \\ 1 & 0 & 0 & \ddots & -a_1 - L_2 \\ 0 & 1 & 0 & \ddots & -a_2 - L_3 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & & 0 & 1 & -a_{n-1} - L_n \end{bmatrix} \end{aligned} \tag{4.37}$$

Since this matrix still remains in the observable canonical form its characteristic polynomial is readily written as follows:

$$\begin{aligned} \chi_{A-LC}(s) &= \det(\mathbf{A} - \mathbf{L}\mathbf{C}) \\ &= \det(\mathbf{A}_o - \mathbf{L}_o\mathbf{C}_o) \\ &= s^n + (a_{n-1} + L_n)\, s^{n-1} + \cdots + (a_1 + L_2)\, s + a_0 + L_1 \end{aligned} \tag{4.38}$$

Identifying Equations (4.34) and (4.38) leads to the expression of each

component of the observer matrix $\mathbf{L}_o$:

$$\begin{cases} p_0 = a_0 + L_1 \\ p_1 = a_1 + L_2 \\ \vdots \\ p_{n-1} = a_{n-1} + L_n \end{cases} \Leftrightarrow \mathbf{L}_o = \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{bmatrix} = \begin{bmatrix} p_0 - a_0 \\ p_1 - a_1 \\ \vdots \\ p_{n-1} - a_{n-1} \end{bmatrix} \tag{4.39}$$

4.4 State observer for SISO systems in arbitrary state-space representation

When an arbitrary state-space representation is used the system has to be converted into the observable canonical form via a similarity transformation. Let $\mathbf{P}_o$ be the matrix of the similarity transformation which enables to get the observable canonical form. We get:

$$\underline{\hat{x}}(t) = \mathbf{P}_o \underline{\hat{x}}_o(t) \Leftrightarrow \underline{\hat{x}}_o(t) = \mathbf{P}_o^{-1} \underline{\hat{x}}(t) \tag{4.40}$$

We have seen in the chapter dedicated to *Realization of transfer functions* that $\mathbf{P}_o$ is a constant nonsingular change of basis matrix which is obtained through the following relation:

$$\mathbf{P}_o = \mathbf{Q}_o^{-1} \mathbf{Q}_{oo} \tag{4.41}$$

Where:

- $\mathbf{Q}_o$ is the observability matrix in the actual basis:

$$\mathbf{Q}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \tag{4.42}$$

- and $\mathbf{Q}_{oo}$ the observability matrix expressed in the observable canonical basis (which is readily obtained through $\det(s\mathbb{I} - \mathbf{A})$).

Thus the state equation of the observer (4.9) reads:

$$\begin{aligned} &\dot{\underline{\hat{x}}}_o(t) = \mathbf{A}_o \underline{\hat{x}}_o(t) + \mathbf{B}_o \underline{u}(t) + \mathbf{L}_o \left(\underline{y}(t) - \underline{\hat{y}}(t)\right) \\ &\Leftrightarrow \mathbf{P}_o^{-1} \dot{\underline{\hat{x}}}(t) = \mathbf{A}_o \mathbf{P}_o^{-1} \underline{\hat{x}}(t) + \mathbf{B}_o \underline{u}(t) + \mathbf{L}_o \left(\underline{y}(t) - \underline{\hat{y}}(t)\right) \\ &\Leftrightarrow \dot{\underline{\hat{x}}}(t) = \mathbf{P}_o \mathbf{A}_o \mathbf{P}_o^{-1} \underline{\hat{x}}(t) + \mathbf{P}_o \mathbf{B}_o \underline{u}(t) + \mathbf{P}_o \mathbf{L}_o \left(\underline{y}(t) - \underline{\hat{y}}(t)\right) \end{aligned} \tag{4.43}$$

That is:

$$\dot{\underline{\hat{x}}}(t) = \mathbf{A} \underline{\hat{x}}(t) + \mathbf{B} \underline{u}(t) + \mathbf{L} \left(\underline{y}(t) - \underline{\hat{y}}(t)\right) \tag{4.44}$$

Where:

$$\mathbf{L} = \mathbf{P}_o \mathbf{L}_o \tag{4.45}$$

And:

$$\begin{cases} \mathbf{A} = \mathbf{P}_o \mathbf{A}_o \mathbf{P}_o^{-1} \\ \mathbf{B} = \mathbf{P}_o \mathbf{B}_o \end{cases} \tag{4.46}$$

Example 4.1. *Design an observer for the following plant:*

$$\begin{cases} \dot{\underline{x}}(t) = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u(t) \\ \\ y(t) = \begin{bmatrix} 3 & 5 \end{bmatrix} \underline{x}(t) \end{cases} \tag{4.47}$$

As far as this is a modal (or diagonal) state space representation, plant's observability is readily checked by inspecting row of output matrix $\mathbf{C}$*: because there is no null element in the output matrix* $\mathbf{C}$ *we conclude that the plant is observable by applying Gilbert's criteria.*

The poles of the observer shall be chosen faster than the dynamics of the plant, whose modes are -1 *and* -2*. We choose (for example) to locate the poles of the observer at* $\lambda_{L1} = -10$ *and* $\lambda_{L2} = -20$*.*

We will first design the observer assuming that we have the observable canonical form of the SISO system. The observable canonical form is readily obtained through $\det(s\mathbb{I} - \mathbf{A})$*:*

$$\begin{aligned} &\det(s\mathbb{I} - \mathbf{A}) = (s+1)(s+2) = s^2 + 3\,s + 2 := s^2 + a_1\,s + a_0 \\ &\Rightarrow \mathbf{A}_o = \begin{bmatrix} 0 & -2 \\ 1 & -3 \end{bmatrix} \text{ and } \mathbf{C}_o = \begin{bmatrix} 0 & 1 \end{bmatrix} \end{aligned} \tag{4.48}$$

On the other hand the characteristic polynomial of the observer is formed thanks to the predefined eigenvalues assigned for the dynamics of the observer:

$$\begin{aligned} \chi_{A-LC}(s) &= (s - \lambda_{L1})\,(s - \lambda_{L2}) = (s+10)\,(s+20) \\ &= s^2 + 30\,s + 200 := s^2 + p_1\,s + p_0 \end{aligned} \tag{4.49}$$

Applying relation (4.39) *we get:*

$$\mathbf{L}_o = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} = \begin{bmatrix} p_0 - a_0 \\ p_1 - a_1 \end{bmatrix} = \begin{bmatrix} 200 - 2 \\ 30 - 3 \end{bmatrix} = \begin{bmatrix} 198 \\ 27 \end{bmatrix} \tag{4.50}$$

Now let's compute the similarity transformation matrix $\mathbf{P}_o$ *which enables to get the observable canonical form.*

$$\mathbf{P}_o = \mathbf{Q}_o^{-1}\mathbf{Q}_{oo} \tag{4.51}$$

Where:

- $\mathbf{Q}_o$ *is the observability matrix in the actual basis:*

$$\mathbf{Q}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \end{bmatrix} = \left[\begin{array}{cc} 3 & 5 \\ \hline \begin{bmatrix} 3 & 5 \end{bmatrix} & \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \end{array}\right] = \begin{bmatrix} 3 & 5 \\ -3 & -10 \end{bmatrix} \tag{4.52}$$

- *and* $\mathbf{Q}_{oo}$ *the observability matrix expressed in the observable canonical basis (which is readily obtained through* $\det(s\mathbb{I} - \mathbf{A})$*):*

$$\mathbf{Q}_{oo} = \begin{bmatrix} \mathbf{C}_o \\ \mathbf{C}_o\mathbf{A}_o \end{bmatrix} = \left[\begin{array}{cc} 0 & 1 \\ \hline \begin{bmatrix} 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & -2 \\ 1 & -3 \end{bmatrix} \end{array}\right] = \begin{bmatrix} 0 & 1 \\ 1 & -3 \end{bmatrix} \tag{4.53}$$

Thus:

$$\begin{aligned}\mathbf{P}_o &= \mathbf{Q}_o^{-1}\mathbf{Q}_{oo} = \begin{bmatrix} 3 & 5 \\ -3 & -10 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 \\ 1 & -3 \end{bmatrix} \\ &= \tfrac{-1}{15} \begin{bmatrix} -10 & -5 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -3 \end{bmatrix} \\ &= \tfrac{-1}{15} \begin{bmatrix} -5 & 5 \\ 3 & -6 \end{bmatrix} = \tfrac{1}{15} \begin{bmatrix} 5 & -5 \\ -3 & 6 \end{bmatrix}\end{aligned} \tag{4.54}$$

We finally get:

$$\mathbf{L} = \mathbf{P}_o\mathbf{L}_o = \frac{1}{15} \begin{bmatrix} 5 & -5 \\ -3 & 6 \end{bmatrix} \begin{bmatrix} 198 \\ 27 \end{bmatrix} = \begin{bmatrix} 57 \\ -28.8 \end{bmatrix} \tag{4.55}$$

■

4.5 Ackermann's formula

Ackermann's formula states that the observer gain matrix $\mathbf{L}$ of a SISO system in arbitrary state-space representation can be obtained as the product between the assigned characteristic polynomial $\chi_{A-LC}(s)$ of matrix $\mathbf{A} - \mathbf{LC}$ evaluated at matrix $\mathbf{A}$ and vector $\underline{q}_o$:

$$\mathbf{L} = \chi_{A-LC}(\mathbf{A})\underline{q}_o \tag{4.56}$$

To get this result we first recall that similarity transformation generates equivalent state-space representations. Let $\mathbf{P}_o$ be the matrix of the similarity transformation which enables to get the observable canonical form. Starting from a state-space representations $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$ in an arbitrary basis, the observable canonical form is obtained through the following relations:

$$\begin{cases} \mathbf{A}_o = \mathbf{P}_o^{-1}\mathbf{A}\mathbf{P}_o \\ \mathbf{B}_o = \mathbf{P}_o^{-1}\mathbf{B} \\ \mathbf{C}_o = \mathbf{C}\mathbf{P}_o \end{cases} \tag{4.57}$$

Consequently matrix $\mathbf{A}_o - \mathbf{L}_o\mathbf{C}_o$ reads:

$$\begin{aligned}\mathbf{A}_o - \mathbf{L}_o\mathbf{C}_o &= \mathbf{P}_o^{-1}\mathbf{A}\mathbf{P}_o - \mathbf{L}_o\mathbf{C}\mathbf{P}_o \\ &= \mathbf{P}_o^{-1}\left(\mathbf{A} - \mathbf{P}_o\mathbf{L}_o\mathbf{C}\right)\mathbf{P}_o\end{aligned} \tag{4.58}$$

This equation indicates that the observer gain matrix $\mathbf{L}$ in arbitrary state-space representation reads:

$$\mathbf{L} = \mathbf{P}_o\mathbf{L}_o \tag{4.59}$$

We have seen in the chapter dedicated to *Realization of transfer functions* that $\mathbf{P}_o$ is a constant nonsingular change of basis matrix which is obtained through the state matrix $\mathbf{A}$ and vector $\underline{q}_o$:

$$\mathbf{P}_o = \begin{bmatrix} \underline{q}_o & \mathbf{A}\underline{q}_o & \cdots & \mathbf{A}^{n-1}\underline{q}_o \end{bmatrix} \tag{4.60}$$

Vector $\underline{q}_o$ is the last column of the inverse of the observability matrix $\mathbf{Q}_o$:

$$\mathbf{Q}_o^{-1} = \begin{bmatrix} * & \cdots & * & \underline{q}_o \end{bmatrix} \quad \text{where} \quad \mathbf{Q}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \tag{4.61}$$

Then we recall Cayley–Hamilton theorem. Let $\chi_{A_o}(s)$ be the characteristic polynomial of the $n \times n$ matrix $\mathbf{A}_o$. We write the characteristic polynomial $\chi_{A_o}(s)$ of matrix $\mathbf{A}_o$ as follows:

$$\chi_{A_o}(s) := \det(s\mathbb{I} - \mathbf{A}_o) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \tag{4.62}$$

The Cayley–Hamilton theorem states that substituting matrix $\mathbf{A}_o$ for s in the characteristic polynomial $\chi_{A_o}(s)$ of matrix $\mathbf{A}_o$ results in the zero matrix [3]:

$$\chi_{A_o}(\mathbf{A}_o) = \mathbf{0} = \mathbf{A}_o^n + a_{n-1}\mathbf{A}_o^{n-1} + \cdots + a_1\mathbf{A}_o + a_0\mathbb{I} \tag{4.63}$$

Let $\chi_{A-LC}(s)$ be the characteristic polynomial of matrix $\mathbf{A}_o - \mathbf{L}_o\mathbf{C}_o$. We have seen that when predefined eigenvalues $\lambda_{L1}, \cdots, \lambda_{Ln}$ are assigned for the dynamics of the observer the characteristic polynomial of matrix $\mathbf{A}_o - \mathbf{L}_o\mathbf{C}_o$ reads:

$$\begin{aligned} \chi_{A-LC}(s) &= \det\left(s\mathbb{I} - (\mathbf{A}_o - \mathbf{L}_o\mathbf{C}_o)\right) \\ &= (s - \lambda_{L1}) \cdots (s - \lambda_{Ln}) \\ &= s^n + p_{n-1}s^{n-1} + \cdots + p_1 s + p_0 \end{aligned} \tag{4.64}$$

Substituting s for matrix $\mathbf{A}_o$ leads to the following relation:

$$\chi_{A-LC}(\mathbf{A}_o) = \mathbf{A}_o^n + p_{n-1}\mathbf{A}_o^{n-1} + \cdots + p_1\mathbf{A}_o + p_0\mathbb{I} \tag{4.65}$$

Note that $\chi_{A-LC}(\mathbf{A}_o)$ is not equal to $\mathbf{0}$ because $\chi_{A-LC}(s)$ is not the characteristic polynomial of matrix $\mathbf{A}_o$.

Thanks to Equation (4.39) and the relation $p_i = a_i + L_i$ we get:

$$\begin{aligned} \chi_{A-LC}(\mathbf{A}_o) = \mathbf{A}_o^n + (a_{n-1} + L_{n-1})\,\mathbf{A}_o^{n-1} + \cdots \\ + (a_1 + L_1)\,\mathbf{A}_o + (a_0 + L_0)\,\mathbb{I} \end{aligned} \tag{4.66}$$

By subtracting Equations (4.63) to (4.66) we get:

$$\chi_{A-LC}(\mathbf{A}_o) = L_{n-1}\mathbf{A}_o^{n-1} + \cdots + L_1\mathbf{A}_o + L_0\mathbb{I} \tag{4.67}$$

Let $\underline{u}$ be the vector defined by:

$$\underline{u} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{4.68}$$

[3] https://en.wikipedia.org/wiki/Cayley–Hamilton_theorem

Due to the *special* form of matrix $\mathbf{A}_o$ we have:

$$\mathbf{A}_o\underline{u} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}, \quad \mathbf{A}_o^2\underline{u} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}, \quad \cdots, \quad \mathbf{A}_o^{n-1}\underline{u} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{4.69}$$

Thus multiplying Equation (4.67) by $\underline{u}$ leads to the following relation:

$$\chi_{A-LC}(\mathbf{A}_o)\underline{u} = \begin{bmatrix} L_0 \\ L_1 \\ \vdots \\ L_{n-1} \end{bmatrix} = \mathbf{L}_o \tag{4.70}$$

Thus we get the expression of the observer matrix $\mathbf{L}_o$ when we use the observable canonical form.

We multiply Equation (4.70) by $\mathbf{P}_o$ and use the fact that $\mathbf{A}_o^k = \left(\mathbf{P}_o^{-1}\mathbf{A}\mathbf{P}_o\right)^k = \mathbf{P}_o^{-1}\mathbf{A}^k\mathbf{P}_o$ to get the expression of the observer gain matrix $\mathbf{L}$ in arbitrary state space representation:

$$\begin{aligned} \mathbf{L} &= \mathbf{P}_o\mathbf{L}_o \\ &= \mathbf{P}_o\chi_{A-LC}(\mathbf{A}_o)\underline{u} \\ &= \chi_{A-LC}(\mathbf{P}_o\mathbf{A}_o)\underline{u} \\ &= \chi_{A-LC}(\mathbf{P}_o\mathbf{A}_o\mathbf{P}_o^{-1}\mathbf{P}_o)\underline{u} \\ &= \chi_{A-LC}(\mathbf{P}_o\mathbf{A}_o\mathbf{P}_o^{-1})\mathbf{P}_o\underline{u} \\ &= \chi_{A-LC}(\mathbf{A})\mathbf{P}_o\underline{u} \end{aligned} \tag{4.71}$$

Because $\underline{u}$ is the vector defined by $\underline{u} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ we get using (4.60):

$$\mathbf{P}_o\underline{u} = \underline{q}_o \tag{4.72}$$

Consequently Equation (4.71) reduces to be the Ackermann's formula (4.56):

$$\mathbf{L} = \chi_{A-LC}(\mathbf{A})\underline{q}_o \tag{4.73}$$

4.6 State observer for MIMO systems - Roppenecker's formula

We have seen in Equation (4.7) that the dynamics of the estimation error $\underline{e}(t)$ reads:

$$\dot{\underline{e}}(t) = (\mathbf{A} - \mathbf{LC})\,\underline{e}(t) \tag{4.74}$$

The purpose of this section is to design the observer gain matrix $\mathbf{L}$ such that the eigenvalues of matrix $\mathbf{A} - \mathbf{LC}$ are assigned to predefined eigenvalues $\lambda_{L1}, \cdots, \lambda_{Ln}$ where n is the size of matrix $\mathbf{A}$.

Let $\lambda_{K1}, \cdots, \lambda_{Kn}$ be n distinct specified eigenvalues of the closed loop state matrix $\mathbf{A} - \mathbf{LC}$. Furthermore we assume that eigenvalues of matrix $\mathbf{A}$ do not shift (meaning that they are different) the eigenvalues of the closed loop state matrix $\mathbf{A} - \mathbf{LC}$. Then it can be shown that the transpose of the observer gain matrix $\mathbf{L}$ can be computed as follows where $\underline{p}_i$ denotes parameter vectors:

$$\mathbf{L}^T = -\left[\begin{array}{ccc} \underline{p}_1 & \cdots & \underline{p}_n \end{array}\right]\left[\begin{array}{ccc} \underline{w}_{L1} & \cdots & \underline{w}_{Ln} \end{array}\right]^{-1} \tag{4.75}$$

Where vector $\underline{w}_{Li}$ is related to parameter vector $\underline{p}_i$ through the following relation:

$$\left[\begin{array}{cc} \mathbf{A}^T - \lambda_{Li}\mathbb{I} & \mathbf{C}^T \end{array}\right]\left[\begin{array}{c} \underline{w}_{Li} \\ \underline{p}_i \end{array}\right] = \underline{0} \tag{4.76}$$

This is the so-called Roppenecker's formula to get the observer gain matrix $\mathbf{L}$.

To get this result we rewrite the dynamics of the estimation error $\underline{e}(t)$ as follows:

$$\dot{\underline{e}}^T(t) = (\mathbf{A} - \mathbf{LC})^T \underline{e}^T(t) \Leftrightarrow \dot{\underline{e}}^T(t) = \left(\mathbf{A}^T - \mathbf{C}^T\mathbf{L}^T\right)\underline{e}^T(t) \tag{4.77}$$

Let $\underline{w}_{Li}$ be an eigenvector of matrix $(\mathbf{A} - \mathbf{LC})^T$. As far as $(\mathbf{A} - \mathbf{LC})^T$ and $(\mathbf{A} - \mathbf{LC})$ have the same eigenvalues, which are the predefined eigenvalues λ_{Li}, we can write:

$$\left(\mathbf{A}^T - \mathbf{C}^T\mathbf{L}^T\right)\underline{w}_{Li} = \lambda_{Li}\underline{w}_{Li} \tag{4.78}$$

The preceding equation can be written as follows:

$$\left(\mathbf{A}^T - \lambda_{Li}\mathbb{I}\right)\underline{w}_{Li} = \mathbf{C}^T\mathbf{L}^T\underline{w}_{Li} \tag{4.79}$$

That is:

$$\underline{w}_{Li} = \left(\mathbf{A}^T - \lambda_{Li}\mathbb{I}\right)^{-1}\mathbf{C}^T\mathbf{L}^T\underline{w}_{Li} \tag{4.80}$$

Then we introduce n parameter vectors $\underline{p}_i$ defined as follows;

$$\underline{p}_i = -\mathbf{L}^T\underline{w}_{Li} \tag{4.81}$$

Each parameter vector $\underline{p}_i$ is a $p \times 1$ vector where p is the number of rows of matrix $\mathbf{C}$, that is the number of outputs of the system.

Using parameter vector $\underline{p}_i$ Equation (4.80) reads:

$$\underline{w}_{Li} = -\left(\mathbf{A}^T - \lambda_{Li}\mathbb{I}\right)^{-1}\mathbf{C}^T\underline{p}_i \tag{4.82}$$

Writing Equation (4.81) for the n distinct predefined eigenvalues $\lambda_{L1}, \cdots, \lambda_{Ln}$ leads to the following relation:

$$\left[\begin{array}{ccc} \underline{p}_1 & \cdots & \underline{p}_n \end{array}\right] = -\mathbf{L}^T\left[\begin{array}{ccc} \underline{w}_{L1} & \cdots & \underline{w}_{Ln} \end{array}\right] \tag{4.83}$$

Finally the transpose of the observer gain matrix $\mathbf{L}$ can be computed as follows:

$$\mathbf{L}^T = -\left[\begin{array}{ccc} \underline{p}_1 & \cdots & \underline{p}_n \end{array}\right]\left[\begin{array}{ccc} \underline{w}_{L1} & \cdots & \underline{w}_{Ln} \end{array}\right]^{-1} \tag{4.84}$$

We have retrieved the so-called Roppenecker's formula to get the observer gain matrix $\mathbf{L}$.

It is worth noticing the following facts:

- Using equation (4.81) within Equation (4.79) it is clear that parameter vectors $\underline{p}_i$ shall satisfy the following relation:

$$\left(\mathbf{A}^T - \lambda_{Li}\mathbb{I}\right)\underline{w}_{Li} = -\mathbf{C}^T\underline{p}_i \tag{4.85}$$

  That is:

$$\begin{bmatrix} \mathbf{A}^T - \lambda_{Li}\mathbb{I} & \mathbf{C}^T \end{bmatrix}\begin{bmatrix} \underline{w}_{Li} \\ \underline{p}_i \end{bmatrix} = \underline{0} \tag{4.86}$$

  It is clear from the previous equation that each $(n+p)\times 1$ vector $\begin{bmatrix} \underline{w}_{Li} \\ \underline{p}_i \end{bmatrix}$ belong to the null-space of matrix $\begin{bmatrix} \mathbf{A}^T - \lambda_{Li}\mathbb{I} & \mathbf{C}^T \end{bmatrix}$. So once any $(n+p)\times 1$ vector which belongs to the null-space of matrix $\begin{bmatrix} \mathbf{A}^T - \lambda_{Li}\mathbb{I} & \mathbf{C}^T \end{bmatrix}$ has been found, its p bottom rows are used to form vector parameter $\underline{p}_i$. In the MIMO case several possibilities are offered.

- By taking the transpose of equation (4.86) we get the following expression:

$$\begin{bmatrix} \underline{w}_{Li}^T & \underline{p}_i^T \end{bmatrix}\begin{bmatrix} \mathbf{A} - \lambda_{Li}\mathbb{I} \\ \mathbf{C} \end{bmatrix} = \underline{0} \tag{4.87}$$

  We recognize in matrix $\begin{bmatrix} \mathbf{A} - \lambda_{Li}\mathbb{I} \\ \mathbf{C} \end{bmatrix}$ the key matrix used in the PBH observability test;

- If we wish to keep an eigenvalue λ_i of matrix $\mathbf{A}$ within the set of eigenvalues of $\mathbf{A} - \mathbf{LC}$ then Equation (4.79) is equal to zero because in that case $\underline{w}_{Li}$ is also an eigenvector of $\mathbf{A}$:

$$\lambda_{Li} = \lambda_i \Rightarrow \left(\mathbf{A}^T - \lambda_{Li}\mathbb{I}\right)\underline{w}_{Li} = -\mathbf{C}^T\underline{p}_i = \left(\mathbf{A}^T - \lambda_i\mathbb{I}\right)\underline{w}_{Li} = \underline{0} \tag{4.88}$$

  Consequently we have to replace $\underline{p}_i$ by $\underline{0}$ and $\underline{w}_{Li}$ by eigenvector $\underline{w}_i$ of $\mathbf{A}$ corresponding to λ_i in the Roppenecker's formula (4.75);

- If we chose a complex eigenvalue λ_{Li} then its complex conjugate must also be chosen. Let's λ_{LiR} and λ_{LiI} be the real part and the imaginary part of λ_{Li}, $\underline{w}_{LiR}$ and $\underline{w}_{LiI}$ be the real part and the imaginary part of $\underline{w}_{Li}$ and $\underline{p}_{iR}$ and $\underline{p}_{iI}$ be the real part and the imaginary part of $\underline{p}_i$ respectively:

$$\begin{cases} \lambda_{Li} = \lambda_{LiR} + j\lambda_{LiI} \\ \underline{w}_{Li} = \underline{w}_{LiR} + j\underline{w}_{LiI} \\ \underline{p}_i = \underline{p}_{iR} + j\underline{p}_{iI} \end{cases} \tag{4.89}$$

  Then equation (4.86) reads:

$$\left(\mathbf{A}^T - (\lambda_{LiR} + j\lambda_{LiI})\,\mathbb{I}\right)(\underline{w}_{LiR} + j\underline{w}_{LiI}) + \mathbf{C}^T\left(\underline{p}_{iR} + j\underline{p}_{iI}\right) = \underline{0} \tag{4.90}$$

  Taking the complex conjugate of the preceding equation reads:

$$\left(\mathbf{A}^T - (\lambda_{LiR} - j\lambda_{LiI})\,\mathbb{I}\right)(\underline{w}_{LiR} - j\underline{w}_{LiI}) + \mathbf{C}^T\left(\underline{p}_{iR} - j\underline{p}_{iI}\right) = \underline{0} \tag{4.91}$$

Summing and subtracting Equations (4.90) and (4.91) reads:

$$\begin{cases} \left(\mathbf{A}^T - \lambda_{LiR}\mathbb{I}\right)\underline{w}_{LiR} + \lambda_{LiI}\underline{w}_{LiI} + \mathbf{C}^T\underline{p}_{iR} = \underline{0} \\ \left(\mathbf{A}^T - \lambda_{LiR}\mathbb{I}\right)\underline{w}_{LiI} - \lambda_{LiI}\underline{w}_{LiR} + \mathbf{C}^T\underline{p}_{iI} = \underline{0} \end{cases} \tag{4.92}$$

That is in vector form:

$$\begin{bmatrix} \mathbf{A}^T - \lambda_{LiR}\mathbb{I} & \lambda_{LiI}\mathbb{I} & \mathbf{C}^T & \mathbf{0} \\ -\lambda_{LiI}\mathbb{I} & \mathbf{A}^T - \lambda_{LiR}\mathbb{I} & \mathbf{0} & \mathbf{C}^T \end{bmatrix} \begin{bmatrix} \underline{w}_{LiR} \\ \underline{w}_{LiI} \\ \underline{p}_{iR} \\ \underline{p}_{iI} \end{bmatrix} = \underline{0} \tag{4.93}$$

In Equation (4.75) vectors $\underline{p}_i$, $\underline{p}_i^*$ (where * denotes complex conjugate), $\underline{w}_{Li}$ and $\underline{w}_{Li}^*$ are replaced by vectors $\underline{p}_{iR}$, $\underline{p}_{iI}$, $\underline{w}_{LiR}$ and $\underline{w}_{LiI}$, respectively.

– In the SISO case the observer gain matrix $\mathbf{L}$ no more depends on parameter vectors $\underline{p}_i$. Indeed is that case the observer gain matrix $\mathbf{L}$ is obtained as follows:

$$\mathbf{L}^T = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \left(\mathbf{A}^T - \lambda_{L1}\mathbb{I}\right)^{-1}\mathbf{C}^T & \cdots & \left(\mathbf{A}^T - \lambda_{Ln}\mathbb{I}\right)^{-1}\mathbf{C}^T \end{bmatrix}^{-1} \tag{4.94}$$

To get this result we start by observing that in the SISO case parameter vector are scalars; they will be denoted p_i. Let vector $\underline{l}_i$ be defined as follows:

$$\underline{l}_i = -\left(\mathbf{A}^T - \lambda_{Li}\mathbb{I}\right)^{-1}\mathbf{C}^T \tag{4.95}$$

Then Equation (4.75) reads:

$$\mathbf{L}^T = -\begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix} \begin{bmatrix} \underline{l}_1 p_1 & \cdots & \underline{l}_n p_n \end{bmatrix}^{-1} \tag{4.96}$$

Let's rearrange the term $\begin{bmatrix} \underline{l}_1 p_1 & \cdots & \underline{l}_n p_n \end{bmatrix}^{-1}$ as follows:

$$\begin{aligned} \begin{bmatrix} \underline{l}_1 p_1 & \cdots & \underline{l}_n p_n \end{bmatrix}^{-1} &= \left(\begin{bmatrix} \underline{l}_1 & \cdots & \underline{l}_n \end{bmatrix} \begin{bmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_n \end{bmatrix} \right)^{-1} \\ &= \begin{bmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_n \end{bmatrix}^{-1} \begin{bmatrix} \underline{l}_1 & \cdots & \underline{l}_n \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \frac{\prod_{i\neq 1} p_i}{\prod_{i=1}^n p_i} & & 0 \\ & \ddots & \\ 0 & & \frac{\prod_{i\neq n} p_i}{\prod_{i=1}^n p_i} \end{bmatrix} \begin{bmatrix} \underline{l}_1 & \cdots & \underline{l}_n \end{bmatrix}^{-1} \end{aligned} \tag{4.97}$$

Multiplying this expression by $-\left[\begin{array}{ccc} p_1 & \cdots & p_n \end{array}\right]$ leads to the expression of $\mathbf{L}^T$:

$$\begin{aligned}
\mathbf{L}^T &= -\left[\begin{array}{ccc} p_1 & \cdots & p_n \end{array}\right]\left[\begin{array}{ccc} \underline{l}_1 p_1 & \cdots & \underline{l}_n p_n \end{array}\right]^{-1} \\
&= -\left[\begin{array}{ccc} p_1 & \cdots & p_n \end{array}\right]\left[\begin{array}{ccc} \frac{\prod_{i\neq 1} p_i}{\prod_{i=1}^n p_i} & & 0 \\ & \ddots & \\ 0 & & \frac{\prod_{i\neq n} p_i}{\prod_{i=1}^n p_i} \end{array}\right]\left[\begin{array}{ccc} \underline{l}_1 & \cdots & \underline{l}_n \end{array}\right]^{-1} \\
&= -\left[\begin{array}{ccc} 1 & \cdots & 1 \end{array}\right]\left[\begin{array}{ccc} \underline{l}_1 & \cdots & \underline{l}_n \end{array}\right]^{-1}
\end{aligned} \tag{4.98}$$

Using the expression of vector $\underline{l}_i = -\left(\mathbf{A}^T - \lambda_{Li}\mathbb{I}\right)^{-1}\mathbf{C}^T$ provided by Equation (4.95) we finally get:

$$\begin{aligned}
\mathbf{L}^T = &\left[\begin{array}{ccc} 1 & \cdots & 1 \end{array}\right] \\
&\left[\begin{array}{ccc} \left(\mathbf{A}^T - \lambda_{L1}\mathbb{I}\right)^{-1}\mathbf{C}^T & \cdots & \left(\mathbf{A}^T - \lambda_{Ln}\mathbb{I}\right)^{-1}\mathbf{C}^T \end{array}\right]^{-1}
\end{aligned} \tag{4.99}$$

We conclude that in the SISO case the observer gain matrix $\mathbf{L}$ no more depends on parameter vectors $\underline{p}_i$.

4.7 Reduced-order observer

Consider the following state space representation where $\mathbf{D} = \mathbf{0}$ and where $\underline{x}(t)$ denotes the state vector, $\underline{y}(t)$ the measured output and $\underline{u}(t)$ the control input:

$$\left\{\begin{array}{l} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{array}\right. \tag{4.100}$$

We will show in this section how to derive an observer of reduced dimension by exploiting the output equation $\underline{y}(t) = \mathbf{C}\underline{x}(t)$.

Let p be the rank of matrix $\mathbf{C}$. Then, from the output equation $\underline{y}(t) = \mathbf{C}\underline{x}(t)$, we can extract p linearly independent equations and then compute directly p components of the state vector $\underline{x}(t)$. Assuming that n is the dimension of the state vector, only the remaining $n - r$ components of the state vector have to be estimated and then the order of the observer can be reduced to $n - r$.

More precisely, since matrix $\mathbf{C}$ is of rank p, there exists a $n \times n$ nonsingular matrix $\mathbf{P}$ such that the following relation holds, where $\mathbb{I}_p$ denotes the identity matrix of size p and $\mathbf{0}_{p,n-p}$ the $p \times (n-p)$ matrix of zeros:

$$\mathbf{CP} = \left[\begin{array}{cc} \mathbb{I}_p & \mathbf{0}_{p,n-p} \end{array}\right] \tag{4.101}$$

Indeed, let $\overline{\mathbf{C}}$ be a $(n-p)\times n$ matrix such that matrix $\left[\begin{array}{c} \mathbf{C} \\ \hline \overline{\mathbf{C}} \end{array}\right]$ is nonsingular. Then a possible choice for $\mathbf{P}$ is the following:

$$\mathbf{P} = \left[\begin{array}{c} \mathbf{C} \\ \hline \overline{\mathbf{C}} \end{array}\right]^{-1} \tag{4.102}$$

Hence, using (4.101), we can write:

$$\underline{y}(t) = \mathbf{C}\underline{x}(t) = \mathbf{C}\mathbf{P}\mathbf{P}^{-1}\underline{x}(t) = \begin{bmatrix} \mathbb{I}_p & \mathbf{0}_{p,n-p} \end{bmatrix} \underline{x}^*(t) \tag{4.103}$$

where:

$$\underline{x}^*(t) = \mathbf{P}^{-1}\underline{x}(t) \Leftrightarrow \underline{x}(t) = \mathbf{P}\underline{x}^*(t) \tag{4.104}$$

Hence, mapping the system in the new state vector $\underline{x}^*(t)$ via the similarity transformation $\mathbf{P}$, we obtain the following state space representation:

$$\begin{cases} \dot{\underline{x}}^*(t) = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\underline{x}^*(t) + \mathbf{P}^{-1}\mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\mathbf{P}\underline{x}^*(t) \end{cases} \tag{4.105}$$

From the fact that $\mathbf{C}\mathbf{P} = \begin{bmatrix} \mathbb{I}_p & \mathbf{0}_{p,n-p} \end{bmatrix}$, it can be seen that the first p components of the new state vector $\underline{x}^*(t)$ are equal to $\underline{y}(t)$. Thus we can write:

$$\mathbf{C}\mathbf{P} = \begin{bmatrix} \mathbb{I}_p & \mathbf{0}_{p,n-p} \end{bmatrix} \Rightarrow \underline{x}^*(t) = \begin{bmatrix} \underline{y}(t) \\ \underline{x}_r(t) \end{bmatrix} \tag{4.106}$$

As far as the p first components of the new state vector $\underline{x}^*(t)$ are equal to $\underline{y}(t)$, they are available through measurements and thus there is no need to estimate those components. Consequently the reduced-order observer focuses on the estimation of the remaining state vector $\underline{x}_r(t)$.

The state equation (4.105) can be written as follows:

$$\begin{cases} \dot{\underline{x}}^*(t) := \begin{bmatrix} \dot{\underline{y}}(t) \\ \dot{\underline{x}}_r(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}^*_{11} & \mathbf{A}^*_{12} \\ \mathbf{A}^*_{21} & \mathbf{A}^*_{22} \end{bmatrix} \begin{bmatrix} \underline{y}(t) \\ \underline{x}_r(t) \end{bmatrix} + \begin{bmatrix} \mathbf{B}^*_1 \\ \mathbf{B}^*_2 \end{bmatrix} \underline{u}(t) \\ \underline{y}(t) = \begin{bmatrix} \mathbb{I}_p & \mathbf{0}_{p,n-p} \end{bmatrix} \begin{bmatrix} \underline{y}(t) \\ \underline{x}_r(t) \end{bmatrix} := \mathbf{C}^*\underline{x}^*(t) \end{cases} \tag{4.107}$$

Let's split matrix $\mathbf{P}$ as follows, where $\mathbf{P}_1$ is a $n \times p$ matrix and $\mathbf{P}_2$ a $n \times (n-p)$ matrix:

$$\mathbf{P} = \begin{bmatrix} \mathbf{C} \\ \overline{\mathbf{C}} \end{bmatrix}^{-1} := \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_2 \end{bmatrix} \tag{4.108}$$

Then we get:

$$\begin{cases} \begin{bmatrix} \mathbf{A}^*_{11} & \mathbf{A}^*_{12} \\ \mathbf{A}^*_{21} & \mathbf{A}^*_{22} \end{bmatrix} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \mathbf{C}\mathbf{A}\mathbf{P}_1 & \mathbf{C}\mathbf{A}\mathbf{P}_2 \\ \overline{\mathbf{C}}\mathbf{A}\mathbf{P}_1 & \overline{\mathbf{C}}\mathbf{A}\mathbf{P}_2 \end{bmatrix} \\ \begin{bmatrix} \mathbf{B}^*_1 \\ \mathbf{B}^*_2 \end{bmatrix} = \mathbf{P}^{-1}\mathbf{B} = \begin{bmatrix} \mathbf{C}\mathbf{B} \\ \overline{\mathbf{C}}\mathbf{B} \end{bmatrix} \end{cases} \tag{4.109}$$

To design an observer for $\underline{x}_r(t)$, we use the knowledge that *an observer has the same structure as the system plus the driving feedback term whose role is to reduce the estimation error to zero*[4]. Hence, an observer for $\underline{x}_r(t)$ reads:

$$\dot{\hat{\underline{x}}}_r(t) = \mathbf{A}^*_{21}\underline{y}(t) + \mathbf{A}^*_{22}\hat{\underline{x}}_r(t) + \mathbf{B}^*_2\underline{u}(t) + \mathbf{L}_r\left(\underline{y}(t) - \hat{\underline{y}}(t)\right) \tag{4.110}$$

[4]Zoran Gajic, Introduction to Linear and Nonlinear Observers, Rutgers University, https://www.ece.rutgers.edu/~gajic/

Unfortunately, since $\underline{y}(t) = \hat{\underline{y}}(t)$, the difference $\underline{y}(t) - \hat{\underline{y}}(t)$ does not carry any information about $\underline{x}_r(t)$. Nevertheless, by taking the time derivative of $\underline{y}(t)$, we get the first equation of (4.107) which carries information about $\underline{x}_r(t)$:

$$\begin{array}{l} \dot{\underline{y}}(t) = \mathbf{A}_{11}^* \underline{y}(t) + \mathbf{A}_{12}^* \underline{x}_r(t) + \mathbf{B}_1^* \underline{u}(t) \\ \Rightarrow \mathbf{A}_{12}^* \underline{x}_r(t) = \dot{\underline{y}}(t) - \mathbf{A}_{11}^* \underline{y}(t) - \mathbf{B}_1^* \underline{u}(t) \end{array} \tag{4.111}$$

Regarding $\underline{y}_r(t) := \mathbf{A}_{12}^* \underline{x}_r(t)$ as a *virtual* output of the reduced state equation, the observer for $\underline{x}_r(t)$ finally reads:

$$\left\{ \begin{array}{ll} \dot{\hat{\underline{x}}}_r(t) & = \mathbf{A}_{21}^* \underline{y}(t) + \mathbf{A}_{22}^* \hat{\underline{x}}_r(t) + \mathbf{B}_2^* \underline{u}(t) + \mathbf{L}_r \left(\underline{y}_r(t) - \mathbf{A}_{12}^* \hat{\underline{x}}_r(t) \right) \\ \underline{y}_r(t) & = \dot{\underline{y}}(t) - \mathbf{A}_{11}^* \underline{y}(t) - \mathbf{B}_1^* \underline{u}(t) \end{array} \right. \tag{4.112}$$

Furthermore the dynamics of the error $\underline{e}_r(t) = \underline{x}_r(t) - \hat{\underline{x}}_r(t)$ reads as follows:

$$\begin{array}{rl} \underline{e}_r(t) & = \underline{x}_r(t) - \hat{\underline{x}}_r(t) \\ \Rightarrow \dot{\underline{e}}_r(t) & = \dot{\underline{x}}_r(t) - \dot{\hat{\underline{x}}}_r(t) \\ & = \mathbf{A}_{21}^* \underline{y}(t) + \mathbf{A}_{22}^* \underline{x}_r(t) + \mathbf{B}_2^* \underline{u}(t) \\ & \quad - \left(\mathbf{A}_{21}^* \underline{y}(t) + \mathbf{A}_{22}^* \hat{\underline{x}}_r(t) + \mathbf{B}_2^* \underline{u}(t) + \mathbf{L}_r \left(\underline{y}_r(t) - \mathbf{A}_{12}^* \hat{\underline{x}}_r(t) \right) \right) \\ & = \left(\mathbf{A}_{22}^* - \mathbf{L}_r \mathbf{A}_{12}^* \right) \underline{e}_r(t) \end{array} \tag{4.113}$$

Consequently, designing the observer gain $\mathbf{L}_r$ such that the characteristic polynomial of matrix $\mathbf{A}_{22}^* - \mathbf{L}_r \mathbf{A}_{12}^*$ is Hurwitz leads to the asymptotic convergence of the estimates $\hat{\underline{x}}_r(t)$ towards $\underline{x}_r(t)$. Such a design is always possible as soon as the pair $(\mathbf{A}_{22}^*, \mathbf{A}_{12}^*)$ is observable, which is a consequence of the observability of the pair $(\mathbf{A}, \mathbf{C})$ (this can be shown using PBH test[4]).

Since it is not wise to use $\dot{\underline{y}}(t)$ because in practice the differentiation process introduces noise, we will estimate vector $\hat{\underline{x}}_{ry}(t)$ rather than $\underline{x}_r(t)$. Vector $\hat{\underline{x}}_{ry}(t)$ is defined as follows:

$$\hat{\underline{x}}_{ry}(t) := \hat{\underline{x}}_r(t) - \mathbf{L}_r \underline{y}(t) \tag{4.114}$$

From (4.112), we get the following observer for $\hat{\underline{x}}_{ry}(t)$:

$$\begin{array}{rl} \dot{\hat{\underline{x}}}_{ry}(t) & = \dot{\hat{\underline{x}}}_r(t) - \mathbf{L}_r \dot{\underline{y}}(t) \\ & = \mathbf{A}_{21}^* \underline{y}(t) + \mathbf{A}_{22}^* \hat{\underline{x}}_r(t) + \mathbf{B}_2^* \underline{u}(t) + \mathbf{L}_r \left(\underline{y}_r(t) - \mathbf{A}_{12}^* \hat{\underline{x}}_r(t) \right) - \mathbf{L}_r \dot{\underline{y}}(t) \\ & = \mathbf{A}_{21}^* \underline{y}(t) + \mathbf{A}_{22}^* \left(\hat{\underline{x}}_{ry}(t) + \mathbf{L}_r \underline{y}(t) \right) + \mathbf{B}_2^* \underline{u}(t) \\ & \quad + \mathbf{L}_r \left(-\mathbf{A}_{11}^* \underline{y}(t) - \mathbf{B}_1^* \underline{u}(t) - \mathbf{A}_{12}^* \left(\hat{\underline{x}}_{ry}(t) + \mathbf{L}_r \underline{y}(t) \right) \right) \\ & := \mathbf{A}_y \underline{y}(t) + \mathbf{A}_{ry} \hat{\underline{x}}_{ry}(t) + \mathbf{B}_u \underline{u}(t) \end{array} \tag{4.115}$$

where:

$$\left\{ \begin{array}{l} \mathbf{A}_y = \mathbf{A}_{21}^* + \mathbf{A}_{22}^* \mathbf{L}_r - \mathbf{L}_r \mathbf{A}_{11}^* - \mathbf{L}_r \mathbf{A}_{12}^* \mathbf{L}_r \\ \mathbf{A}_{ry} = \mathbf{A}_{22}^* - \mathbf{L}_r \mathbf{A}_{12}^* \\ \mathbf{B}_u = \mathbf{B}_2^* - \mathbf{L}_r \mathbf{B}_1^* \end{array} \right. \tag{4.116}$$

Assembling the previous results, the estimation of state vector $\underline{x}(t)$ finally reads as follows where the dynamics of $\hat{\underline{x}}_{ry}(t)$ is given by (4.115):

$$\hat{\underline{x}}(t) = \mathbf{P} \hat{\underline{x}}^*(t) = \mathbf{P} \left[\begin{array}{c} \underline{y}(t) \\ \hat{\underline{x}}_r(t) \end{array} \right] = \mathbf{P} \left[\begin{array}{c} \underline{y}(t) \\ \hat{\underline{x}}_{ry}(t) + \mathbf{L}_r \underline{y}(t) \end{array} \right] \tag{4.117}$$

Chapter 5

Controller design

5.1 Introduction

Controller enables to obtain stable systems which meet performances specifications. In the case where the full state vector $\underline{x}(t)$ is available then controller design involves state feedback. In the more usual case where only some components of the state vector are available through the output vector $\underline{y}(t)$ then controller design involves output feedback in association with a state observer.

This chapter focuses on controllers design. More specifically state feedback controller for SISO systems in controllable canonical form, state feedback controller for SISO systems in arbitrary state-space representation, static state feedback controller and static output feedback controller for MIMO systems will be presented. We will also present controller with integral action.

5.2 Static state feedback controller

Consider the following state equation where $\underline{x}(t)$ denotes the state vector and $\underline{u}(t)$ the control input:

$$\dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \tag{5.1}$$

We will assume in the following that the full state vector $\underline{x}(t)$ is available for control.

Let $\underline{r}(t)$ be a reference input signal. A static state feedback (or full state) controller computes the control input $\underline{u}(t)$ as a function of a state-feedback gain $\mathbf{K}$ and a feedforward gain matrix $\mathbf{H}$ as follows:

$$\underline{u}(t) = \mathbf{H}\underline{r}(t) - \mathbf{K}\underline{x}(t) \tag{5.2}$$

Substituting the control law (5.2) into the state equation (5.1) of the system yields:

$$\dot{\underline{x}}(t) = (\mathbf{A} - \mathbf{BK})\,\underline{x}(t) + \mathbf{BH}\underline{r}(t) \tag{5.3}$$

- The purpose of the controller gain $\mathbf{K}$ is at least to assign the eigenvalues of the closed-loop state matrix $\mathbf{A} - \mathbf{BK}$ at predefined locations. For MIMO

systems there are additional degrees of freedom which may be used for others purposes like eigenstructure assignment;

– The purpose of the feedforward gain matrix $\mathbf{H}$ is to pick up the desired final value such that the closed-loop system has no steady state error to any constant value of the reference input $\underline{r}(t)$.

  In steady state conditions we have $\dot{\underline{x}}(t) = \underline{0}$ and consequently (5.3) reads:

$$\dot{\underline{x}}(t) = \underline{0} \Rightarrow \underline{x}(t) = -\left(\mathbf{A} - \mathbf{BK}\right)^{-1} \mathbf{BH}\underline{r}(t) \tag{5.4}$$

  On the other hand, using (5.2) the output equation $\underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t)$ reads:

$$\begin{cases} \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \\ \underline{u}(t) = -\mathbf{K}\underline{x}(t) + \mathbf{H}\underline{r}(t) \end{cases} \Rightarrow \underline{y}(t) = \left(\mathbf{C} - \mathbf{DK}\right)\underline{x}(t) + \mathbf{DH}\underline{r}(t) \tag{5.5}$$

  Inserting (5.4) into (5.5) yields:

$$\begin{aligned} \underline{y}(t) &= -\left(\mathbf{C} - \mathbf{DK}\right)\left(\mathbf{A} - \mathbf{BK}\right)^{-1} \mathbf{BH}\underline{r}(t) + \mathbf{DH}\underline{r}(t) \\ &= \Big(\mathbf{D} - \left(\mathbf{C} - \mathbf{DK}\right)\left(\mathbf{A} - \mathbf{BK}\right)^{-1} \mathbf{B}\Big) \mathbf{H}\underline{r}(t) \end{aligned} \tag{5.6}$$

  Then matrix $\mathbf{H}$ is computed such that the closed-loop system has no steady state error to any constant value of the reference input $\underline{r}(t)$. So imposing $\underline{y}(t) = \underline{r}(t)$ leads to the following expression of the feedforward gain matrix $\mathbf{H}$:

$$\underline{y}(t) = \underline{r}(t) \Rightarrow \mathbf{H} = \Big(\mathbf{D} - \left(\mathbf{C} - \mathbf{DK}\right)\left(\mathbf{A} - \mathbf{BK}\right)^{-1} \mathbf{B}\Big)^{-1} \tag{5.7}$$

  In the usual case where matrix $\mathbf{D}$ is null the preceding relation simplifies as follows:

$$\mathbf{H} = -\Big(\mathbf{C}\left(\mathbf{A} - \mathbf{BK}\right)^{-1} \mathbf{B}\Big)^{-1} \tag{5.8}$$

We will see in section 5.5 that adding an integral action within the controller is an alternative method which avoid the computation of feedforward gain matrix $\mathbf{H}$.

In the following we will assume that the system is controllable, or at least stabilizable, such that it is possible to design a state feedback controller. Indeed Wonham[1] has shown that controllability of an open-loop system is equivalent to the possibility of assigning an arbitrary set of poles to the transfer matrix of the closed-loop system, formed by means of suitable linear feedback of the state.

[1] Wonham W., On pole assignment in multi-input controllable linear systems, IEEE Transactions on Automatic Control, vol. 12, no. 6, pp. 660-665, December 1967. doi: 10.1109/TAC.1967.1098739

5.3 Control of SISO systems

5.3.1 State feedback controller in controllable canonical form

Let $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$ be an controllable Single-Input Single-Output (SISO) linear time-invariant system of order n and let $\chi_{A-BK}(s)$ be an imposed n^{th} order monic polynomial (a monic polynomial is a polynomial in which the leading coefficient, that is the nonzero coefficient of highest degree, is equal to 1). Polynomial $\chi_{A-BK}(s)$ corresponds to the characteristic polynomial of matrix $\mathbf{A} - \mathbf{BK}$. It is formed thanks to the predefined eigenvalues $\lambda_{K_1}, \cdots, \lambda_{K_n}$ assigned for the dynamics of the controller:

$$\chi_{A-BK}(s) = \det\left(s\mathbb{I} - (\mathbf{A} - \mathbf{BK})\right) = (s - \lambda_{K_1}) \cdots (s - \lambda_{K_n}) \tag{5.9}$$

When expanding the preceding product we get:

$$\chi_{A-BK}(s) = s^n + p_{n-1}s^{n-1} + \cdots + p_1 s + p_0 \tag{5.10}$$

We begin with the case where the system realization $\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$ is the controllable canonical form. Then matrices $\mathbf{A}_c$ and $\mathbf{B}_c$ are the following:

$$\begin{cases} \mathbf{A}_c = \begin{bmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ & & \ddots & \ddots & 0 \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} \\ \mathbf{B}_c = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \end{cases} \tag{5.11}$$

Matrices $\mathbf{C}_c$ and $\mathbf{D}$ will not be used.

Let $\mathbf{K}_c$ be the controller gain matrix when the controllable canonical form of the system is used. For a SISO system this is a $1 \times n$ matrix. Let K_i be the (scalar) component in the ith row of matrix $\mathbf{K}_c$:

$$\mathbf{K}_c = \begin{bmatrix} K_1 & \cdots & K_n \end{bmatrix} \tag{5.12}$$

Using the duality principle we can infer that the expression of the state feedback controller for SISO systems in controllable canonical form has the following expression:

$$\mathbf{K}_c = \mathbf{L}_o^T = \begin{bmatrix} p_0 - a_0 & \cdots & p_{n-1} - a_{n-1} \end{bmatrix} \tag{5.13}$$

To check it just notice that when the controllable canonical form of the system is used then matrix $\mathbf{A}_c - \mathbf{B}_c\mathbf{K}_c$ reads:

$$\begin{aligned}
\mathbf{A}_c - \mathbf{B}_c\mathbf{K}_c &= \begin{bmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ & & \ddots & \ddots & 0 \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} K_1 & \cdots & K_n \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ & & \ddots & \ddots & 0 \\ 0 & 0 & 0 & & 1 \\ -a_0 - K_1 & -a_1 - K_2 & -a_2 - K_3 & \cdots & -a_{n-1} - K_n \end{bmatrix}
\end{aligned} \tag{5.14}$$

Since this matrix still remains in the controllable canonical form its characteristic polynomial is readily written as follows:

$$\begin{aligned}
\chi_{A-BK}(s) &= \det(\mathbf{A} - \mathbf{BK}) \\
&= \det(\mathbf{A}_c - \mathbf{B}_c\mathbf{K}_c) \\
&= s^n + (a_{n-1} + K_n)\, s^{n-1} + \cdots + (a_1 + K_2)\, s + a_0 + K_1
\end{aligned} \tag{5.15}$$

Identifying Equations (5.10) and (5.15) leads to the expression of each component of the controller matrix $\mathbf{K}_c$:

$$\begin{cases} p_0 = a_0 + K_1 \\ p_1 = a_1 + K_2 \\ \vdots \\ p_{n-1} = a_{n-1} + K_n \end{cases} \Leftrightarrow \begin{aligned} \mathbf{K}_c &= \begin{bmatrix} K_1 & K_2 & \cdots & K_n \end{bmatrix} \\ &= \begin{bmatrix} p_0 - a_0 & p_1 - a_1 & \cdots & p_{n-1} - a_{n-1} \end{bmatrix} \end{aligned} \tag{5.16}$$

5.3.2 State feedback controller in arbitrary state-space representation

When an arbitrary state-space representation is used the system has to be converted into the controllable canonical form via a similarity transformation. Let $\mathbf{P}_c$ be the matrix of the similarity transformation which enables to get the controllable canonical form. We get:

$$\underline{x}(t) = \mathbf{P}_c\underline{x}_c(t) \Leftrightarrow \underline{x}_c(t) = \mathbf{P}_c^{-1}\underline{x}(t) \tag{5.17}$$

We have seen in the chapter dedicated to *Realization of transfer functions* that $\mathbf{P}_c^{-1}$ is a constant nonsingular change of basis matrix which is obtained through the following relation:

$$\mathbf{P}_c^{-1} = \mathbf{Q}_{cc}\mathbf{Q}_c^{-1} \tag{5.18}$$

Where:

– $\mathbf{Q}_c$ is the controllability matrix in the actual basis:

$$\mathbf{Q}_c = \left[\begin{array}{cccc} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{array}\right] \tag{5.19}$$

– and $\mathbf{Q}_{cc}$ the controllability matrix expressed in the controllable canonical basis (which is readily obtained through $\det(s\mathbb{I} - \mathbf{A})$).

Thus the control law $\underline{u}(t)$ reads:

$$\underline{u}(t) = -\mathbf{K}_c\underline{x}_c(t) + \mathbf{H}\underline{r}(t) = -\mathbf{K}_c\mathbf{P}_c^{-1}\underline{x}(t) + \mathbf{H}\underline{r}(t) \tag{5.20}$$

That is:

$$\underline{u}(t) = -\mathbf{K}\underline{x}(t) + \mathbf{H}\underline{r}(t) \tag{5.21}$$

Where:

$$\mathbf{K} = \mathbf{K}_c\mathbf{P}_c^{-1} \tag{5.22}$$

Example 5.1. *Design a* state *feedback controller for the following unstable plant:*

$$\begin{cases} \dot{\underline{x}}(t) = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u(t) \\ \\ y(t) = \begin{bmatrix} 3 & 5 \end{bmatrix} \underline{x}(t) \end{cases} \tag{5.23}$$

As far as this is a modal (or diagonal) state space representation, plant's controllability is readily checked by inspecting column of control matrix $\mathbf{B}$*: because there is no null element in the control matrix* $\mathbf{B}$ *we conclude that the plant is controllable by applying Gilbert's criteria.*

The poles of the controller shall be chosen to render the closed-loop stable and to satisfy some specifications. We choose (for example) to locate the poles of the controller at $\lambda_{K_1} = -1$ *and* $\lambda_{K_2} = -2$.

We will first design the controller assuming that we have the controllable canonical form of the SISO system. The controllable canonical form is readily obtained through $\det(s\mathbb{I} - \mathbf{A})$*:*

$$\begin{aligned} &\det(s\mathbb{I} - \mathbf{A}) = (s-1)(s-2) = s^2 - 3\,s + 2 := s^2 + a_1\,s + a_0 \\ &\Rightarrow \mathbf{A}_c = \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} \; \textit{and}\; \mathbf{B}_c = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned} \tag{5.24}$$

On the other hand the characteristic polynomial of the controller is formed thanks to the predefined eigenvalues assigned for the dynamics of the controller:

$$\begin{aligned} \chi_{A-BK}(s) &= (s - \lambda_{K_1})\,(s - \lambda_{K_2}) = (s+1)\,(s+2) \\ &= s^2 + 3\,s + 2 := s^2 + p_1\,s + p_0 \end{aligned} \tag{5.25}$$

Applying relation (5.16) *we get:*

$$\begin{aligned} \mathbf{K}_c &= \begin{bmatrix} K_1 & K_2 \end{bmatrix} = \begin{bmatrix} p_0 - a_0 & p_1 - a_1 \end{bmatrix} \\ &= \begin{bmatrix} 2-2 & 3+3 \end{bmatrix} = \begin{bmatrix} 0 & 6 \end{bmatrix} \end{aligned} \tag{5.26}$$

Now let's compute the inverse of the similarity transformation matrix $\mathbf{P}_c^{-1}$ *which enables to get the controllable canonical form.*

$$\mathbf{P}_c^{-1} = \mathbf{Q}_{cc}\mathbf{Q}_c^{-1} \tag{5.27}$$

Where:

- $\mathbf{Q}_c$ *is the controllability matrix in the actual basis:*

$$\mathbf{Q}_c = \begin{bmatrix} \mathbf{B} & \mathbf{AB} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} \tag{5.28}$$

- *and* $\mathbf{Q}_{cc}$ *the controllability matrix expressed in the controllable canonical basis (which is readily obtained through* $\det(s\mathbb{I} - \mathbf{A})$*):*

$$\mathbf{Q}_{cc} = \begin{bmatrix} \mathbf{B}_c & \mathbf{A}_c\mathbf{B}_c \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \tag{5.29}$$

Thus:

$$\begin{aligned} \mathbf{P}_c^{-1} &= \mathbf{Q}_{cc}\mathbf{Q}_c^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}^{-1} \\ &= \tfrac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ -2 & 1 \end{bmatrix} \\ &= \tfrac{1}{2} \begin{bmatrix} -2 & 1 \\ -2 & 2 \end{bmatrix} \end{aligned} \tag{5.30}$$

We finally get:

$$\mathbf{K} = \mathbf{K}_c\mathbf{P}_c^{-1} = \frac{1}{2} \begin{bmatrix} 0 & 6 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ -2 & 2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -12 & 12 \end{bmatrix} = \begin{bmatrix} -6 & 6 \end{bmatrix} \tag{5.31}$$

The feedforward gain matrix $\mathbf{H}$ *is computed thanks to* (5.7) *(where* $\mathbf{D} = \mathbf{0}$*):*

$$\mathbf{H} = -\left(\mathbf{C}\left(\mathbf{A} - \mathbf{BK}\right)^{-1}\mathbf{B}\right)^{-1} = -0.125 \tag{5.32}$$

■

5.3.3 Ackermann's formula

Ackermann's formula states that the controller gain matrix $\mathbf{K}$ of a SISO system in arbitrary state-space representation can be obtained as the product between vector $\underline{q}_c^T$ and the assigned characteristic polynomial $\chi_{A-BK}(s)$ of matrix $\mathbf{A} - \mathbf{KC}$ evaluated at matrix $\mathbf{A}$:

$$\mathbf{K} = \underline{q}_c^T \chi_{A-BK}(\mathbf{A}) \tag{5.33}$$

To get this result we first recall that similarity transformation generates equivalent state-space representations. Let $\mathbf{P}_c$ be the matrix of the similarity transformation which enables to get the controllable canonical form. Starting

from a state-space representations $\left(\begin{array}{c|c}\mathbf{A} & \mathbf{B}\\ \hline \mathbf{C} & \mathbf{D}\end{array}\right)$ in an arbitrary basis, the controllable canonical form is obtained through the following relations:

$$\begin{cases} \mathbf{A}_c = \mathbf{P}_c^{-1}\mathbf{A}\mathbf{P}_c \\ \mathbf{B}_c = \mathbf{P}_c^{-1}\mathbf{B} \\ \mathbf{C}_c = \mathbf{C}\mathbf{P}_c \end{cases} \tag{5.34}$$

Consequently matrix $\mathbf{A}_c - \mathbf{B}_c\mathbf{K}_c$ reads:

$$\begin{aligned} \mathbf{A}_c - \mathbf{B}_c\mathbf{K}_c &= \mathbf{P}_c^{-1}\mathbf{A}\mathbf{P}_c - \mathbf{P}_c^{-1}\mathbf{B}\mathbf{K}_c \\ &= \mathbf{P}_c^{-1}\left(\mathbf{A} - \mathbf{B}\mathbf{K}_c\mathbf{P}_c^{-1}\right)\mathbf{P}_c \end{aligned} \tag{5.35}$$

This equation indicates that the controller gain matrix $\mathbf{K}$ in arbitrary state-space representation reads:

$$\mathbf{K} = \mathbf{K}_c\mathbf{P}_c^{-1} \tag{5.36}$$

We have seen in the chapter dedicated to *Realization of transfer functions* that $\mathbf{P}_c^{-1}$ is a constant nonsingular change of basis matrix which is obtained through the state matrix $\mathbf{A}$ and vector $\underline{q}_c^T$:

$$\mathbf{P}_c^{-1} = \begin{bmatrix} \underline{q}_c^T \\ \underline{q}_c^T\mathbf{A} \\ \vdots \\ \underline{q}_c^T\mathbf{A}^{n-1} \end{bmatrix} \tag{5.37}$$

Vector $\underline{q}_c^T$ is the last row of the inverse of the controllability matrix $\mathbf{Q}_c$:

$$\mathbf{Q}_c^{-1} = \begin{bmatrix} * \\ \vdots \\ * \\ \underline{q}_c^T \end{bmatrix} \quad \text{where} \quad \mathbf{Q}_c = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \tag{5.38}$$

Then we recall Cayley–Hamilton theorem. Let $\chi_{A_c}(s)$ be the characteristic polynomial of the $n \times n$ matrix $\mathbf{A}_c$. We write the characteristic polynomial $\chi_{A_c}(s)$ of matrix $\mathbf{A}_c$ as follows:

$$\chi_{A_c}(s) := \det(s\mathbb{I} - \mathbf{A}_c) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \tag{5.39}$$

The Cayley–Hamilton theorem states that substituting matrix $\mathbf{A}_c$ for s in the characteristic polynomial $\chi_{A_c}(s)$ of matrix $\mathbf{A}_c$ results in the zero matrix[2]:

$$\chi_{A_c}(\mathbf{A}_c) = \mathbf{0} = \mathbf{A}_c^n + a_{n-1}\mathbf{A}_c^{n-1} + \cdots + a_1\mathbf{A}_c + a_0\mathbb{I} \tag{5.40}$$

Let $\chi_{A-BK}(s)$ be the characteristic polynomial of matrix $\mathbf{A}_c - \mathbf{B}_c\mathbf{K}_c$. We have seen that when predefined eigenvalues $\lambda_{K_1}, \cdots, \lambda_{K_n}$ are assigned for the

[2] https://en.wikipedia.org/wiki/Cayley–Hamilton_theorem

dynamics of the controller the characteristic polynomial of matrix $\mathbf{A}_c - \mathbf{B}_c\mathbf{K}_c$ reads:

$$\begin{aligned}\chi_{A-BK}(s) &= \det\left(s\mathbb{I} - (\mathbf{A}_c - \mathbf{B}_c\mathbf{K}_c)\right)\\ &= (s-\lambda_{K_1})\cdots(s-\lambda_{K_n})\\ &= s^n + p_{n-1}s^{n-1} + \cdots + p_1 s + p_0\end{aligned} \tag{5.41}$$

Substituting s for matrix $\mathbf{A}_c$ leads to the following relation:

$$\chi_{A-BK}(\mathbf{A}_c) = \mathbf{A}_c^n + p_{n-1}\mathbf{A}_c^{n-1} + \cdots + p_1\mathbf{A}_c + p_0\mathbb{I} \tag{5.42}$$

Note that $\chi_{A-BK}(\mathbf{A}_c)$ is not equal to $\mathbf{0}$ because $\chi_{A-BK}(s)$ is not the characteristic polynomial of matrix $\mathbf{A}_c$.

Thanks to Equation (5.16) and the relation $p_i = a_i + K_i$ we get:

$$\begin{aligned}\chi_{A-BK}(\mathbf{A}_c) = \mathbf{A}_c^n + (a_{n-1}+K_{n-1})\,\mathbf{A}_c^{n-1} + \cdots\\ + (a_1+K_1)\,\mathbf{A}_c + (a_0+K_0)\,\mathbb{I}\end{aligned} \tag{5.43}$$

By subtracting Equations (5.40) to (5.43) we get:

$$\chi_{A-BK}(\mathbf{A}_c) = K_{n-1}\mathbf{A}_c^{n-1} + \cdots + K_1\mathbf{A}_c + K_0\mathbb{I} \tag{5.44}$$

Due to the fact that coefficients K_i are scalar we can equivalently write:

$$\chi_{A-BK}(\mathbf{A}_c) = \mathbf{A}_c^{n-1}K_{n-1} + \cdots + \mathbf{A}_c K_1 + \mathbb{I}K_0 \tag{5.45}$$

Let $\underline{u}^T$ be the vector defined by:

$$\underline{u}^T = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \tag{5.46}$$

Due to the *special* form of matrix $\mathbf{A}_c$ we have:

$$\begin{aligned}\underline{u}^T\mathbf{A}_c &= \begin{bmatrix} 0 & 1 & 0 & \cdots & \end{bmatrix}\\ \underline{u}^T\mathbf{A}_c^2 &= \begin{bmatrix} 0 & 0 & 1 & 0 & \cdots \end{bmatrix}\\ &\vdots\\ \underline{u}^T\mathbf{A}_c^{n-1} &= \begin{bmatrix} 0 & \cdots & 0 & 1 & \end{bmatrix}\end{aligned} \tag{5.47}$$

Thus multiplying Equation (5.45) by $\underline{u}^T$ leads to the following relation:

$$\underline{u}^T\chi_{A-BK}(\mathbf{A}_c) = \begin{bmatrix} K_0 & K_1 & \cdots & K_{n-1} \end{bmatrix} = \mathbf{K}_c \tag{5.48}$$

Thus we get the expression of the controller matrix $\mathbf{K}_c$ when we use the controllable canonical form.

We multiply Equation (5.48) by $\mathbf{P}_c^{-1}$ and use the fact that $\mathbf{A}_c^k = \left(\mathbf{P}_c^{-1}\mathbf{A}\mathbf{P}_c\right)^k = \mathbf{P}_c^{-1}\mathbf{A}^k\mathbf{P}_c$ get the expression of the controller gain matrix $\mathbf{K}$ in arbitrary state-space representation:

$$\begin{aligned}\mathbf{K} &= \mathbf{K}_c\mathbf{P}_c^{-1}\\ &= \underline{u}^T\chi_{A-BK}(\mathbf{A}_c)\mathbf{P}_c^{-1}\\ &= \underline{u}^T\chi_{A-BK}(\mathbf{A}_c\mathbf{P}_c^{-1})\\ &= \underline{u}^T\chi_{A-BK}(\mathbf{P}_c^{-1}\mathbf{P}_c\mathbf{A}_c\mathbf{P}_c^{-1})\\ &= \underline{u}^T\mathbf{P}_c^{-1}\chi_{A-BK}(\mathbf{P}_c\mathbf{A}_c\mathbf{P}_c^{-1})\\ &= \underline{u}^T\mathbf{P}_c^{-1}\chi_{A-BK}(\mathbf{A})\end{aligned} \tag{5.49}$$

Because $\underline{u}$ is the vector defined by $\underline{u}^T = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$ we get using (5.37):

$$\underline{u}^T \mathbf{P}_c^{-1} = \underline{q}_c^T \tag{5.50}$$

Consequently Equation (5.49) reduces to be the Ackermann's formula (5.33):

$$\mathbf{K} = \underline{q}_c^T \chi_{A-BK}(\mathbf{A}) \tag{5.51}$$

5.3.4 Invariance of (transmission) zeros under state feedback

It is worth noticing that the zeros of the closed-loop transfer function are the same than the zeros of the open-loop transfer function when state feedback is used. In other words, state feedback $\mathbf{K}$ just changes the values poles of the poles, the zeros remaining unchanged.

To get this result, we can use the controllable canonical form to compute the closed-loop transfer function, $\mathbf{C}_c \left(s\mathbb{I} - \mathbf{A}_c + \mathbf{B}_c\mathbf{K}_c\right)^{-1} \mathbf{B}_c\mathbf{H}$, and notice that its numerator is independent of both state feedback gain $\mathbf{K}_c$ and state matrix $\mathbf{A}_c$. When the controllable canonical form is used, it is clear that coefficients which appear in the numerator of the closed-loop transfer function come from product $\mathbf{C}_c\mathbf{H}$ and from the last column of the adjugate of the closed-loop state matrix, which is $\begin{bmatrix} 1 & s & \cdots & s^{n-1} \end{bmatrix}^T$:

$$\left(s\mathbb{I} - \mathbf{A}_c + \mathbf{B}_c\mathbf{K}_c\right)^{-1} \mathbf{B}_c = \left(\frac{1}{\det\left(s\mathbb{I} - \mathbf{A}_c + \mathbf{B}_c\mathbf{K}_c\right)} \begin{bmatrix} * & * & 1 \\ * & * & s \\ \vdots & \vdots & \vdots \\ * & * & s^{n-1} \end{bmatrix} \right) \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$$\Rightarrow \mathbf{C}_c \left(s\mathbb{I} - \mathbf{A}_c + \mathbf{B}_c\mathbf{K}_c\right)^{-1} \mathbf{B}_c\mathbf{H} = \frac{\mathbf{C}_c}{\det\left(s\mathbb{I} - \mathbf{A}_c + \mathbf{B}_c\mathbf{K}_c\right)} \begin{bmatrix} 1 \\ s \\ \vdots \\ s^{n-1} \end{bmatrix} \mathbf{H} \tag{5.52}$$

Consequently state feedback gain $\mathbf{K}_c$, or $\mathbf{K}$ in general, does not affect the numerator of the closed-loop transfer function. A practical use of this observation is that state feedback gain $\mathbf{K}$ can NOT be used to change zeros location in a state feedback loop.

5.4 Static state feedback controller

5.4.1 Roppenecker's formula

We consider the following state-space representation where $\underline{y}(t) \in \mathbb{R}^p$ denotes the measured output signal and $\underline{u}(t) \in \mathbb{R}^m$ the control input:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{5.53}$$

We will assume in the following that only the output vector $\underline{y}(t)$ is available for control.

Let $\underline{r}(t)$ be a reference input signal. A static output feedback controller computes the control input $\underline{u}(t)$ as a function of a state-feedback gain $\mathbf{K}$ and a feedforward gain matrix $\mathbf{H}$ as follows:

$$\underline{u}(t) = -\mathbf{K}\underline{y}(t) + \mathbf{H}\underline{r}(t) \tag{5.54}$$

Using the output equation $\underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t)$ the control input $\underline{u}(t)$ can be expressed as a function of the state vector $\underline{x}(t)$ and the reference input $\underline{r}(t)$:

$$\begin{aligned} &\underline{u}(t) = -\mathbf{K}\left(\mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t)\right) + \mathbf{H}\underline{r}(t) \\ &\Rightarrow \underline{u}(t) = \left(\mathbb{I} + \mathbf{KD}\right)^{-1}\left(-\mathbf{KC}\underline{x}(t) + \mathbf{H}\underline{r}(t)\right) \end{aligned} \tag{5.55}$$

Substituting the control law (5.55) into the state equation (5.53) of the system reads:

$$\begin{aligned} \dot{\underline{x}}(t) &= \mathbf{A}\underline{x}(t) + \mathbf{B}\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\left(-\mathbf{KC}\underline{x}(t) + \mathbf{H}\underline{r}(t)\right) \\ &= \left(\mathbf{A} - \mathbf{B}\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{KC}\right)\underline{x}(t) + \mathbf{B}\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{H}\underline{r}(t) \end{aligned} \tag{5.56}$$

We denote by $\mathbf{A}_{cl}$ the closed-loop state matrix:

$$\mathbf{A}_{cl} = \mathbf{A} - \mathbf{B}\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{KC} \tag{5.57}$$

It is worth noticing that in the special case where the feedforward gain matrix $\mathbf{D}$ is zero ($\mathbf{D} = \mathbf{0}$) and where the output matrix $\mathbf{C}$ is equal to identity ($\mathbf{C} = \mathbb{I}$) then the static *output* feedback controller $\mathbf{K}$ reduces to be a static *state* feedback controller.

Let $\lambda_{K_1}, \cdots, \lambda_{K_n}$ be n distinct specified eigenvalues of the closed-loop state matrix $\mathbf{A}_{cl}$. Furthermore we assume that eigenvalues of matrix $\mathbf{A}$ do not shift (meaning that they are different) the eigenvalues λ_{K_i} of the closed-loop state matrix $\mathbf{A}_{cl}$. Let $\underline{v}_{K_i}$ be an eigenvector of the closed-loop state matrix $\mathbf{A}_{cl}$ corresponding to eigenvalue λ_{K_i}:

$$\left(\mathbf{A} - \mathbf{B}\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{KC}\right)\underline{v}_{K_i} = \lambda_{K_i}\underline{v}_{K_i} \tag{5.58}$$

The preceding relation can be rewritten as follows:

$$\left(\mathbf{A} - \lambda_{K_i}\mathbb{I}\right)\underline{v}_{K_i} - \mathbf{B}\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{KC}\underline{v}_{K_i} = \underline{0} \tag{5.59}$$

Let's $\underline{p}_i$ be the parameter vector which is actually the input direction corresponding to eigenvector $\underline{v}_{K_i}$:

$$\underline{p}_i = -\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{KC}\underline{v}_{K_i} \tag{5.60}$$

Combining Equations (5.59) and (5.60) leads to the following relation:

$$\begin{bmatrix} \mathbf{A} - \lambda_{K_i}\mathbb{I} & \mathbf{B} \end{bmatrix}\begin{bmatrix} \underline{v}_{K_i} \\ \underline{p}_i \end{bmatrix} = \underline{0} \tag{5.61}$$

From (5.61) it is clear that vector $\begin{bmatrix} \underline{v}_{K_i} \\ \underline{p}_i \end{bmatrix}$ belongs to the kernel of matrix $\begin{bmatrix} \mathbf{A} - \lambda_{K_i}\mathbb{I} & \mathbf{B} \end{bmatrix}$:

$$\begin{bmatrix} \underline{v}_{K_i} \\ \underline{p}_i \end{bmatrix} \in \mathcal{N}\left(\begin{bmatrix} \mathbf{A} - \lambda_{K_i}\mathbb{I} & \mathbf{B} \end{bmatrix}\right) \tag{5.62}$$

Assuming that the rank of matrix $\mathbf{A} - \lambda_{K_i}\mathbb{I}$ is n, matrix $\begin{bmatrix} \mathbf{A} - \lambda_{K_i}\mathbb{I} & \mathbf{B} \end{bmatrix}$ is of rank n. For a system with m inputs, $\mathbf{B}$ has m columns and (5.61) represents o system with n linearly independant equations and $n+m$ unknowns $\begin{bmatrix} \underline{v}_{K_i} \\ \underline{p}_i \end{bmatrix}$. Thus we can choose arbitrary m unknowns. In the following, we will assume that $\underline{p}_i$ is arbitrarily chosen. Nevertheless, in general, the m unknowns parameters can be split indifferently either on $\underline{v}_{K_i}$ or $\underline{p}_i$.

Writing (5.61) for all the distinct predefined eigenvalues $\mathbf{\Lambda}_{cl} = diag(\lambda_{K_1}, \cdots, \lambda_{K_p})$ of the closed-loop state matrix gives:

$$\begin{aligned} &\begin{bmatrix} \mathbf{A} - \lambda_{K_i}\mathbb{I} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \underline{v}_{K_i} \\ \underline{p}_i \end{bmatrix} = \underline{0} \\ &\Leftrightarrow \mathbf{A}\underline{v}_{K_i} - \underline{v}_{K_i}\lambda_{K_i} + \mathbf{B}\underline{p}_i = \underline{0} \quad \forall\, i = 1, \cdots, n \end{aligned} \tag{5.63}$$

This leads to the so-called Sylvester matrix equation:

$$\mathbf{AV} - \mathbf{V}\mathbf{\Lambda}_{cl} + \mathbf{BP} = \mathbf{0} \tag{5.64}$$

Where matrices $\mathbf{P}$ and $\mathbf{V}$ are defined as follows:

$$\begin{cases} \mathbf{P} = \begin{bmatrix} \underline{p}_1 & \cdots & \underline{p}_p \end{bmatrix} = -\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{KCV} \\ \mathbf{V} = \begin{bmatrix} \underline{v}_{K_1} & \cdots & \underline{v}_{K_p} \end{bmatrix} \end{cases} \tag{5.65}$$

It is clear that as soon as $\underline{v}_{K_i} = -\left(\mathbf{A} - \lambda_{K_i}\mathbb{I}\right)^{-1}\mathbf{B}\underline{p}_i$ then kernel equation (5.61) is solved. Consequently matrices $\mathbf{P}$ and $\mathbf{V}$ satisfying Sylvester matrix equation (5.64) are obtained as follows where $m \times r$ parameter matrix $\mathbf{P}$ is a real matrix of rank m:

$$\begin{cases} \mathbf{P} = \begin{bmatrix} \underline{p}_1 & \cdots & \underline{p}_p \end{bmatrix} \text{ where } \operatorname{rank}(\mathbf{P}) = m \\ \mathbf{V} = \begin{bmatrix} \mathbf{W}_1\underline{p}_1 & \cdots & \mathbf{W}_p\underline{p}_p \end{bmatrix} \end{cases} \tag{5.66}$$

Where:

$$\begin{cases} \underline{p}_i = \begin{bmatrix} p_{i1} & \cdots & p_{im} \end{bmatrix}^T \\ \mathbf{W}_i = -\left(\mathbf{A} - \lambda_{K_i}\mathbb{I}\right)^{-1}\mathbf{B} \end{cases} \tag{5.67}$$

There are p vectors $\underline{p}_i$, $i = 1, \cdots, p$ to determine, each of size m, thus $m \times p$ unknowns. Thus the number n of required eigenvalues λ_{K_i}, $i = 1, \cdots, n$ shall be such that $m \times p \geq n$.

From the definition of matrix $\mathbf{P}$ given in (5.65) we get:

$$\begin{aligned} &\mathbf{P} = -\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{KCV} \\ &\Leftrightarrow \mathbf{P} = -\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{KCV} \\ &\Leftrightarrow \left(\mathbb{I} + \mathbf{KD}\right)\mathbf{P} = -\mathbf{KCV} \\ &\Leftrightarrow \mathbf{K}\left(\mathbf{CV} + \mathbf{DP}\right) = -\mathbf{P} \end{aligned} \tag{5.68}$$

Usually matrix $\mathbf{CV} + \mathbf{DP}$ is not square. Consequently the static output feedback gain $\mathbf{K}$ is obtained by taking the pseudo-inverse $(\mathbf{CV} + \mathbf{DP})^{\dagger}$ of matrix $\mathbf{CV} + \mathbf{DP}$:

$$\begin{gathered}\mathbf{K} = -\mathbf{P}\,(\mathbf{CV} + \mathbf{DP})^{\dagger} \\ \text{where } (\mathbf{CV} + \mathbf{DP})^{\dagger} := (\mathbf{CV} + \mathbf{DP})^{T} \left((\mathbf{CV} + \mathbf{DP})\,(\mathbf{CV} + \mathbf{DP})^{T}\right)^{-1}\end{gathered} \tag{5.69}$$

In the special case where matrix $\mathbf{CV} + \mathbf{DP}$ is square and invertible the preceding relation reads:

$$\mathbf{K} = -\mathbf{P}\,(\mathbf{CV} + \mathbf{DP})^{-1} \tag{5.70}$$

Or equivalently:

$$\mathbf{K} = -\left[\begin{array}{ccc} \underline{p}_1 & \cdots & \underline{p}_n \end{array}\right]\left(\mathbf{C}\left[\begin{array}{ccc} \underline{v}_{K_1} & \cdots & \underline{v}_{K_n} \end{array}\right] + \mathbf{D}\left[\begin{array}{ccc} \underline{p}_1 & \cdots & \underline{p}_n \end{array}\right]\right)^{-1} \tag{5.71}$$

The preceding relation is the so-called Roppenecker's formula[3] to get the static *output* feedback gain matrix $\mathbf{K}$.

5.4.2 Comments on Roppenecker's formula

We recall that the n distinct eigenvalues λ_{K_i} of the closed-loop state matrix and the corresponding eigenvectors $\underline{v}_{K_i}$ are related to the parameter vectors $\underline{p}_i$ by relation (5.61) which is reported hereafter:

$$\left[\begin{array}{cc} \mathbf{A} - \lambda_{K_i}\mathbb{I} & \mathbf{B} \end{array}\right]\left[\begin{array}{c} \underline{v}_{K_i} \\ \underline{p}_i \end{array}\right] = \underline{0} \tag{5.72}$$

It is worth noticing the following facts:

- From the relation:

$$\left[\begin{array}{cc} \mathbf{A} - \lambda_{K_i}\mathbb{I} & \mathbf{B} \end{array}\right]\left[\begin{array}{c} \underline{v}_{K_i} \\ \underline{p}_i \end{array}\right] = \underline{0} \tag{5.73}$$

  It is clear that each $(n+m) \times 1$ vector $\left[\begin{array}{c} \underline{v}_{K_i} \\ \underline{p}_i \end{array}\right]$ belongs to the kernel of matrix $\left[\begin{array}{cc} \mathbf{A} - \lambda_{K_i}\mathbb{I} & \mathbf{B} \end{array}\right]$. So once any $(n+m) \times 1$ vector which belongs to the kernel of matrix $\left[\begin{array}{cc} \mathbf{A} - \lambda_{K_i}\mathbb{I} & \mathbf{B} \end{array}\right]$ has been found, its m bottom rows are used to form vector parameter $\underline{p}_i$. In the MIMO case several possibilities are offered.

- We recognize in matrix $\left[\begin{array}{cc} \mathbf{A} - \lambda_{K_i}\mathbb{I} & \mathbf{B} \end{array}\right]$ the key matrix used in the PBH controllability test;

[3] G. Roppenecker, On Parametric State Feedback Design, International Journal of Control, Volume 43, 1986 - Issue 3

- In the special case of *state* feedback where $\mathbf{D} = \mathbf{0}$ and $\mathbf{C} = \mathbb{I}$ then matrix $\mathbf{CV} + \mathbf{DP} = \mathbf{V}$ where $\mathbf{V}$ is a square and invertible matrix. In that case Equation (5.70) leads to the expression of the static *state* feedback gain matrix $\mathbf{K}$:

$$\left\{ \begin{array}{l} \mathbf{C} = \mathbb{I} \\ \mathbf{D} = \mathbf{0} \end{array} \right. \Rightarrow \mathbf{K} = -\mathbf{P}\mathbf{V}^{-1} = -\left[\begin{array}{ccc} \underline{p}_1 & \cdots & \underline{p}_n \end{array} \right] \left(\left[\begin{array}{ccc} \underline{v}_{K_1} & \cdots & \underline{v}_{K_n} \end{array} \right] \right)^{-1} \tag{5.74}$$

- If we wish to keep an eigenvalue λ_i of matrix $\mathbf{A}$ within the set of eigenvalues of the closed-loop state matrix $\mathbf{A}_{cl}$ then $(\mathbf{A} - \lambda_{K_i}\mathbb{I})\,\underline{v}_{K_i}$ is equal to zero because in that case $\underline{v}_{K_i}$ is also an eigenvector of $\mathbf{A}$:

$$\lambda_{K_i} = \lambda_i \Rightarrow (\mathbf{A} - \lambda_{K_i}\mathbb{I})\,\underline{v}_{K_i} = -\mathbf{B}\underline{p}_i = (\mathbf{A} - \lambda_i\mathbb{I})\,\underline{v}_{K_i} = \underline{0} \tag{5.75}$$

  Consequently we have to replace $\underline{p}_i$ by $\underline{0}$ and $\underline{v}_{K_i}$ by eigenvector $\underline{v}_i$ of $\mathbf{A}$ corresponding to λ_i in the Roppenecker's formula (5.70);

- The static output feedback gain $\mathbf{K}$ satisfy the following relation:

$$\mathbf{K}\left(\mathbf{C}\underline{v}_{K_i} + \mathbf{D}\underline{p}_i\right) = -\underline{p}_i \tag{5.76}$$

  Indeed by combining Equations (5.61) and (5.76) we retrieve Equation (5.58):

$$\begin{array}{l} \left\{ \begin{array}{l} \mathbf{A}\underline{v}_{K_i} + \mathbf{B}\underline{p}_i = \lambda_{K_i}\underline{v}_{K_i} \\ \mathbf{K}\left(\mathbf{C}\underline{v}_{K_i} + \mathbf{D}\underline{p}_i\right) = -\underline{p}_i \Leftrightarrow \underline{p}_i = -\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{KC}\underline{v}_{K_i} \end{array} \right. \\ \Rightarrow \left(\mathbf{A} - \mathbf{B}\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{KC}\right)\underline{v}_{K_i} = \lambda_{K_i}\underline{v}_{K_i} \end{array} \tag{5.77}$$

  Conversely we can write the preceding equation as follows:

$$\left[\begin{array}{cc} \mathbf{A} - \lambda_{K_i}\mathbb{I} & \mathbf{B} \end{array} \right] \left[\begin{array}{c} \underline{v}_{K_i} \\ -\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{KC}\underline{v}_{K_i} \end{array} \right] = \underline{0} \tag{5.78}$$

  Thus by defining parameter vector $\underline{p}_i$ as $\underline{p}_i = -\left(\mathbb{I} + \mathbf{KD}\right)^{-1}\mathbf{KC}\underline{v}_{K_i}$ we retrieve $\mathbf{K}\left(\mathbf{C}\underline{v}_{K_i} + \mathbf{D}\underline{p}_i\right) = -\underline{p}_i$ which is exactly Equation (5.76).

- In the SISO case where $\mathbf{D} = \mathbf{0}$ and $\mathbf{C} = \mathbb{I}$, that is where a *state* feedback is assumed, the controller gain matrix $\mathbf{K}$ no more depends on parameter vectors $\underline{p}_i$. Indeed is that case the controller gain matrix $\mathbf{K}$ is obtained as follows:

$$\mathbf{K} = \left[\begin{array}{ccc} 1 & \cdots & 1 \end{array} \right] \left[\begin{array}{ccc} (\mathbf{A} - \lambda_{K_1}\mathbb{I})^{-1}\mathbf{B} & \cdots & (\mathbf{A} - \lambda_{K_n}\mathbb{I})^{-1}\mathbf{B} \end{array} \right]^{-1} \tag{5.79}$$

To get this result we start by observing that in the SISO case where $\mathbf{D} = \mathbf{0}$ and $\mathbf{C} = \mathbb{I}$ parameter vector are scalars; they will be denoted p_i. Let vector $\underline{K}_i$ be defined as follows:

$$\underline{K}_i = -\left(\mathbf{A} - \lambda_{K_i}\mathbb{I}\right)^{-1}\mathbf{B} \tag{5.80}$$

Then Equation (5.71) reads:

$$\mathbf{K} = -\begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix}\begin{bmatrix} \underline{K}_1 p_1 & \cdots & \underline{K}_n p_n \end{bmatrix}^{-1} \tag{5.81}$$

Let's rearrange the term $\begin{bmatrix} \underline{K}_1 p_1 & \cdots & \underline{K}_n p_n \end{bmatrix}^{-1}$ as follows:

$$\begin{aligned}
\begin{bmatrix} \underline{K}_1 p_1 & \cdots & \underline{K}_n p_n \end{bmatrix}^{-1} &= \left(\begin{bmatrix} \underline{K}_1 & \cdots & \underline{K}_n \end{bmatrix}\begin{bmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_n \end{bmatrix}\right)^{-1} \\
&= \begin{bmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_n \end{bmatrix}^{-1}\begin{bmatrix} \underline{K}_1 & \cdots & \underline{K}_n \end{bmatrix}^{-1} \\
&= \begin{bmatrix} \frac{\prod_{i\neq 1} p_i}{\prod_{i=1}^{n} p_i} & & 0 \\ & \ddots & \\ 0 & & \frac{\prod_{i\neq n} p_i}{\prod_{i=1}^{n} p_i} \end{bmatrix}\begin{bmatrix} \underline{K}_1 & \cdots & \underline{K}_n \end{bmatrix}^{-1}
\end{aligned} \tag{5.82}$$

Multiplying this expression by $-\begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix}$ leads to the expression of $\mathbf{K}$:

$$\begin{aligned}
\mathbf{K} &= -\begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix}\begin{bmatrix} \underline{K}_1 p_1 & \cdots & \underline{K}_n p_n \end{bmatrix}^{-1} \\
&= -\begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix}\begin{bmatrix} \frac{\prod_{i\neq 1} p_i}{\prod_{i=1}^{n} p_i} & & 0 \\ & \ddots & \\ 0 & & \frac{\prod_{i\neq n} p_i}{\prod_{i=1}^{n} p_i} \end{bmatrix}\begin{bmatrix} \underline{K}_1 & \cdots & \underline{K}_n \end{bmatrix}^{-1} \\
&= -\begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}\begin{bmatrix} \underline{K}_1 & \cdots & \underline{K}_n \end{bmatrix}^{-1}
\end{aligned} \tag{5.83}$$

Using the expression of vector $\underline{K}_i = -\left(\mathbf{A} - \lambda_{K_i}\mathbb{I}\right)^{-1}\mathbf{B}$ provided by Equation (5.80) we finally get:

$$\begin{aligned}
\mathbf{K} = &\begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \\
&\begin{bmatrix} \left(\mathbf{A} - \lambda_{K_1}\mathbb{I}\right)^{-1}\mathbf{B} & \cdots & \left(\mathbf{A} - \lambda_{K_n}\mathbb{I}\right)^{-1}\mathbf{B} \end{bmatrix}^{-1}
\end{aligned} \tag{5.84}$$

We conclude that in the SISO case where $\mathbf{D} = \mathbf{0}$ and $\mathbf{C} = \mathbb{I}$, that is where a *state* feedback is assumed, the controller gain matrix $\mathbf{K}$ no more depends on parameter vectors $\underline{p}_i$.

5.4.3 Solving Sylvester equation

The Sylvester equation reads as follows where $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{m\times m}$, $\mathbf{C} \in \mathbb{R}^{n\times m}$ and $\mathbf{X} \in \mathbb{R}^{n\times m}$ is the sought-after solution:

$$\mathbf{AX} + \mathbf{XB} + \mathbf{C} = \mathbf{0} \tag{5.85}$$

Following Peter Benner[4], Sylvester equation (5.85) has a unique solution if and only if $\alpha + \beta \neq 0$ for all $\alpha \in \sigma(\mathbf{A})$ and $\beta \in \sigma(\mathbf{B})$, where $\sigma(\mathbf{Z})$ denotes the spectrum of the matrix $\mathbf{Z}$. In particular, this property holds for *stable* Sylvester equations, where both $\sigma(\mathbf{A})$ and $\sigma(\mathbf{B})$ lie in the open left half plane. The anti-stable case, where $\sigma(\mathbf{A})$ and $\sigma(\mathbf{B})$ are contained in the open right half plane, is trivially turned into a stable one by multiplying (5.85) by -1. Furthermore, the unique solution $\mathbf{X}$ of a *stable* Sylvester equation can be found from the block upper triangular matrix $\mathbf{H}$ as follows[4]:

$$\mathbf{H} = \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0} & -\mathbf{B} \end{bmatrix} \Rightarrow \frac{1}{2}\left(sign(\mathbf{H}) + \mathbb{I}_{n+m}\right) = \begin{bmatrix} \mathbf{0} & \mathbf{X} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \tag{5.86}$$

We recall that if $\mathbf{H} = \mathbf{P}\,\mathbf{J}\,\mathbf{P}^{-1}$ is a Jordan decomposition of $\mathbf{H}$ with $diag(\mathbf{J}) = diag(\lambda_i)$ then $sign(\mathbf{H})$ is defined as follows:

$$\mathbf{H} = \mathbf{P}\,\mathbf{J}\,\mathbf{P}^{-1} \Rightarrow sign(\mathbf{H}) = \mathbf{P}\,sign(\mathbf{J})\,\mathbf{P}^{-1} = \mathbf{P}\,diag\left(sign(\lambda_i)\right)\,\mathbf{P}^{-1} \tag{5.87}$$

For a scalar complex variable λ, the sign function is defined as follows:

$$sign(\lambda) = \begin{cases} +1 \text{ if } Re(\lambda) > 0 \\ -1 \text{ if } Re(\lambda) < 0 \end{cases} \tag{5.88}$$

Alternatively, the unique solution of a *stable* Sylvester matrix equation (5.85) can be computed as follows[5]:

$$\mathbf{X} = \int_0^\infty e^{\mathbf{A}t}\,\mathbf{C}\,e^{\mathbf{B}t}\,dt \tag{5.89}$$

5.4.4 Solving general algebraic Riccati and Lyapunov equations

The general algebraic Riccati equation reads as follows where all matrices are square of dimension $n \times n$:

$$\mathbf{AX} + \mathbf{XB} + \mathbf{C} + \mathbf{XDX} = \mathbf{0} \tag{5.90}$$

Matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are known whereas matrix $\mathbf{X}$ has to be determined.

The general algebraic Lyapunov equation is obtained as a special case of the algebraic Riccati by setting $\mathbf{D} = \mathbf{0}$.

[4]Peter Benner, Factorized Solution of Sylvester Equations with Applications in Control, January 2004

[5]Zeyad Abdel Aziz Al-Zhour, New techniques for solving some matrix and matrix differential equations, Ain Shams Engineering Journal (2015) 6, 347–354

The general algebraic Riccati equation can be solved[6] by considering the following $2n \times 2n$ matrix $\mathbf{H}$:

$$\mathbf{H} = \begin{bmatrix} \mathbf{B} & \mathbf{D} \\ -\mathbf{C} & -\mathbf{A} \end{bmatrix} \tag{5.91}$$

Let the eigenvalues of matrix $\mathbf{H}$ be denoted λ_1, $i = 1, \cdots, 2n$, and the corresponding eigenvectors be denoted $\underline{v}_i$. Furthermore let $\mathbf{M}$ be the $2n \times 2n$ matrix composed of all real eigenvectors of matrix $\mathbf{H}$; for complex conjugate eigenvectors, the corresponding columns of matrix $\mathbf{M}$ are changed into the real and imaginary parts of such eigenvectors. Note that there are many ways to form matrix $\mathbf{M}$.

Then we can write the following relation:

$$\mathbf{HM} = \mathbf{M\Lambda} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{bmatrix} \tag{5.92}$$

Matrix $\mathbf{M}_1$ contains the n first columns of $\mathbf{M}$ whereas matrix $\mathbf{M}_2$ contains the n last columns of $\mathbf{M}$.

Matrices $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ are diagonal matrices formed by the eigenvalues of $\mathbf{H}$ as soon as there are distinct; for eigenvalues with multiplicity greater than 1, the corresponding part in matrix $\mathbf{\Lambda}$ represents the Jordan form.

Thus we have:

$$\begin{cases} \mathbf{HM}_1 = \mathbf{M}_1\mathbf{\Lambda}_1 \\ \mathbf{HM}_2 = \mathbf{M}_2\mathbf{\Lambda}_2 \end{cases} \tag{5.93}$$

We will focus our attention on the first equation and split matrix $\mathbf{M}_1$ as follows:

$$\mathbf{M}_1 = \begin{bmatrix} \mathbf{M}_{11} \\ \mathbf{M}_{12} \end{bmatrix} \tag{5.94}$$

Using the expression of $\mathbf{H}$ in (5.91), the relation $\mathbf{HM}_1 = \mathbf{M}_1\mathbf{\Lambda}_1$ reads as follows:

$$\mathbf{HM}_1 = \mathbf{M}_1\mathbf{\Lambda}_1 \Rightarrow \begin{cases} \mathbf{BM}_{11} + \mathbf{DM}_{12} = \mathbf{M}_{11}\mathbf{\Lambda}_1 \\ -\mathbf{CM}_{11} - \mathbf{AM}_{12} = \mathbf{M}_{12}\mathbf{\Lambda}_1 \end{cases} \tag{5.95}$$

Assuming that matrix $\mathbf{M}_{11}$ is not singular, we can check that a solution $\mathbf{X}$ of the general algebraic Riccati equation (5.90) reads:

$$\mathbf{X} = \mathbf{M}_{12}\mathbf{M}_{11}^{-1} \tag{5.96}$$

[6] Optimal Control of Singularly Perturbed Linear Systems with Applications: High Accuracy Techniques, Z. Gajic and M. Lim, Marcel Dekker, New York, 2001

Indeed:

$$\begin{cases} \mathbf{B}\mathbf{M}_{11} + \mathbf{D}\mathbf{M}_{12} = \mathbf{M}_{11}\mathbf{\Lambda}_1 \\ \mathbf{C}\mathbf{M}_{11} + \mathbf{A}\mathbf{M}_{12} = -\mathbf{M}_{12}\mathbf{\Lambda}_1 \\ \mathbf{X} = \mathbf{M}_{12}\mathbf{M}_{11}^{-1} \end{cases}$$

$$\begin{aligned} \Rightarrow \mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{B} + \mathbf{C} + \mathbf{X}\mathbf{D}\mathbf{X} &= \mathbf{A}\mathbf{M}_{12}\mathbf{M}_{11}^{-1} + \mathbf{M}_{12}\mathbf{M}_{11}^{-1}\mathbf{B} + \mathbf{C} \\ &\quad + \mathbf{M}_{12}\mathbf{M}_{11}^{-1}\mathbf{D}\mathbf{M}_{12}\mathbf{M}_{11}^{-1} \\ &= (\mathbf{A}\mathbf{M}_{12} + \mathbf{C}\mathbf{M}_{11})\,\mathbf{M}_{11}^{-1} \\ &\quad + \mathbf{M}_{12}\mathbf{M}_{11}^{-1}(\mathbf{B}\mathbf{M}_{11} + \mathbf{D}\mathbf{M}_{12})\,\mathbf{M}_{11}^{-1} \\ &= -\mathbf{M}_{12}\mathbf{\Lambda}_1\mathbf{M}_{11}^{-1} + \mathbf{M}_{12}\mathbf{M}_{11}^{-1}\mathbf{M}_{11}\mathbf{\Lambda}_1\mathbf{M}_{11}^{-1} \\ &= \mathbf{0} \end{aligned} \tag{5.97}$$

It is worth noticing that each selection of eigenvectors within matrix $\mathbf{M}_1$ leads to a new solution of the general algebraic Riccati equation (5.90). Consequently the solution to the general algebraic Riccati equation (5.90) is not unique. The same statement holds for different choice of matrix $\mathbf{M}_2$ and the corresponding solution of (5.90) obtained from $\mathbf{X} = \mathbf{M}_{21}\mathbf{M}_{22}^{-1}$.

5.5 Control with integral action

5.5.1 Adding an integrator in the state-space realization

We consider the state-space representation (5.98) where the state vector $\underline{x}$ is of dimension n (that is the size of state matrix $\mathbf{A}$). In addition $\underline{y}(t)$ denotes the output vector and $\underline{u}(t)$ the input vector. We will assume that the feedforward gain matrix $\mathbf{D}$ is zero ($\mathbf{D} = \mathbf{0}$):

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.98}$$

In some circumstances it may be helpful to use integral action in the controller design. This can be achieved by adding to the state vector of the state-space realization (5.98) the integral of the tracking error $\underline{e}_I(t)$ which is defined as follows where $\mathbf{T}$ is a design matrix (usually $\mathbf{T} = \mathbb{I}$) and where $\underline{r}(t)$ is the reference input signal:

$$\underline{e}_I(t) = \int_0^t \mathbf{T}\left(\underline{r}(\tau) - \underline{y}(\tau)\right)\, d\tau \tag{5.99}$$

As far as the feedforward gain matrix $\mathbf{D}$ is zero ($\mathbf{D} = \mathbf{0}$) we get:

$$\dot{\underline{e}}_I(t) = \mathbf{T}\left(\underline{r}(t) - \underline{y}(t)\right) = \mathbf{T}\underline{r}(t) - \mathbf{T}\mathbf{C}\underline{x}(t) \tag{5.100}$$

This leads to the following augmented state-space realization:

$$\begin{cases} \begin{bmatrix} \dot{\underline{x}}(t) \\ \dot{\underline{e}}_I(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ -\mathbf{T}\mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \underline{x}(t) \\ \underline{e}_I(t) \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix} \underline{u}(t) + \begin{bmatrix} \mathbf{0} \\ \mathbf{T} \end{bmatrix} \underline{r}(t) \\ \underline{y}(t) = \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \underline{x}(t) \\ \underline{e}_I(t) \end{bmatrix} \end{cases} \tag{5.101}$$

The regulation problem deals with the case where $\underline{r}(t) = \underline{0}$. In that situation the preceding augmented state-space realization has the same structure than the state-space realization (5.98). Thus the same techniques may be applied for the purpose of regulator design.

On the other hand the tracking problem deals with the case where $\underline{r}(t) \neq \underline{0}$. Let's denote $\underline{x}_a(t)$ the augmented state-space vector:

$$\underline{x}_a(t) = \begin{bmatrix} \underline{x}(t) \\ \underline{e}_I(t) \end{bmatrix} \tag{5.102}$$

Thus the augmented state-space realization (5.101) reads:

$$\begin{cases} \dot{\underline{x}}_a(t) = \mathbf{A}_a\, \underline{x}_a(t) + \mathbf{B}_a\, \underline{u}(t) + \mathbf{E}\, \underline{r}(t) \\ \underline{y}(t) = \mathbf{C}_a\, \underline{x}_a(t) \end{cases} \tag{5.103}$$

Let the control $\underline{u}(t)$ be chosen as follows (here pre-filter gain doesn't exists):

$$\underline{u}(t) = -\mathbf{K}_a\, \underline{x}_a(t) = -\begin{bmatrix} \mathbf{K}_p & \mathbf{K}_i \end{bmatrix} \begin{bmatrix} \underline{x}(t) \\ \underline{e}_I(t) \end{bmatrix} = -\mathbf{K}_p\, \underline{x}(t) - \mathbf{K}_i\, \underline{e}_I(t) \tag{5.104}$$

Obviously, term $\mathbf{K}_p$ represents the *proportional* gain of the controller whereas term $\mathbf{K}_i$ represents the *integral* gain of the controller.

The state space equation of closed-loop system is obtained by setting $\underline{u}(t) = -\mathbf{K}_a\, \underline{x}_a(t) = -\mathbf{K}_p\, \underline{x}(t) - \mathbf{K}_i\, \underline{e}_I(t)$ in (5.101):

$$\begin{cases} \frac{d}{dt}\begin{bmatrix} \underline{x}(t) \\ \underline{e}_I(t) \end{bmatrix} &= (\mathbf{A}_a - \mathbf{B}_a\mathbf{K}_a)\, \underline{x}_a(t) + \mathbf{E}\underline{r}(t) \\ &= \begin{bmatrix} \mathbf{A} - \mathbf{B}\mathbf{K}_p & -\mathbf{B}\mathbf{K}_i \\ -\mathbf{T}\mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \underline{x}(t) \\ \underline{e}_I(t) \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{T} \end{bmatrix} \underline{r}(t) \\ \underline{y}(t) &= \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \underline{x}(t) \\ \underline{e}_I(t) \end{bmatrix} \\ \underline{e}_I(t) &= \mathbf{T} \int_0^t \left(\underline{r}(\tau) - \underline{y}(\tau)\right)\, d\tau \end{cases} \tag{5.105}$$

The corresponding bloc diagram is shown in Figure 5.1 where $\Phi_a(s) = (s\mathbb{I} - \mathbf{A}_a)^{-1}$.

5.5.2 Proof of the cancellation of the steady state error through integral augmentation

In order to proof that integrator cancels the steady state error when $\underline{r}(t)$ is a step input, let us compute the final value of the error $\underline{e}(t)$ using the final value theorem where s denotes the Laplace variable:

$$\lim_{t\to\infty} \underline{e}(t) = \lim_{s\to 0} s\underline{E}(s) \tag{5.106}$$

When $\underline{r}(t)$ is a step input with amplitude one, we have:

$$\underline{r}(t) = 1\ \forall\ t \geq 0 \Rightarrow \underline{R}(s) = \frac{1}{s} \tag{5.107}$$

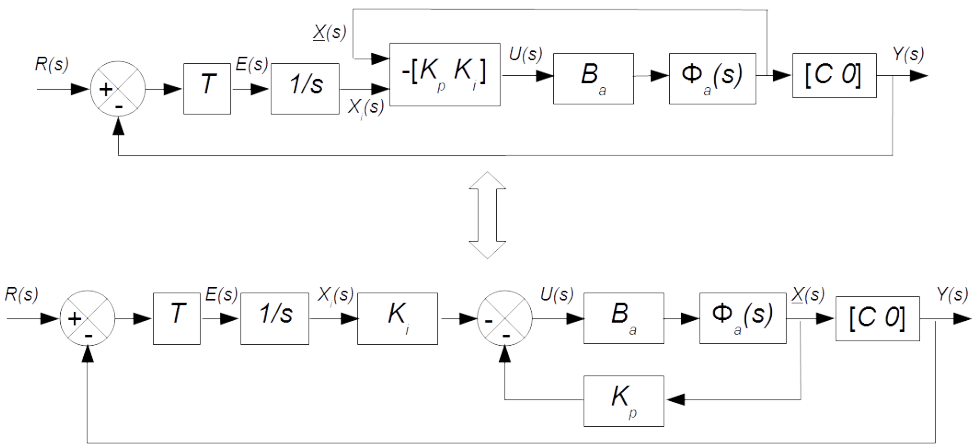

Figure 5.1: Plant augmented with integrator

Using the feedback $\underline{u} = -\mathbf{K}_a \underline{x}_a$ the dynamics of the closed-loop system is:

$$\begin{aligned} \dot{\underline{x}}_a &= (\mathbf{A}_a - \mathbf{B}_a\mathbf{K}_a)\,\underline{x}_a + \begin{bmatrix}\mathbf{0}\\ \mathbf{T}\end{bmatrix}\underline{r}(t) \\ \Rightarrow \underline{e}(t) &= \mathbf{T}\left(\underline{r}(t) - \underline{y}(t)\right) \\ &= \mathbf{T}\left(\underline{r}(t) - \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\underline{x}\\ \underline{x}_i\end{bmatrix}\right) \\ &= \mathbf{T}\left(\underline{r}(t) - \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}\underline{x}_a\right) \end{aligned} \tag{5.108}$$

Using the Laplace transform, and denoting by $\mathbb{I}$ the identity matrix, we get:

$$\begin{cases} \underline{X}_a(s) = (s\mathbb{I} - \mathbf{A}_a + \mathbf{B}_a\mathbf{K}_a)^{-1}\begin{bmatrix}\mathbf{0}\\ \mathbf{T}\end{bmatrix}\underline{R}(s) \\ \underline{E}(s) = \mathbf{T}\left(\underline{R}(s) - \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}\underline{X}_a(s)\right) \end{cases} \tag{5.109}$$

Inserting (5.107) in (5.109) we get:

$$\underline{E}(s) = \mathbf{T}\left(\mathbb{I} - \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}(s\mathbb{I} - \mathbf{A}_a + \mathbf{B}_a\mathbf{K}_a)^{-1}\begin{bmatrix}\mathbf{0}\\ \mathbf{T}\end{bmatrix}\right)\frac{1}{s} \tag{5.110}$$

Then the final value theorem (5.106) takes the following expression:

$$\begin{aligned} \lim_{t\to\infty}\underline{e}(t) &= \lim_{s\to 0} s\underline{E}(s) \\ &= \lim_{s\to 0}\mathbf{T}\left(\mathbb{I} - \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}(s\mathbb{I} - \mathbf{A}_a + \mathbf{B}_a\mathbf{K}_a)^{-1}\begin{bmatrix}\mathbf{0}\\ \mathbf{T}\end{bmatrix}\right) \\ &= \mathbf{T}\left(\mathbb{I} - \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}(-\mathbf{A}_a + \mathbf{B}_a\mathbf{K}_a)^{-1}\begin{bmatrix}\mathbf{0}\\ \mathbf{T}\end{bmatrix}\right) \end{aligned} \tag{5.111}$$

Let us focus on the inverse of the matrix $-\mathbf{A}_a + \mathbf{B}_a\mathbf{K}_a$. First we write $\mathbf{K}_a$ as $\mathbf{K}_a = \begin{bmatrix}\mathbf{K}_p & \mathbf{K}_i\end{bmatrix}$, where $\mathbf{K}_p$ and $\mathbf{K}_i$ represents respectively the proportional and the integral gains. Then using (5.101) we get:

$$-\mathbf{A}_a + \mathbf{B}_a\mathbf{K}_a = \begin{bmatrix}-\mathbf{A} & \mathbf{0}\\ \mathbf{TC} & \mathbf{0}\end{bmatrix} + \begin{bmatrix}\mathbf{B}\\ \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{K}_p & \mathbf{K}_i\end{bmatrix} = \begin{bmatrix}-\mathbf{A} + \mathbf{BK}_p & \mathbf{BK}_i\\ \mathbf{TC} & \mathbf{0}\end{bmatrix} \tag{5.112}$$

Assuming that $\mathbf{X}$ is a square invertible matrix, it can be shown that the inverse of the matrix $\begin{bmatrix}\mathbf{X} & \mathbf{Y}\\ \mathbf{Z} & \mathbf{0}\end{bmatrix}$ is the following:

$$\begin{bmatrix}\mathbf{X} & \mathbf{Y}\\ \mathbf{Z} & \mathbf{0}\end{bmatrix}^{-1} = \begin{bmatrix}\mathbf{0} & \mathbf{Y}\left(\mathbf{ZY}\right)^{-1}\\ \mathbf{Y}^T\left(\mathbf{YY}^T\right)^{-1} & \mathbf{W}\end{bmatrix}$$
$$\text{where } \mathbf{XY}\left(\mathbf{ZY}\right)^{-1} + \mathbf{YW} = \mathbf{0} \quad (5.113)$$

Thus:

$$\begin{aligned}\left(-\mathbf{A}_a + \mathbf{B}_a\mathbf{K}_a\right)^{-1} &= \begin{bmatrix}-\mathbf{A}+\mathbf{BK}_p & \mathbf{BK}_i\\ \mathbf{TC} & \mathbf{0}\end{bmatrix}^{-1}\\ &= \begin{bmatrix}\mathbf{0} & \mathbf{BK}_i\left(\mathbf{TCBK}_i\right)^{-1}\\ \left(\mathbf{BK}_i\right)^T\left(\mathbf{BK}_i\left(\mathbf{BK}_i\right)^T\right)^{-1} & \mathbf{W}\end{aligned} \quad (5.114)$$

And:

$$\begin{aligned}\left(-\mathbf{A}_a + \mathbf{B}_a\mathbf{K}\right)^{-1}\begin{bmatrix}\mathbf{0}\\ \mathbf{T}\end{bmatrix} &= \begin{bmatrix}\mathbf{0} & \mathbf{BK}_i\left(\mathbf{TCBK}_i\right)^{-1}\\ * & \mathbf{W}\end{bmatrix}\begin{bmatrix}\mathbf{0}\\ \mathbf{T}\end{bmatrix}\\ &= \begin{bmatrix}\mathbf{BK}_i\left(\mathbf{TCBK}_i\right)^{-1}\mathbf{T}\\ \mathbf{WT}\end{bmatrix}\\ \Rightarrow \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}\left(-\mathbf{A}_a + \mathbf{B}_a\mathbf{K}\right)^{-1}\begin{bmatrix}\mathbf{0}\\ \mathbf{T}\end{bmatrix} &= \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{BK}_i\left(\mathbf{TCBK}_i\right)^{-1}\mathbf{T}\\ \mathbf{WT}\end{bmatrix}\\ &= \mathbf{CBK}_i\left(\mathbf{TCBK}_i\right)^{-1}\mathbf{T}\end{aligned} \quad (5.115)$$

Consequently, using (5.115) in (5.111), the final value of the error $\underline{e}(t)$ becomes:

$$\begin{aligned}\lim_{t\to\infty}\underline{e}(t) &= \mathbf{T}\left(\mathbb{I} - \begin{bmatrix}\mathbf{C} & \mathbf{0}\end{bmatrix}\left(-\mathbf{A}_a + \mathbf{B}_a\mathbf{K}_a\right)^{-1}\begin{bmatrix}\mathbf{0}\\ \mathbf{T}\end{bmatrix}\right)\\ &= \mathbf{T}\left(\mathbb{I} - \mathbf{CBK}_i\left(\mathbf{TCBK}_i\right)^{-1}\mathbf{T}\right)\\ &= \mathbf{T} - \mathbf{TCBK}_i\left(\mathbf{TCBK}_i\right)^{-1}\mathbf{T}\\ &= \mathbf{T} - \mathbf{T}\\ &= \mathbf{0}\end{aligned} \quad (5.116)$$

As a consequence, the integrator allows to cancel the steady state error whatever the input step $\underline{r}(t)$.

5.6 Observer-based controller

5.6.1 Separation principle

We consider the following state-space representation of dimension n (that is the size of state matrix $\mathbf{A}$) where $\underline{y}(t)$ denotes the output vector and $\underline{u}(t)$ the input vector:

$$\begin{cases}\dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t)\\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t)\end{cases} \quad (5.117)$$

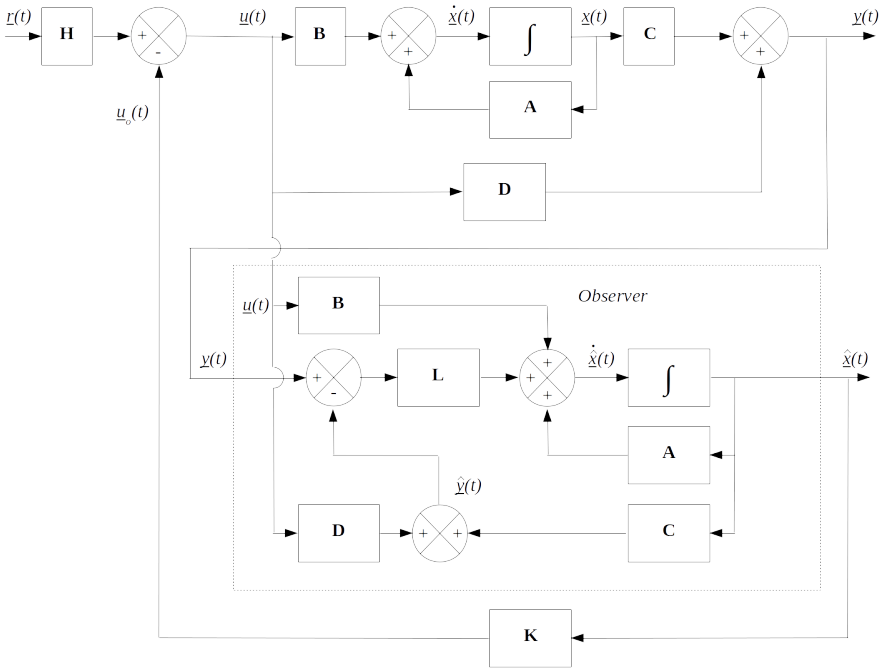

Figure 5.2: Block diagram of the observer-based controller

When the full state $\underline{x}(t)$ cannot be measured then it is estimated thanks to an observer. This leads to an observer-based controller with inputs $\underline{y}(t)$ and $\underline{u}(t)$. The output of the observer-based controller will be denoted $\underline{u}_o(t)$. Its dynamics reads as follows:

$$\begin{cases} \dot{\hat{\underline{x}}}(t) = \mathbf{A}\hat{\underline{x}}(t) + \mathbf{B}\underline{u}(t) + \mathbf{L}\left(\underline{y}(t) - (\mathbf{C}\hat{\underline{x}}(t) + \mathbf{D}\underline{u}(t))\right) \\ \underline{u}_o(t) = \mathbf{K}\hat{\underline{x}}(t) \end{cases} \tag{5.118}$$

Gain matrices $\mathbf{L}$, $\mathbf{K}$ and $\mathbf{H}$ are degrees of freedom which shall be set to achieve some performance criteria.

The block diagram corresponding to the observer-based controller is shown in Figure 5.2.

The estimation error $\underline{e}(t)$ is defined as follows:

$$\underline{e}(t) = \underline{x}(t) - \hat{\underline{x}}(t) \tag{5.119}$$

The time derivative of the estimation error reads:

$$\begin{aligned} \dot{\underline{e}}(t) &= \dot{\underline{x}}(t) - \dot{\hat{\underline{x}}}(t) \\ &= \mathbf{A}\underline{x}(t) - \mathbf{A}\hat{\underline{x}}(t) - \mathbf{L}\left(\underline{y}(t) - \mathbf{C}\hat{\underline{x}}(t) - \mathbf{D}\underline{u}(t)\right) \\ &= \mathbf{A}\underline{x}(t) - \mathbf{A}\hat{\underline{x}}(t) - \mathbf{L}\left(\mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) - \mathbf{C}\hat{\underline{x}}(t) - \mathbf{D}\underline{u}(t)\right) \\ &= \mathbf{A}\underline{e}(t) - \mathbf{L}\mathbf{C}\underline{e}(t) \\ &= (\mathbf{A} - \mathbf{L}\mathbf{C})\,\underline{e}(t) \end{aligned} \tag{5.120}$$

Combining the dynamics of the state vector $\underline{x}(t)$ in (5.117) and dynamics of the estimation error $\underline{e}(t)$, and using the fact that $\hat{\underline{x}}(t) = \underline{x}(t) - \underline{e}(t)$, yields to

the following state-space representation for the closed-loop system:

$$\begin{bmatrix} \dot{\underline{x}}(t) \\ \dot{\underline{e}}(t) \end{bmatrix} = \mathbf{A}_{cl} \begin{bmatrix} \underline{x}(t) \\ \underline{e}(t) \end{bmatrix} + \begin{bmatrix} \mathbf{BH} \\ \mathbf{0} \end{bmatrix} \underline{r}(t) \tag{5.121}$$

where:

$$\mathbf{A}_{cl} = \begin{bmatrix} \mathbf{A} - \mathbf{BK} & \mathbf{BK} \\ \mathbf{0} & \mathbf{A} - \mathbf{LC} \end{bmatrix} \tag{5.122}$$

Gain matrices $\mathbf{L}$ and $\mathbf{K}$ shall be chosen such that the eigenvalues of matrices $\mathbf{A} - \mathbf{BK}$ and $\mathbf{A} - \mathbf{LC}$ are situated in the left half complex plane so that the closed-loop system is asymptotically stable.

Furthermore it is worth noticing that matrix is block triangular. Consequently we can write:

$$\begin{aligned} \det\left(s\mathbb{I} - \mathbf{A}_{cl}\right) &= \det\left(\begin{bmatrix} s\mathbb{I} - \mathbf{A} + \mathbf{BK} & -\mathbf{BK} \\ \mathbf{0} & s\mathbb{I} - \mathbf{A} + \mathbf{LC} \end{bmatrix}\right) \\ &= \det\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right) \det\left(s\mathbb{I} - \mathbf{A} + \mathbf{LC}\right) \end{aligned} \tag{5.123}$$

In other words the eigenvalues of the closed-loop system are obtained by the union of the eigenvalues of matrix $\mathbf{A} - \mathbf{BK}$, that is the state matrix of the closed-loop system without the observer, and the eigenvalues of matrix $\mathbf{A} - \mathbf{LC}$, that is the state matrix of the closed-loop system without the controller. This result is known as the separation principle. As a consequence the observer and the controller can be designed separately: the eigenvalues obtained thanks to the controller gain $\mathbf{K}$ assuming full state feedback are independent of the eigenvalues obtained thanks to the observer gain $\mathbf{L}$ assuming no controller.

Gain matrices $\mathbf{L}$ and $\mathbf{K}$ shall be chosen such that the eigenvalues of matrices $\mathbf{A} - \mathbf{BK}$ and $\mathbf{A} - \mathbf{LC}$ are situated in the left half complex plane so that the closed-loop system is asymptotically stable.

Usually observer gain $\mathbf{L}$ is chosen such that the eigenvalues of matrix $\mathbf{A} - \mathbf{LC}$ are around 5 to 10 times faster than the eigenvalues of matrix $\mathbf{A} - \mathbf{BK}$, so that the state estimation moves towards the actual state as early as possible.

Furthermore the reference input $\underline{r}(t)$ has no influence on the dynamics of the estimation error $\underline{e}(t)$. Consequently the feedforward gain matrix $\mathbf{H}$ is still given by Equation (5.7).

5.6.2 Example

Design an *output* feedback controller for the following unstable plant:

$$\begin{cases} \dot{\underline{x}}(t) = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u(t) \\ \\ y(t) = \begin{bmatrix} 3 & 5 \end{bmatrix} \underline{x}(t) \end{cases} \tag{5.124}$$

The poles of the controller shall be chosen to render the closed-loop stable and to satisfy some specifications. We choose (for example) to locate the poles of the controller at $\lambda_{K_1} = -1$ and $\lambda_{K_2} = -2$.

First we check that is system is observable and controllable.

We have seen in example 5.1 how to design a *state* feedback controller. By applying the separation principle the observer which estimates the state vector $\underline{\hat{x}}(t)$ which will feed the controller can be designed separately from the controller. We have obtained:

$$\begin{cases} \mathbf{K} = \begin{bmatrix} -6 & 6 \end{bmatrix} \\ \mathbf{H} = -0.125 \end{cases} \tag{5.125}$$

As mentioned the eigenvalues of the observer are chosen around 5 to 10 times faster than the eigenvalues achieved thanks to the controller. As far as the closed-loop poles obtained thanks to the controller are located at $\lambda_{K_1} = -1$ and $\lambda_{K_2} = -2$ we choose (for example) to locate the poles of the observer at $\lambda_{L1} = -10$ and $\lambda_{L2} = -20$.

As described in the chapter dedicated to *Observer design*, we first design the observer assuming that we have the observable canonical form of the SISO system. The observable canonical form is readily obtained through $\det(s\mathbb{I} - \mathbf{A})$:

$$\begin{aligned} &\det(s\mathbb{I} - \mathbf{A}) = (s-1)(s-2) = s^2 - 3\,s + 2 := s^2 + a_1\,s + a_0 \\ &\Rightarrow \mathbf{A}_o = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix} \text{ and } \mathbf{C}_o = \begin{bmatrix} 0 & 1 \end{bmatrix} \end{aligned} \tag{5.126}$$

On the other hand the characteristic polynomial of the observer is formed thanks to the predefined eigenvalues assigned for the dynamics of the observer:

$$\begin{aligned} \chi_{A-LC}(s) &= (s - \lambda_{L1})\,(s - \lambda_{L2}) = (s+10)\,(s+20) \\ &= s^2 + 30\,s + 200 := s^2 + p_1\,s + p_0 \end{aligned} \tag{5.127}$$

Applying relation (4.39) we get:

$$\mathbf{L}_o = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} = \begin{bmatrix} p_0 - a_0 \\ p_1 - a_1 \end{bmatrix} = \begin{bmatrix} 200 - 2 \\ 30 + 3 \end{bmatrix} = \begin{bmatrix} 198 \\ 33 \end{bmatrix} \tag{5.128}$$

Now let's compute the similarity transformation matrix $\mathbf{P}_o$ which enables to get the observable canonical form.

$$\mathbf{P}_o = \mathbf{Q}_o^{-1}\mathbf{Q}_{oo} \tag{5.129}$$

Where:

- $\mathbf{Q}_o$ is the observability matrix in the actual basis:

$$\mathbf{Q}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \end{bmatrix} = \left[\begin{array}{cc} 3 & 5 \\ \hline \begin{bmatrix} 3 & 5 \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \end{array}\right] = \begin{bmatrix} 3 & 5 \\ 3 & 10 \end{bmatrix} \tag{5.130}$$

- and $\mathbf{Q}_{oo}$ the observability matrix expressed in the observable canonical basis (which is readily obtained through $\det(s\mathbb{I} - \mathbf{A})$):

$$\mathbf{Q}_{oo} = \begin{bmatrix} \mathbf{C}_o \\ \mathbf{C}_o\mathbf{A}_o \end{bmatrix} = \left[\begin{array}{cc} 0 & 1 \\ \hline \begin{bmatrix} 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix} \end{array}\right] = \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \tag{5.131}$$

Thus:

$$\begin{aligned}\mathbf{P}_o &= \mathbf{Q}_o^{-1}\mathbf{Q}_{oo} = \begin{bmatrix} 3 & 5 \\ 3 & 10 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \\ &= \tfrac{1}{15}\begin{bmatrix} 10 & -5 \\ -3 & 3 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \\ &= \tfrac{1}{15}\begin{bmatrix} -5 & -5 \\ 3 & 6 \end{bmatrix}\end{aligned} \tag{5.132}$$

We finally get:

$$\mathbf{L} = \mathbf{P}_o\mathbf{L}_o = \frac{1}{15}\begin{bmatrix} -5 & -5 \\ 3 & 6 \end{bmatrix}\begin{bmatrix} 198 \\ 33 \end{bmatrix} = \begin{bmatrix} -77 \\ 52.8 \end{bmatrix} \tag{5.133}$$

5.6.3 Feedback controller

It is possible to avoid the need to feed the observer (5.118) with control $\underline{u}(t)$. In the alternative design presented hereafter, the output-based controller is feed by the output $\underline{y}(t)$ of the plant to be controlled. The counterpart of this design is that the estimation $\underline{\hat{x}}(t)$ of the state vector of the plant is now not provided by the output-based controller. The output of the output-based controller will be denoted $\underline{u}_o(t)$. Its dynamics reads as follows:

$$\begin{cases} \dot{\underline{x}}_c(t) = (\mathbf{A} - \mathbf{BK})\,\underline{x}_c(t) + \mathbf{L}\left(\underline{y}(t) - (\mathbf{C} - \mathbf{DK})\,\underline{x}_c(t)\right) \\ \underline{u}_o(t) = \mathbf{K}\underline{x}_c(t) \end{cases} \tag{5.134}$$

The block diagram corresponding to the observer-based controller is shown in Figure 5.3.

By combining the dynamics of the plant (5.117) and the dynamics of the output-based controller (5.134), we get the dynamics of the closed-loop system, which reads as follows:

$$\begin{cases} \underline{u}(t) = \mathbf{H}\underline{r}(t) - \underline{u}_o(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \end{cases} \Rightarrow \begin{bmatrix} \dot{\underline{x}}(t) \\ \dot{\underline{x}}_c(t) \end{bmatrix} = \mathbf{A}_{cl}\begin{bmatrix} \underline{x}(t) \\ \underline{x}_c(t) \end{bmatrix} + \begin{bmatrix} \mathbf{BH} \\ \mathbf{LD} \end{bmatrix}\underline{r}(t) \tag{5.135}$$

where the closed-loop state matrix $\mathbf{A}_{cl}$ reads:

$$\mathbf{A}_{cl} = \begin{bmatrix} \mathbf{A} & -\mathbf{BK} \\ \mathbf{LC} & \mathbf{A} - \mathbf{BK} - \mathbf{LC} \end{bmatrix} \tag{5.136}$$

The closed-loop eigenvalues are the roots of the characteristic polynomial $\chi_{A_{cl}}(s)$ defined as follows:

$$\chi_{A_{cl}}(s) = \det\left(s\mathbb{I} - \mathbf{A}_{cl}\right) = \det\left(\begin{bmatrix} s\mathbb{I} - \mathbf{A} & \mathbf{BK} \\ -\mathbf{LC} & s\mathbb{I} - \mathbf{A} + \mathbf{BK} + \mathbf{LC} \end{bmatrix}\right) \tag{5.137}$$

Now we will use the fact that adding one column / row to another column / row does not change the value of the determinant. Thus subtracting the second row to the first row of $\begin{bmatrix} s\mathbb{I} - \mathbf{A} & \mathbf{BK} \\ -\mathbf{LC} & s\mathbb{I} - \mathbf{A} + \mathbf{BK} + \mathbf{LC} \end{bmatrix}$ leads to the following expression of $\chi_{A_{cl}}(s)$:

$$\chi_{A_{cl}}(s) = \det\left(\begin{bmatrix} s\mathbb{I} - \mathbf{A} + \mathbf{LC} & -s\mathbb{I} + \mathbf{A} - \mathbf{LC} \\ -\mathbf{LC} & s\mathbb{I} - \mathbf{A} + \mathbf{BK} + \mathbf{LC} \end{bmatrix}\right) \tag{5.138}$$

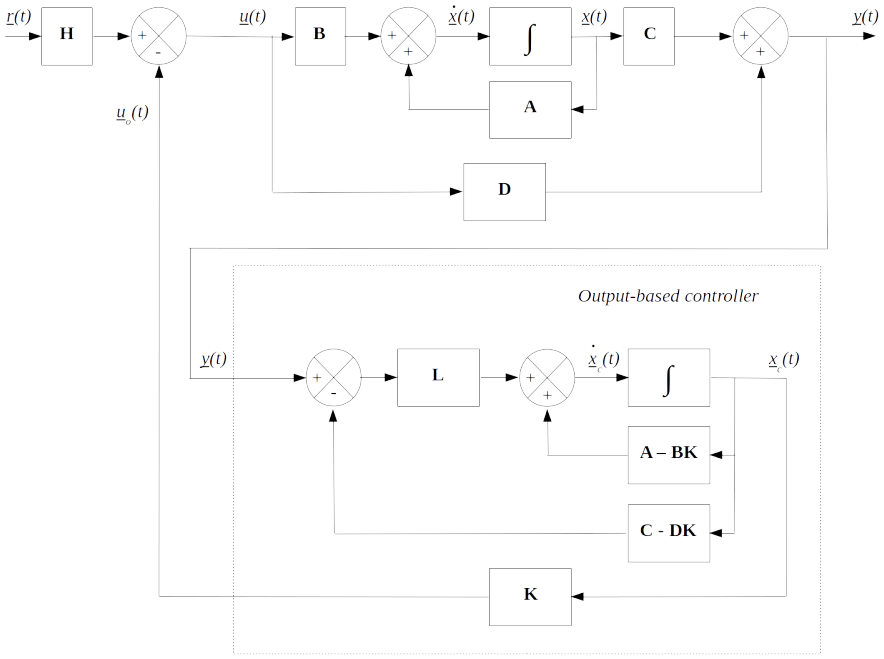

Figure 5.3: Block diagram of the output-based controller

Now adding the first column to the second column of $\begin{bmatrix} s\mathbb{I} - \mathbf{A} + \mathbf{LC} & -s\mathbb{I} + \mathbf{A} - \mathbf{LC} \\ -\mathbf{LC} & s\mathbb{I} - \mathbf{A} + \mathbf{BK} + \mathbf{LC} \end{bmatrix}$ leads to the following expression of $\chi_{A_{cl}}(s)$:

$$\chi_{A_{cl}}(s) = \det\left(\begin{bmatrix} s\mathbb{I} - \mathbf{A} + \mathbf{LC} & \mathbf{0} \\ -\mathbf{LC} & s\mathbb{I} - \mathbf{A} + \mathbf{BK} \end{bmatrix}\right) \tag{5.139}$$

Furthermore it is worth noticing that matrix is block triangular. Consequently we can write:

$$\begin{aligned} \det\left(s\mathbb{I} - \mathbf{A}_{cl}\right) &= \det\left(\begin{bmatrix} s\mathbb{I} - \mathbf{A} + \mathbf{BK} & -\mathbf{BK} \\ \mathbf{0} & s\mathbb{I} - \mathbf{A} + \mathbf{LC} \end{bmatrix}\right) \\ &= \det\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)\det\left(s\mathbb{I} - \mathbf{A} + \mathbf{LC}\right) \end{aligned} \tag{5.140}$$

In other words the eigenvalues of the closed-loop system are obtained by the union of the eigenvalues of matrix $\mathbf{A} - \mathbf{BK}$, that is the state matrix of the closed-loop system assuming $\mathbf{L} = \mathbf{0}$, and the eigenvalues of matrix $\mathbf{A} - \mathbf{LC}$, that is the state matrix of the closed-loop system assuming $\mathbf{K} = \mathbf{0}$. This result is known (again) as the separation principle. As a consequence, gains $\mathbf{K}$ and $\mathbf{L}$ can be designed separately: the eigenvalues obtained thanks to the gain $\mathbf{K}$ assuming $\mathbf{L} = \mathbf{0}$ are independent of the eigenvalues obtained thanks to the gain $\mathbf{L}$ assuming assuming $\mathbf{K} = \mathbf{0}$.

Furthermore the reference input $\underline{r}(t)$ has no influence on the dynamics of the closed-loop. Consequently matrix $\mathbf{H}$ is still given by Equation (5.7).

Consequently, the block diagram in Figure 5.3 is equivalent to the block diagram in Figure 5.4 where the controller $C(s)$ is put in the feedback loop.

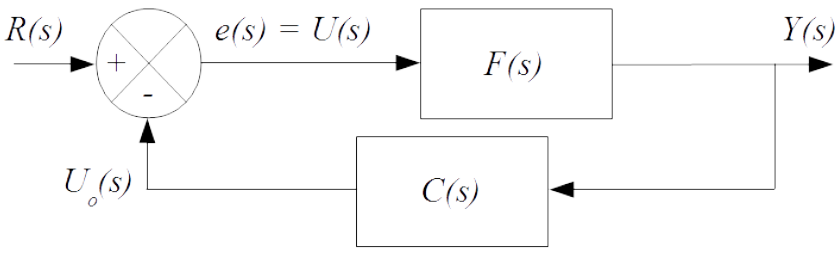

Figure 5.4: Feedback loop with $C(s)$

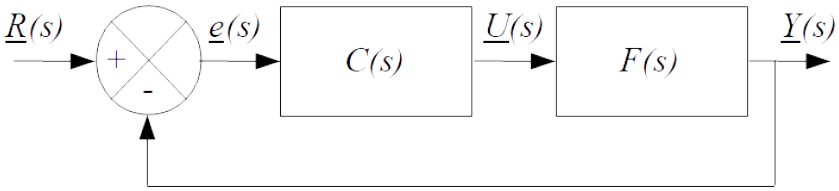

Figure 5.5: Unity feedback loop

Assuming that $\mathbf{D} = \mathbf{0}$, the transfer function of the output-based controller can be obtained by taking the Laplace transform (assuming no initial conditions) of its state space representation:

$$\begin{cases} \dot{\underline{x}}_c(t) = (\mathbf{A} - \mathbf{BK})\,\underline{x}_c(t) + \mathbf{L}\left(\underline{y}(t) - \mathbf{C}\underline{x}_c(t)\right) \\ \underline{u}_o(t) = \mathbf{K}\underline{x}_c(t) \end{cases} \tag{5.141}$$

We get:

$$\underline{U}_o(s) = \mathbf{C}(s)\underline{Y}(s) \tag{5.142}$$

Where:

$$\mathbf{C}(s) = \mathbf{K}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK} + \mathbf{LC}\right)^{-1}\mathbf{L} \tag{5.143}$$

5.6.4 Unity feedback loop

We consider now the unity feedback loop shown in Figure 5.5. The realization of the plant transfer function $F(s)$ is assumed to be the following:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.144}$$

On the other hand, the realization of the controller transfer function $C(s)$ is assumed to be the following, where gain matrices $\mathbf{K}$ and $\mathbf{L}$ are the design parameters of the controller :

$$\boxed{\begin{cases} \dot{\underline{x}}_c(t) = (\mathbf{A} - \mathbf{BK} - \mathbf{LC})\,\underline{x}_c(t) + \mathbf{L}\underline{e}(t) \\ \underline{u}(t) = \mathbf{K}\underline{x}_c(t) \end{cases}} \tag{5.145}$$

To get the transfer function of the controller we take the Laplace transform (assuming no initial conditions) of its state space representation (5.145). We

get:

$$\boxed{\underline{U}(s) = \mathbf{C}(s)\underline{e}(s)} \tag{5.146}$$

Where:

$$\boxed{\mathbf{C}(s) = \mathbf{K}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK} + \mathbf{LC}\right)^{-1}\mathbf{L}} \tag{5.147}$$

In order to get the closed-loop realization, we get the following relations from Figure 5.5 :

$$\begin{cases} \underline{e}(t) = \underline{r}(t) - \underline{y}(t) = \underline{r}(t) - \mathbf{C}\underline{x}(t) \\ \underline{u}(t) = \mathbf{K}\underline{x}_c(t) \end{cases} \tag{5.148}$$

Thus the state space realization of the unity feedback loop reads:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{BK}\underline{x}_c(t) \\ \dot{\underline{x}}_c(t) = \left(\mathbf{A} - \mathbf{BK} - \mathbf{LC}\right)\underline{x}_c(t) + \mathbf{L}\left(\underline{r}(t) - \mathbf{C}\underline{x}(t)\right) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.149}$$

That is:

$$\begin{cases} \begin{bmatrix} \dot{\underline{x}}(t) \\ \dot{\underline{x}}_c(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{BK} \\ -\mathbf{LC} & \mathbf{A} - \mathbf{BK} - \mathbf{LC} \end{bmatrix} \begin{bmatrix} \underline{x}(t) \\ \underline{x}_c(t) \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{L} \end{bmatrix} \underline{r}(t) \\ \underline{y}(t) = \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \underline{x}(t) \\ \underline{x}_c(t) \end{bmatrix} \end{cases} \tag{5.150}$$

It is worth noticing that the following relations hold:

$$\boxed{\begin{array}{c} \mathbf{P}_1^{-1} \begin{bmatrix} \mathbf{A} & \mathbf{BK} \\ -\mathbf{LC} & \mathbf{A} - \mathbf{BK} - \mathbf{LC} \end{bmatrix} \mathbf{P}_1 = \begin{bmatrix} \mathbf{A} - \mathbf{BK} & \mathbf{BK} \\ \mathbf{0} & \mathbf{A} - \mathbf{LC} \end{bmatrix} \\ \text{where} \begin{cases} \mathbf{P}_1^{-1} = \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ \mathbb{I} & \mathbb{I} \end{bmatrix} \\ \mathbf{P}_1 = \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbb{I} & \mathbb{I} \end{bmatrix} \end{cases} \end{array}} \tag{5.151}$$

Let $\underline{\xi}_1$ be another state vector defined as follows:

$$\begin{bmatrix} \underline{x}(t) \\ \underline{x}_c(t) \end{bmatrix} = \mathbf{P}_1\underline{\xi}_1 \Leftrightarrow \underline{\xi}_1 = \mathbf{P}_1^{-1} \begin{bmatrix} \underline{x}(t) \\ \underline{x}_c(t) \end{bmatrix} \tag{5.152}$$

Then state space representation (5.150) becomes:

$$\begin{cases} \dot{\underline{\xi}}_1(t) & = \mathbf{P}_1^{-1} \begin{bmatrix} \mathbf{A} & \mathbf{BK} \\ -\mathbf{LC} & \mathbf{A} - \mathbf{BK} - \mathbf{LC} \end{bmatrix} \mathbf{P}_1\underline{\xi}_1(t) + \mathbf{P}_1^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{L} \end{bmatrix} \underline{r}(t) \\ & = \begin{bmatrix} \mathbf{A} - \mathbf{BK} & \mathbf{BK} \\ \mathbf{0} & \mathbf{A} - \mathbf{LC} \end{bmatrix} \underline{\xi}_1(t) + \begin{bmatrix} \mathbf{0} \\ \mathbf{L} \end{bmatrix} \underline{r}(t) \\ \underline{y}(t) & = \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix} \mathbf{P}_1\underline{\xi}_1(t) = \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix} \underline{\xi}_1(t) \end{cases} \tag{5.153}$$

Using the fact that the state space matrix is now block triangular, the characteristic polynomial $\chi_{A_{cl}}(s)$ of the unity feedback loop reads as follows:

$$\begin{aligned}\chi_{A_{cl}}(s) &= \det\left(s\mathbb{I} - \begin{bmatrix} \mathbf{A}-\mathbf{BK} & \mathbf{BK} \\ \mathbf{0} & \mathbf{A}-\mathbf{LC} \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} s\mathbb{I}-\mathbf{A}+\mathbf{BK} & -\mathbf{BK} \\ \mathbf{0} & s\mathbb{I}-\mathbf{A}+\mathbf{LC} \end{bmatrix}\right) \\ &= \det\left(s\mathbb{I}-\mathbf{A}+\mathbf{LC}\right)\det\left(s\mathbb{I}-\mathbf{A}+\mathbf{BK}\right)\end{aligned} \tag{5.154}$$

In other words the eigenvalues of the closed-loop system are obtained by the union of the eigenvalues of matrix $\mathbf{A}-\mathbf{BK}$, that is the state matrix of the closed-loop system assuming $\mathbf{L}=\mathbf{0}$, and the eigenvalues of matrix $\mathbf{A}-\mathbf{LC}$, that is the state matrix of the closed-loop system assuming $\mathbf{K}=\mathbf{0}$. This result is known (again) as the separation principle. As a consequence, gains $\mathbf{K}$ and $\mathbf{L}$ can be designed separately: the eigenvalues obtained thanks to the gain $\mathbf{K}$ assuming $\mathbf{L}=\mathbf{0}$ are independent of the eigenvalues obtained thanks to the gain $\mathbf{L}$ assuming assuming $\mathbf{K}=\mathbf{0}$.

As far as the closed-loop transfer function $\mathbf{G}(s)$ is concerned, we get from (5.153):

$$\begin{aligned}\mathbf{G}(s) &= \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix}\left(s\mathbb{I} - \begin{bmatrix} \mathbf{A}-\mathbf{BK} & \mathbf{BK} \\ \mathbf{0} & \mathbf{A}-\mathbf{LC} \end{bmatrix}\right)^{-1}\begin{bmatrix} \mathbf{0} \\ \mathbf{L} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix}\begin{bmatrix} s\mathbb{I}-\mathbf{A}+\mathbf{BK} & -\mathbf{BK} \\ \mathbf{0} & s\mathbb{I}-\mathbf{A}+\mathbf{LC} \end{bmatrix}^{-1}\begin{bmatrix} \mathbf{0} \\ \mathbf{L} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix}\begin{bmatrix} (s\mathbb{I}-\mathbf{A}+\mathbf{BK})^{-1} & \mathbf{X}(s) \\ \mathbf{0} & (s\mathbb{I}-\mathbf{A}+\mathbf{LC})^{-1} \end{bmatrix}\begin{bmatrix} \mathbf{0} \\ \mathbf{L} \end{bmatrix}\end{aligned} \tag{5.155}$$

where:

$$\mathbf{X}(s) = (s\mathbb{I}-\mathbf{A}+\mathbf{BK})^{-1}\,\mathbf{BK}\,(s\mathbb{I}-\mathbf{A}+\mathbf{LC})^{-1} \tag{5.156}$$

We finally get:

$$\begin{aligned}\mathbf{G}(s) &= \mathbf{C}\,\mathbf{X}(s)\,\mathbf{L} \\ &= \Big(\mathbf{C}\,(s\mathbb{I}-\mathbf{A}+\mathbf{BK})^{-1}\,\mathbf{B}\Big)\Big(\mathbf{K}\,(s\mathbb{I}-\mathbf{A}+\mathbf{LC})^{-1}\,\mathbf{L}\Big)\end{aligned} \tag{5.157}$$

The preceding relation indicates that the closed-loop transfer function $\mathbf{G}(s)$ can be seen as the series interconnection between transfer functions $\mathbf{G}_1(s)$ and $\mathbf{G}_2(s)$:

$$\mathbf{G}(s) = \mathbf{G}_1(s)\,\mathbf{G}_2(s) \text{ where } \begin{cases} \mathbf{G}_1(s) = \mathbf{C}\,(s\mathbb{I}-\mathbf{A}+\mathbf{BK})^{-1}\,\mathbf{B} \\ \mathbf{G}_2(s) = \mathbf{K}\,(s\mathbb{I}-\mathbf{A}+\mathbf{LC})^{-1}\,\mathbf{L} \end{cases} \tag{5.158}$$

Similar results can be obtained through the following alternative relation:

$$\boxed{\begin{gathered}\mathbf{P}_2^{-1}\begin{bmatrix} \mathbf{A} & \mathbf{BK} \\ -\mathbf{LC} & \mathbf{A}-\mathbf{BK}-\mathbf{LC} \end{bmatrix}\mathbf{P}_2 = \begin{bmatrix} \mathbf{A}-\mathbf{LC} & \mathbf{0} \\ -\mathbf{LC} & \mathbf{A}-\mathbf{BK} \end{bmatrix} \\ \text{where } \begin{cases} \mathbf{P}_2^{-1} = \begin{bmatrix} \mathbb{I} & \mathbb{I} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \\ \mathbf{P}_2 = \begin{bmatrix} \mathbb{I} & -\mathbb{I} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \end{cases}\end{gathered}} \tag{5.159}$$

5.7 Static output feedback controller

5.7.1 Partial eigenvalues assignment

We consider in this section the following linear dynamical system where $\underline{u}(t) \in \mathbb{R}^m$, $\underline{x}(t) \in \mathbb{R}^n$ and $\underline{y}(t) \in \mathbb{R}^p$. Furthermore we assume that $(\mathbf{A}, \mathbf{B})$ is controllable and that $(\mathbf{A}, \mathbf{C})$ is observable:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.160}$$

Let's consider a static output feedback where the control $\underline{u}(t)$ is proportional to output $\underline{y}(t)$ through gain $\mathbf{K}$ as well as reference input $\underline{r}(t)$:

$$\underline{u}(t) = -\mathbf{K}\underline{y}(t) + \mathbf{H}\underline{r}(t) \tag{5.161}$$

Let $\mathbf{\Lambda}$ be the following diagonal matrix:

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{n-p} \end{bmatrix} \text{ where } \begin{cases} \mathbf{\Lambda}_p = diag\left(\lambda_{K_1}, \cdots, \lambda_{K_p}\right) \\ \mathbf{\Lambda}_{n-p} = diag\left(\lambda_{K_{p+1}}, \cdots, \lambda_{K_n}\right) \end{cases} \tag{5.162}$$

It is assumed that $\mathbf{\Lambda}_p$ and $\mathbf{\Lambda}_{n-p}$ are self-conjugate sets and that $\mathbf{\Lambda}$ contains distinct eigenvalues. The problem considered is to find a real matrix $\mathbf{K}$ such that the eigenvalues of $\mathbf{A} - \mathbf{BKC}$ are those of the set $\mathbf{\Lambda}$.

Brasch & Pearson[11] have shown that the transfer function of the closed-loop plant can be written as follows:

$$\begin{aligned} \mathbf{G}(s) &= \mathbf{C}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BKC}\right)^{-1}\mathbf{BH} \\ &= \left(\mathbb{I} + \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{BK}\right)^{-1}\mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{BH} \\ &= \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B}\left(\mathbb{I} + \mathbf{KC}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B}\right)^{-1}\mathbf{H} \end{aligned} \tag{5.163}$$

Then, given any set $\mathbf{\Lambda}_p$ there exists a static output feedback gain $\mathbf{K}$ such that the eigenvalues of $\mathbf{A} - \mathbf{BKC}$ are precisely the elements of the set $\mathbf{\Lambda}_p$. Furthermore, in view of (5.163), the same methodology than in section 5.12.2 can be applied to compute $\mathbf{K}$.

Let $\mathbf{N}_{ol}(s) := \operatorname{adj}\left(s\mathbb{I} - \mathbf{A}\right)\mathbf{B} \in \mathbb{R}^{n\times m}$, where $\operatorname{adj}\left(s\mathbb{I} - \mathbf{A}\right)$ stands for the adjugate matrix of $s\mathbb{I} - \mathbf{A}$, and $D(s) := \det\left(s\mathbb{I} - \mathbf{A}\right)$ is the determinant of $s\mathbb{I} - \mathbf{A}$, that is the characteristic polynomial of the open-loop plant :

$$\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} = \frac{\operatorname{adj}\left(s\mathbb{I} - \mathbf{A}\right)\mathbf{B}}{\det\left(s\mathbb{I} - \mathbf{A}\right)} := \frac{\mathbf{N}_{ol}(s)}{D(s)} \tag{5.164}$$

Consequently, we get from (5.163) the expression of the characteristic polynomial of the closed-loop transfer function $\mathbf{G}(s)$:

$$\det\left(s\mathbb{I}_n - \mathbf{A} + \mathbf{BKC}\right) = \det\left(D(s)\mathbb{I}_m + \mathbf{KCN}_{ol}(s)\right) \tag{5.165}$$

As soon as λ_{K_i} is a desired closed-loop eigenvalue then the following relation holds:

$$\det\left(D(s)\mathbb{I}_m + \mathbf{KCN}_{ol}(s)\right)\big|_{s=\lambda_{K_i}} = 0 \tag{5.166}$$

Consequently it is desired that matrix $D(s)\mathbb{I} + \mathbf{KCN}_{ol}(s)|_{s=\lambda_{K_i}}$ is singular. Following Shieh & al.[7], let $\underline{\omega}_i \neq \underline{0}$ be a vector of size $m \times 1$, where m is the number of columns of $\mathbf{B}$ (that is the size of the input vector $\underline{u}(t)$ of the plant), belonging to the kernel of matrix $D(s)\mathbb{I} + \mathbf{KCN}_{ol}(s)|_{s=\lambda_{K_i}}$. Thus changing s by λ_{K_i} we can write:

$$(D(\lambda_{K_i})\mathbb{I}_m + \mathbf{KCN}_{ol}(\lambda_{K_i}))\,\underline{\omega}_i = \underline{0} \tag{5.167}$$

Actually, vector $\underline{\omega}_i \neq \underline{0} \in \mathbb{C}^m$ is used as a design parameter.

In order to get gain $\mathbf{K}$ the preceding relation is rewritten as follows:

$$\mathbf{KCN}_{ol}(\lambda_{K_i})\underline{\omega}_i = -D(\lambda_{K_i})\underline{\omega}_i \tag{5.168}$$

This relation does not lead to the value of gain $\mathbf{K}$ as soon as $\mathbf{N}_{ol}(\lambda_{K_i})\underline{\omega}_i$ is a vector which is not invertible. Nevertheless assuming that n denotes the order of state matrix $\mathbf{A}$ we can apply this relation for the p closed-loop eigenvalues given by $\mathbf{\Lambda}_p$. We get:

$$\mathbf{KC}\left[\begin{array}{ccc} \underline{v}_{K_1} & \cdots & \underline{v}_{K_p} \end{array}\right] = -\left[\begin{array}{ccc} \underline{p}_1 & \cdots & \underline{p}_p \end{array}\right] \tag{5.169}$$

Where vectors $\underline{v}_{K_i}$ and $\underline{p}_i$ are given by:

$$\left\{\begin{array}{l} \underline{v}_{K_i} = \mathbf{N}_{ol}(\lambda_{K_i})\,\underline{\omega}_i \\ \underline{p}_i = D(\lambda_{K_i})\,\underline{\omega}_i \end{array}\right. \quad \forall\, i = 1, \cdots, p \tag{5.170}$$

We finally retrieve expression (5.74) of the static *state* feedback gain matrix $\mathbf{K}$ to get the p closed-loop eigenvalues given by $\mathbf{\Lambda}_p$:

$$\mathbf{K} = -\mathbf{P}\,(\mathbf{CV})^{-1} \tag{5.171}$$

where:

$$\left\{\begin{array}{l} \mathbf{P} = \left[\begin{array}{ccc} D(\lambda_{K_1})\,\underline{\omega}_1 & \cdots & D(\lambda_{K_p})\,\underline{\omega}_p \end{array}\right] := \left[\begin{array}{ccc} \underline{p}_1 & \cdots & \underline{p}_p \end{array}\right] \\ \mathbf{V} = \left[\begin{array}{ccc} \mathbf{N}_{ol}(\lambda_{K_1})\,\underline{\omega}_1 & \cdots & \mathbf{N}_{ol}(\lambda_{K_p})\,\underline{\omega}_p \end{array}\right] := \left[\begin{array}{ccc} \underline{v}_{K_1} & \cdots & \underline{v}_{K_p} \end{array}\right] \end{array}\right. \tag{5.172}$$

As shown by Duan[8], by duality (5.171) can be changed as follows:

$$\mathbf{K} = -\mathbf{P}\left(\mathbf{B}^T\mathbf{V}\right)^{-1} \tag{5.173}$$

Then relation (5.172) still holds when vectors $\underline{v}_{K_i}$ and $\underline{p}_i$ are defined as follows where vector $\underline{\nu}_i \neq \underline{0} \in \mathbb{C}^p$ is used as a design parameter:

$$\left\{\begin{array}{l} \underline{v}_{K_i} = \mathbf{N}^d_{ol}(\lambda_{K_i})\,\underline{\nu}_i \\ \underline{p}_i = D^d(\lambda_{K_i})\,\underline{\nu}_i \end{array}\right. \quad \forall\, i = 1, \cdots, m \tag{5.174}$$

[7] L. S. Shieh, H. M. Dib and R. E. Yates, Sequential design of linear quadratic state regulators via the optimal root-locus techniques, IEE Proceedings D - Control Theory and Applications, vol. 135, no. 4, pp. 289-294, July 1988.

[8] G. R. Duan, Parametric eigenstructure assignment via output feedback based on singular value decompositions, Proceedings of the 40th IEEE Conference on Decision and Control (Cat. No.01CH37228), Orlando, FL, USA, 2001, pp. 2665-2670 vol.3.

where $\mathbf{N}_{ol}^{d}(s) := \operatorname{adj}\left(s\mathbb{I} - \mathbf{A}^T\right)\mathbf{C}^T \in \mathbb{R}^{n\times p}$, where $\operatorname{adj}\left(s\mathbb{I} - \mathbf{A}^T\right)$ stands for the adjugate matrix of $s\mathbb{I} - \mathbf{A}^T$, and $D^d(s) := \det\left(s\mathbb{I} - \mathbf{A}^T\right)$ is the determinant of $s\mathbb{I} - \mathbf{A}^T$, that is the characteristic polynomial of the dual plant :

$$\left(s\mathbb{I} - \mathbf{A}^T\right)^{-1}\mathbf{C}^T = \frac{\operatorname{adj}\left(s\mathbb{I} - \mathbf{A}^T\right)\mathbf{C}^T}{\det\left(s\mathbb{I} - \mathbf{A}^T\right)} := \frac{\mathbf{N}_{ol}^{d}(s)}{D^d(s)} \tag{5.175}$$

Furthermore, and assuming that $\operatorname{rank}(\mathbf{B}) = m$ and $\operatorname{rank}(\mathbf{C}) = p$, the remaining $n - p$ eigenvalues of the closed-loop matrix $\mathbf{A} - \mathbf{BKC}$ can be achieved by selecting parameter vectors $\underline{\omega}_i \neq \underline{0}$ and $\underline{\nu}_j \neq \underline{0}$ such that the following constraints hold:

$$\underline{\nu}_j^T \mathbf{N}_{ji}\,\underline{\omega}_i = 0 \text{ where } \begin{cases} \underline{\omega}_i \neq \underline{0} \in \mathbb{C}^{m\times 1},\ i = 1, \cdots, p \\ \underline{\nu}_j \neq \underline{0} \in \mathbb{C}^{p\times 1},\ j = p+1, \cdots, n \end{cases} \tag{5.176}$$

where $p \times m$ matrix $\mathbf{N}_{ji}$ is defined as follows:

$$\mathbf{N}_{ji} = \left(\mathbf{N}_{ol}^{d}(\lambda_{K_j})\right)^T \mathbf{N}_{ol}(\lambda_{K_i}) \tag{5.177}$$

Matrices $\mathbf{N}_{ol}^{d}(\lambda_{K_i})$ and $\mathbf{N}_{ol}(\lambda_{K_j})$ are defined in (5.164) and (5.175).
The last component of each parameter vectors as follows is set as follows:

- If the eigenvalue λ_{K_i} is real, the last component of parameter vectors $\underline{\omega}_i$ and $\underline{\nu}_i$ is set to 1 ;
- If the eigenvalue λ_{K_i} and λ_{K_j} are complex conjugate, the last component of parameter vectors $\underline{\omega}_i$ and $\underline{\nu}_i$ is set to $1 + j$ whereas the last component of parameter vectors $\underline{\omega}_j$ and $\underline{\nu}_j$ is set to $1 - j$;
- More generally, Duan[8] has shown that to fulfill (5.176) parameter vectors $\underline{\omega}_i$ and $\underline{\nu}_i$ are real as soon as λ_{K_i} is real. But if λ_{K_i} and λ_{K_j} are complex conjugate, that is $\lambda_{K_i} = \bar{\lambda}_{K_j}$, then $\underline{\omega}_i = \bar{\underline{\omega}}_j$ and $\underline{\nu}_i = \bar{\underline{\nu}}_j$.

Alexandridis & al.[9] have shown that given a set $\mathbf{\Lambda}$ of n eigenvalues λ_{K_i} for the closed-loop system, we have to determine p parameter vectors $\underline{\omega}_i$ such that there exits $n - p$ parameter vectors $\underline{\nu}_i$ which solve the set of bilinear algebraic equations (5.176).

From (5.176) there is $p \times (n-p)$ equality constraints which shall be fulfilled. On the other hand, p parameter vectors $\underline{\omega}_i$ with $m-1$ free parameters (the last component is set) and $n-p$ parameter vectors $\underline{\nu}_j$ with $p-1$ free parameters (the last component is set) have to be found. A necessary condition for constraints (5.176) to be solvable is that the number of equations must be equal or less than the sum of the free parameters:

$$p \times (n-p) \leq p \times (m-1) + (n-p) \times (p-1) \Leftrightarrow m \times p \geq n \tag{5.178}$$

[9] A. T. Alexandridis and P. N. Paraskevopoulos, A new approach to eigenstructure assignment by output feedback, IEEE Transactions on Automatic Control, vol. 41, no. 7, pp. 1046-1050, July 1996

Thus a necessary condition for this set to be solvable is that $m \times p \geq n$. However the condition $m \times p \geq n$ does not assure that a solution exists (it is not a sufficient condition).

Furthermore, in the particular case where $m+p \geq n+1$, parameter vectors $\underline{\nu}_j,\ j = p+1, \cdots, n$ can be arbitrarily set. In that case the set of equations (5.176) reduce to a linear system of algebraic equations with $\underline{\omega}_i,\ i = 1, \cdots, p$ as unknown parameters.

As mentioned by Duan[8], an efficient manner to solve constraints (5.176) is to use a Singular Value Decomposition (SVD) of matrices $\mathbf{N}_{ji}$ which reads as follows:

$$\begin{cases} \mathbf{N}_{ji} = \mathbf{U}_j \mathbf{\Sigma}_{ji} \mathbf{V}_i^H \\ \mathbf{\Sigma}_{ji} = \begin{bmatrix} \mathbf{\Lambda}_q & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \\ \mathbf{\Lambda}_q = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \dots \\ 0 & 0 & 0 & \sigma_q \end{bmatrix} \end{cases} \tag{5.179}$$

where:

$$\begin{cases} \mathbf{U}_j \text{ and } \mathbf{V}_i \text{ are unitary matrices} \\ \sigma_i \in \mathbb{R}^+ \quad \forall i = 1, 2, \cdots, q \\ \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_q > 0 \\ q = \min(m, p) \text{ assuming that } \mathbf{N}_{ji} \text{ has no eigenvalue equal to } 0 \end{cases} \tag{5.180}$$

In all cases, and assuming that $\underline{\omega}_i$ and possibly $\underline{\nu}_j$ have be chosen such that $\det(\mathbf{CV}) \neq 0$, static output feedback $\mathbf{K}$ is computed thanks to (5.171).

5.7.2 Changing PID controller into static output feedback

We present hereafter some results provided by Zheng & al.[10] which transforms a PID controller to static output feedback.

We consider the following linear time-invariant system:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.181}$$

And the following PID controller where matrices $\mathbf{K}_p$, $\mathbf{K}_i$ and $\mathbf{K}_d$ have to be designed:

$$\underline{u}(t) = -\left(\mathbf{K}_p\, \underline{e}(t) + \mathbf{K}_i \int_0^t \underline{e}(\tau) d\tau + \mathbf{K}_d \frac{d}{dt}\underline{e}(t)\right) \tag{5.182}$$

where:

$$\underline{e}(t) = \underline{y}(t) - \underline{r}(t) \tag{5.183}$$

Let's denote $\underline{x}_a(t)$ the augmented state-space vector defined as follows:

$$\underline{x}_a(t) = \begin{bmatrix} \underline{x}(t) \\ \int_0^t \underline{e}(\tau) d\tau \end{bmatrix} \tag{5.184}$$

[10] Zheng, F., Wang, Q.-G. & Lee, T. H. (2002). On the design of multivariable PID controllers via LMI approach. Automatica 38, 517-526

Thus:

$$\dot{\underline{x}}_a(t) = \mathbf{A}_a \underline{x}_a(t) + \mathbf{B}_a \underline{u}(t) + \begin{bmatrix} \mathbf{0} \\ -\mathbb{I} \end{bmatrix} \underline{r}(t) \tag{5.185}$$

where:

$$\begin{cases} \mathbf{A}_a = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \\ \mathbf{B}_a = \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix} \end{cases} \tag{5.186}$$

Furthermore, assuming that $\dot{\underline{r}}(t) = \underline{0}$, we have:

$$\dot{\underline{r}}(t) = \underline{0} \Rightarrow \frac{d}{dt}\underline{e}(t) = \mathbf{C}\dot{\underline{x}}(t) = \mathbf{CA}\underline{x}(t) + \mathbf{CB}\underline{u}(t) \tag{5.187}$$

Using the definition of $\underline{x}_a(t)$, the PID controller reads:

$$\begin{aligned} \underline{u}(t) &= -\left(\mathbf{K}_p\, \underline{e}(t) + \mathbf{K}_i \int_0^t \underline{e}(\tau)d\tau + \mathbf{K}_d \tfrac{d}{dt}\underline{e}(t)\right) \\ &= -\mathbf{K}_p\mathbf{C}\underline{x}(t) + \mathbf{K}_p\mathbf{C}\underline{r}(t) - \mathbf{K}_i \tfrac{d}{dt}\underline{e}(t) - \mathbf{K}_d\left(\mathbf{CA}\underline{x}(t) + \mathbf{CB}\underline{u}(t)\right) \\ &= -\mathbf{K}_p \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix} \underline{x}_a(t) - \mathbf{K}_i \begin{bmatrix} \mathbf{0} & \mathbb{I} \end{bmatrix} \underline{x}_a(t) - \mathbf{K}_d \begin{bmatrix} \mathbf{CA} & \mathbf{0} \end{bmatrix} \underline{x}_a(t) \\ &\qquad\qquad -\mathbf{K}_d\mathbf{CB}\underline{u}(t) + \mathbf{K}_p\mathbf{C}\underline{r}(t) \\ &= -\begin{bmatrix} \mathbf{K}_p & \mathbf{K}_i & \mathbf{K}_d \end{bmatrix} \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \\ \mathbf{CA} & \mathbf{0} \end{bmatrix} \underline{x}_a(t) - \mathbf{K}_d\mathbf{CB}\underline{u}(t) + \mathbf{K}_p\mathbf{C}\underline{r}(t) \end{aligned} \tag{5.188}$$

We will assume that $\mathbb{I}+\mathbf{K}_d\mathbf{CB}$ is invertible and define $\mathbf{C}_a$ and $\mathbf{K}_a$ as follows:

$$\begin{cases} \mathbf{C}_a = \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \\ \mathbf{CA} & \mathbf{0} \end{bmatrix} \\ \mathbf{K}_a = (\mathbb{I} + \mathbf{K}_d\mathbf{CB})^{-1} \begin{bmatrix} \mathbf{K}_p & \mathbf{K}_i & \mathbf{K}_d \end{bmatrix} \end{cases} \tag{5.189}$$

Let $\widetilde{\mathbf{K}}_p$, $\widetilde{\mathbf{K}}_i$ and $\widetilde{\mathbf{K}}_d$ be defined as follows:

$$\begin{cases} \widetilde{\mathbf{K}}_p = (\mathbb{I} + \mathbf{K}_d\mathbf{CB})^{-1}\mathbf{K}_p \\ \widetilde{\mathbf{K}}_i = (\mathbb{I} + \mathbf{K}_d\mathbf{CB})^{-1}\mathbf{K}_i \\ \widetilde{\mathbf{K}}_d = (\mathbb{I} + \mathbf{K}_d\mathbf{CB})^{-1}\mathbf{K}_d \end{cases} \tag{5.190}$$

Assuming that $\widetilde{\mathbf{K}}_p$, $\widetilde{\mathbf{K}}_i$ and $\widetilde{\mathbf{K}}_d$ are known, gains $\mathbf{K}_p$, $\mathbf{K}_i$ and $\mathbf{K}_d$ are obtained as follows where it can be shown[10] that matrix $\mathbb{I} - \mathbf{CB}\widetilde{\mathbf{K}}_d$ is always invertible:

$$\begin{cases} \mathbf{K}_d = \widetilde{\mathbf{K}}_d\left(\mathbb{I} - \mathbf{CB}\widetilde{\mathbf{K}}_d\right)^{-1} \\ \mathbf{K}_p = (\mathbb{I} + \mathbf{K}_d\mathbf{CB})\,\widetilde{\mathbf{K}}_p \\ \mathbf{K}_i = (\mathbb{I} + \mathbf{K}_d\mathbf{CB})\,\widetilde{\mathbf{K}}_i \end{cases} \tag{5.191}$$

Thus the problem of PID controller design is changed into the following static output feedback problem:

$$\begin{cases} \dot{\underline{x}}_a(t) = \mathbf{A}_a\underline{x}_a(t) + \mathbf{B}_a\underline{u}(t) \\ \underline{y}_a(t) = \mathbf{C}_a\underline{x}_a(t) \\ \underline{u}(t) = -\mathbf{K}_a\underline{y}_a(t) + (\mathbb{I} + \mathbf{K}_d\mathbf{CB})^{-1}\mathbf{K}_p\mathbf{C}\underline{r}(t) \end{cases} \tag{5.192}$$

It is worth noticing that the same results are obtained, but without the assumption that $\dot{\underline{r}}(t) = \underline{0}$, when a PI-D controller is used; for such a controller the term multiplied by $\mathbf{K}_d$ is $\underline{y}(t)$ rather than $\underline{e}(t)$:

$$\underline{u}(t) = -\left(\mathbf{K}_p\,\underline{e}(t) + \mathbf{K}_i \int_0^t \underline{e}(\tau)d\tau + \mathbf{K}_d \frac{d}{dt}\underline{y}(t)\right) \tag{5.193}$$

5.7.3 Adding integrators, controllability and observability indexes

We consider the following controllable and observable state-space representation where the state vector $\underline{x}$ is of dimension n (that is the size of state matrix $\mathbf{A}$). In addition $\underline{y}(t)$ denotes the output vector and $\underline{u}(t)$ the input vector. We will assume that the feedforward gain matrix $\mathbf{D}$ is zero ($\mathbf{D} = \mathbf{0}$):

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.194}$$

Brasch & Pearson[11] have computed the number n_i of integrators that can be added to increase the size of the output vector:

$$n_i = \min(p_c - 1, p_o - 1) \tag{5.195}$$

where p_c is the controllability index of the plant and p_o the observability index of the plant.

The controllability index p_c of the plant is the smallest integers such that:

$$\text{rank}\left(\begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{p_c-1}\mathbf{B} \end{bmatrix}\right) = n \tag{5.196}$$

Similarly, the observability index p_o of the augmented plant is the smallest integers such that:

$$\text{rank}\left(\begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{p_o-1} \end{bmatrix}\right) = n \tag{5.197}$$

The compensator in cascade with the plant will be taken to be n_i integrations of the component $y_i(t) = \mathbf{C}_i\underline{x}(t)$ of the output vector $\underline{y}(t)$ of the plant:

$$\begin{cases} \dot{x}_{i,1}(t) = y_i(t) = \mathbf{C}_i\underline{x}(t) \\ \quad\vdots \\ \dot{x}_{i,n_i}(t) = x_{i,n_i-1}(t) \end{cases} \tag{5.198}$$

Furthermore the control $\underline{u}(t)$ of the augmented plant, that is the plant and the n_i integrators, will be taken to be the actual input $\underline{u}(t)$ of the plant and the

[11] F. Brasch and J. Pearson, Pole placement using dynamic compensators, IEEE Transactions on Automatic Control, vol. 15, no. 1, pp. 34-43, February 1970.

n_i integrations of the output $y_i(t)$ of the plant:

$$\underline{u}(t) = \begin{bmatrix} \underline{u}(t) \\ x_{i,1}(t) \\ \vdots \\ x_{i,n_i}(t) \end{bmatrix} \tag{5.199}$$

Then we define matrices $\mathbf{A}_{n_i}$, $\mathbf{B}_{n_i}$ and $\mathbf{C}_{n_i}$ of the augmented plant as follows where $\mathbf{0}_{n_i}$ is the null matrix of size n_i and $\mathbb{I}_{n_i}$ is the identity matrix of size n_i:

$$\begin{cases} \mathbf{A}_{n_i} = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0}_{n_i} \end{array}\right] \\ \mathbf{B}_{n_i} = \left[\begin{array}{c|c} \mathbf{B} & \mathbf{0} \\ \hline \mathbf{0} & \mathbb{I}_{n_i} \end{array}\right] \\ \mathbf{C}_{n_i} = \left[\begin{array}{c|c} \mathbf{C} & \mathbf{0} \\ \hline \mathbf{0} & \mathbb{I}_{n_i} \end{array}\right] \end{cases} \tag{5.200}$$

Alternatively matrices $\mathbf{A}_{n_i}$, $\mathbf{B}_{n_i}$ and $\mathbf{C}_{n_i}$ of the augmented plant can be defined as follows where $\mathbf{0}_{p\times r}$ is the null matrix of size $p \times r$ and $\mathbb{I}_{n_i}$ is the identity matrix of size n_i:

$$\begin{cases} \mathbf{A}_{n_i} = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{0} \\ \hline \mathbf{C}_i & \mathbf{0}_{1\times n_i} \\ \mathbf{0}_{(n_i-1)\times n} & \mathbb{I}_{n_i-1} \vdots \mathbf{0}_{(n_i-1)\times 1} \end{array}\right] \\ \mathbf{B}_{n_i} = \left[\begin{array}{c} \mathbf{B} \\ \hline \mathbf{0} \end{array}\right] \\ \mathbf{C}_{n_i} = \left[\begin{array}{c|c} \mathbf{C} & \mathbf{0} \\ \hline \mathbf{0} & \mathbb{I}_{n_i} \end{array}\right] \end{cases} \tag{5.201}$$

The interest of the preceding state space representation of the augmented plant is that its input vector $\underline{u}_{n_i}(t)$ is the same than the input vector $\underline{u}(t)$ of the actual plant :

$$\underline{u}_{n_i}(t) = \underline{u}(t) \tag{5.202}$$

5.8 Dynamical output feedback controller

5.8.1 From dynamical output feedback to observer-based control

Again we consider the state-space representation (5.203) where the state vector $\underline{x}$ is of dimension n (that is the size of state matrix $\mathbf{A}$). In addition $\underline{y}(t)$ denotes the output vector and $\underline{u}(t)$ the input vector. We will assume that the feedforward gain matrix $\mathbf{D}$ is zero ($\mathbf{D} = \mathbf{0}$):

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.203}$$

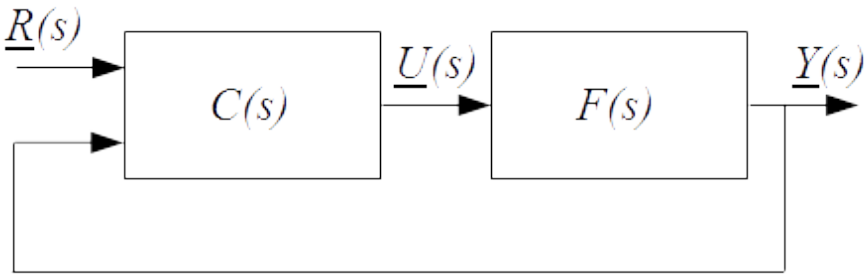

Figure 5.6: Plant with dynamic compensator

Following Radman[12], we assume that this system is controllable and observable and is controlled as shown in Figure 5.6 by a dynamical output feedback controller $\mathbf{C}(s)$ of dimension n_c (that is the size of $\mathbf{A}_c$) whose state space representation reads:

$$\begin{cases} \dot{\underline{x}}_c(t) = \mathbf{A}_c\underline{x}_c(t) + \mathbf{B}_{cy}\underline{y}(t) + \mathbf{B}_{cr}\underline{r}(t) \\ \underline{u}(t) = \mathbf{C}_c\underline{x}_c(t) + \mathbf{D}_{cy}\underline{y}(t) + \mathbf{D}_{cr}\underline{r}(t) \end{cases} \tag{5.204}$$

Assuming that compensator $\mathbf{C}(s)$ has the same dimension than plant $\mathbf{F}(s)$, that is $n_c = n$, and from the following settings:

$$\begin{cases} \mathbf{B}_{cy} = \mathbf{L} \\ \mathbf{B}_{cr} = \mathbf{B} \\ \mathbf{C}_c = -\mathbf{K} \\ \mathbf{D}_{cy} = \mathbf{0} \\ \mathbf{D}_{cr} = \mathbb{I} \end{cases} \tag{5.205}$$

We get:

$$\begin{cases} \dot{\underline{x}}_c(t) = \mathbf{A}_c\underline{x}_c(t) + \mathbf{L}\underline{y}(t) + \mathbf{B}\underline{r}(t) \\ \underline{u}(t) = -\mathbf{K}\underline{x}_c(t) + \underline{r}(t) \end{cases} \tag{5.206}$$

From the second relation we get $\underline{r}(t) = \underline{u}(t) + \mathbf{K}\underline{x}_c(t)$. Thus the previous state space representation reads:

$$\begin{cases} \dot{\underline{x}}_c(t) = (\mathbf{A}_c + \mathbf{BK})\,\underline{x}_c(t) + \mathbf{B}\underline{u}(t) + \mathbf{L}\underline{y}(t) \\ \underline{u}(t) = -\mathbf{K}\underline{x}_c(t) + \underline{r}(t) \end{cases} \tag{5.207}$$

Thus the dynamical output feedback controller $\mathbf{C}(s)$ can be seen as an observer-based controller with gain $\mathbf{K}$ that uses $\underline{x}_c(t)$ as an estimate of the plant state $\underline{x}(t)$.

5.8.2 Dynamical output feedback

We now assume that plant $\mathbf{F}(s)$ is controlled by the dynamical output feedback controller $\mathbf{C}(s)$ defined in (5.204) where $n_c \neq n$ (n_c is the size of $\mathbf{A}_c$):

[12] G. Radman, Design of a dynamic compensator for complete pole-zero placement, The Twentieth Southeastern Symposium on System Theory, Charlotte, NC, USA, 1988, pp. 176-177

$$\begin{cases} \dot{\underline{x}}_c(t) = \mathbf{A}_c\underline{x}_c(t) + \mathbf{B}_{cy}\underline{y}(t) + \mathbf{B}_{cr}\underline{r}(t) \\ \underline{u}(t) = \mathbf{C}_c\underline{x}_c(t) + \mathbf{D}_{cy}\underline{y}(t) + \mathbf{D}_{cr}\underline{r}(t) \end{cases} \tag{5.208}$$

It is worth noticing that we retrieve the static output feedback controller:

$$\underline{u}(t) = \mathbf{H}\underline{r}(t) - \mathbf{K}_c\underline{y}(t) \tag{5.209}$$

When setting:

$$\begin{cases} \mathbf{A}_c = \mathbf{B}_{cy} = \mathbf{B}_{cr} = \mathbf{C}_c = \mathbf{0} \\ \mathbf{D}_{cy} = -\mathbf{K}_c \\ \mathbf{D}_{cr} = \mathbf{H} \end{cases} \tag{5.210}$$

When the dynamical output feedback controller is of order $n_c \leq n$ the design problem can be bring back to a static output feedback controller[13] by introducing a new control $\underline{v}(t)$ defined by:

$$\underline{v}(t) = \dot{\underline{x}}_c(t) \tag{5.211}$$

and by considering the following augmented output vector $\underline{y}_a(t)$, input vector $\underline{u}_a(t)$ and state vector $\underline{x}_a(t)$:

$$\begin{cases} \underline{x}_a(t) = \begin{bmatrix} \underline{x}(t) \\ \underline{x}_c(t) \end{bmatrix} \\ \underline{u}_a(t) = \begin{bmatrix} \underline{u}(t) \\ \underline{v}(t) \end{bmatrix} \\ \underline{y}_a(t) = \begin{bmatrix} \underline{y}(t) \\ \underline{x}_c(t) \end{bmatrix} \end{cases} \tag{5.212}$$

Then the augmented equivalent *open-loop* system reads:

$$\begin{cases} \begin{bmatrix} \dot{\underline{x}}(t) \\ \dot{\underline{x}}_c(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \underline{x}(t) \\ \underline{x}_c(t) \end{bmatrix} + \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{u}(t) \\ \underline{v}(t) \end{bmatrix} \\ \begin{bmatrix} \underline{y}(t) \\ \underline{x}_c(t) \end{bmatrix} = \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{x}(t) \\ \underline{x}_c(t) \end{bmatrix} \end{cases} \tag{5.213}$$

That is in a more compact form:

$$\begin{cases} \dot{\underline{x}}_a(t) = \mathbf{A}_a\underline{x}_a(t) + \mathbf{B}_a\underline{u}_a(t) \\ \underline{y}_a(t) = \mathbf{C}_a\underline{x}_a(t) \end{cases} \tag{5.214}$$

Where:

$$\begin{cases} \mathbf{A}_a = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \\ \mathbf{B}_a = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \\ \mathbf{C}_a = \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \end{cases} \tag{5.215}$$

[13] V.L. Syrmos, C. Abdallah, P. Dorato, Static Output Feedback: a Survey, Proceedings of the 33rd IEEE Conference on Decision and Control, 1994

It is worth noticing that the dynamical output feedback controller $\mathbf{C}(s)$ defined in (5.204) becomes static in the augmented state-space:

$$\begin{bmatrix} \underline{u}(t) \\ \underline{v}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{cy} & \mathbf{C}_c \\ \mathbf{B}_{cy} & \mathbf{A}_c \end{bmatrix} \begin{bmatrix} \underline{y}(t) \\ \underline{x}_c(t) \end{bmatrix} + \begin{bmatrix} \mathbf{D}_{cr} \\ \mathbf{B}_{cr} \end{bmatrix} \underline{r}(t) \tag{5.216}$$

That is in a more compact form:

$$\underline{u}_a(t) = -\mathbf{K}_a \underline{y}_a(t) + \mathbf{H}_a \underline{r}(t) \tag{5.217}$$

Where:

$$\begin{cases} \mathbf{K}_a = -\begin{bmatrix} \mathbf{D}_{cy} & \mathbf{C}_c \\ \mathbf{B}_{cy} & \mathbf{A}_c \end{bmatrix} \\ \mathbf{H}_a = \begin{bmatrix} \mathbf{D}_{cr} \\ \mathbf{B}_{cr} \end{bmatrix} \end{cases} \tag{5.218}$$

Assuming that $\underline{u}(t)$ is of dimension m, $\underline{x}(t)$ is of dimension n, $\underline{x}_c(t)$ is of dimension n_c and $\underline{y}(t)$ is of dimension p we conclude that:

- $\mathbf{A}_a$ is a square matrix of dimension $(n + n_c)$;
- $\mathbf{B}_a$ is a matrix of size $(n + n_c) \times (m + n_c)$;
- $\mathbf{C}_a$ is a matrix of size $(p + n_c) \times (n + n_c)$;
- $\mathbf{K}_a$ is a matrix of size $(m + n_c) \times (p + n_c)$;

If we wish to apply the Roppenecker's formula (5.71) to set the static output feedback gain $\mathbf{K}_a$ so that n_p predefined closed-loop eigenvalues are achieved, we have to notice that matrix $\mathbf{C}_a\mathbf{V}_a$ is a $(p+n_c) \times n_p$ matrix. Consequently matrix $\mathbf{C}_a\mathbf{V}_a$ is square and possibly invertible as soon as:

$$p + n_c = n_p \tag{5.219}$$

- In the case of state feedback we have $n_c = 0$ and $p = n$ thus the number of eigenvalues which can be predefined is $n_p = n$.
- In the case of output feedback the number n_p of eigenvalues which can be predefined is obviously lower or equal to the size $n + n_c$ of the augmented state matrix $\mathbf{A}_a$:

$$n_p \leq n + n_c \tag{5.220}$$

Assuming $p + n_c = n_p$ so that $\mathbf{C}_a\mathbf{V}_a$ is a square matrix we conclude that there are $n - p$ remaining eigenvalues whose location is not controlled through output feedback.

Using the relations $\underline{u}_a(t) = -\mathbf{K}_a \underline{y}_a(t) + \mathbf{H}_a \underline{r}(t)$ and $\underline{y}_a(t) = \mathbf{C}_a \underline{x}_a(t)$ the dynamics of the closed-loop system reads:

$$\begin{aligned} \dot{\underline{x}}_a(t) &= \mathbf{A}_a \underline{x}_a(t) + \mathbf{B}_a \underline{u}_a(t) \\ &= \mathbf{A}_a \underline{x}_a(t) + \mathbf{B}_a \left(-\mathbf{K}_a \underline{y}_a(t) + \mathbf{H}_a \underline{r}(t) \right) \\ &= \left(\mathbf{A}_a - \mathbf{B}_a \mathbf{K}_a \mathbf{C}_a \right) \underline{x}_a(t) + \mathbf{B}_a \mathbf{H}_a \underline{r}(t) \end{aligned} \tag{5.221}$$

The product $\mathbf{B}_a\mathbf{K}_a\mathbf{C}_a$ expands as follows:

$$\begin{aligned}\mathbf{B}_a\mathbf{K}_a\mathbf{C}_a &= \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \mathbf{D}_{cy} & \mathbf{C}_c \\ \mathbf{B}_{cy} & \mathbf{A}_c \end{bmatrix} \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \mathbf{D}_{cy}\mathbf{C} & \mathbf{C}_c \\ \mathbf{B}_{cy}\mathbf{C} & \mathbf{A}_c \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{B}\mathbf{D}_{cy}\mathbf{C} & \mathbf{B}\mathbf{C}_c \\ \mathbf{B}_{cy}\mathbf{C} & \mathbf{A}_c \end{bmatrix}\end{aligned} \tag{5.222}$$

And consequently $\mathbf{A}_a - \mathbf{B}_a\mathbf{K}_a\mathbf{C}_a$ reads:

$$\mathbf{A}_a - \mathbf{B}_a\mathbf{K}_a\mathbf{C}_a = \begin{bmatrix} \mathbf{A} - \mathbf{B}\mathbf{D}_{cy}\mathbf{C} & -\mathbf{B}\mathbf{C}_c \\ -\mathbf{B}_{cy}\mathbf{C} & -\mathbf{A}_c \end{bmatrix} \tag{5.223}$$

The transfer function $\mathbf{G}(s)$ of the closed-loop system between the output vector $\underline{y}(t)$ and the reference input vector $\underline{r}(t)$ reads:

$$\underline{y}(t) = \mathbf{G}(s)\underline{r}(t) \tag{5.224}$$

Where:

$$\begin{aligned}\mathbf{G}(s) &= \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix} (s\mathbb{I} - (\mathbf{A}_a - \mathbf{B}_a\mathbf{K}_a\mathbf{C}_a))^{-1}\, \mathbf{B}_a\mathbf{H}_a \\ &= \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix} (s\mathbb{I} - (\mathbf{A}_a - \mathbf{B}_a\mathbf{K}_a\mathbf{C}_a))^{-1} \begin{bmatrix} \mathbf{B}\mathbf{D}_{cr} \\ \mathbf{B}_{cr} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} s\mathbb{I} - (\mathbf{A} - \mathbf{B}\mathbf{D}_{cy}\mathbf{C}) & \mathbf{B}\mathbf{C}_c \\ \mathbf{B}_{cy}\mathbf{C} & s\mathbb{I} + \mathbf{A}_c \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{B}\mathbf{D}_{cr} \\ \mathbf{B}_{cr} \end{bmatrix}\end{aligned} \tag{5.225}$$

5.9 Mode decoupling controller

5.9.1 Eigenstructure assignment

We consider the following state-space representation of a controllable and observable system where $\underline{u}(t) \in \mathbb{R}^m$ denotes the control input, $\underline{x}(t) \in \mathbb{R}^n$ the state vector and $\underline{y}(t) \in \mathbb{R}^p$ the output vector:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.226}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$, $\mathbf{C} \in \mathbb{R}^{p\times n}$.

Assume the following state feedback control:

$$\underline{u}(t) = -\mathbf{K}\underline{x}(t) + \underline{r}(t) \tag{5.227}$$

Then the closed-loop system reads:

$$\begin{cases} \dot{\underline{x}}(t) = (\mathbf{A} - \mathbf{B}\mathbf{K})\,\underline{x}(t) + \mathbf{B}\underline{r}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.228}$$

Similarly to the open-loop case the transfer function $\mathbf{G}_K(s)$ of the closed-loop system when the control $\underline{u}(t)$ is $-\mathbf{K}\underline{x}(t) + \underline{r}(t)$ reads:

$$\underline{Y}(s) = \mathbf{G}_K(s)\underline{R}(s) \tag{5.229}$$

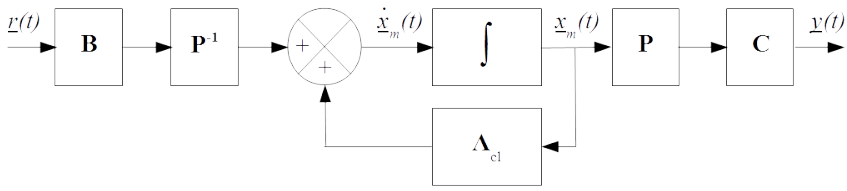

Figure 5.7: Modal decomposition of a transfer function

Where:

$$\mathbf{G}_K(s) = \mathbf{C}\left(s\mathbb{I} - (\mathbf{A} - \mathbf{BK})\right)^{-1}\mathbf{B} \tag{5.230}$$

As in the open-loop case the transfer function $\mathbf{G}_K(s)$ of the closed-loop system may be expressed as a function of the closed-loop eigenvalues λ_{K_i} and the left and right eigenvectors of matrix $\mathbf{A}-\mathbf{BK}$. Assuming that matrix $\mathbf{A}-\mathbf{BK}$ is diagonalizable we have:

$$\mathbf{G}_K(s) = \mathbf{C}\left(\sum_{i=1}^{n} \frac{\underline{v}_{K_i}\underline{w}_{K_i}^T}{s - \lambda_{K_i}}\right)\mathbf{B} \tag{5.231}$$

Figure 5.7 presents the modal decomposition of the transfer function where $\underline{x}_m(t)$ is the state vector expressed in the modal basis and where matrices $\mathbf{\Lambda}_{cl}$, $\mathbf{P}$ and $\mathbf{P}^{-1}$ are defined as follows:

$$\mathbf{A} - \mathbf{BK} = \mathbf{P}\mathbf{\Lambda}_{cl}\mathbf{P}^{-1} \text{ where } \begin{cases} \mathbf{\Lambda}_{cl} = \begin{bmatrix} \lambda_{K_1} & & \\ & \ddots & \\ & & \lambda_{K_n} \end{bmatrix} \\ \mathbf{P} = \begin{bmatrix} \underline{v}_{K_1} & \cdots & \underline{v}_{K_n} \end{bmatrix} \\ \mathbf{P}^{-1} = \mathbf{P}^T = \begin{bmatrix} \underline{w}_{K_1}^T \\ \vdots \\ \underline{w}_{K_n}^T \end{bmatrix} \end{cases} \tag{5.232}$$

Vector $\underline{v}_{K_i}$ is a right eigenvector corresponding to eigenvalue λ_{K_i}: it is a real vector if the eigenvalue λ_{K_i} is real, a complex vector otherwise.

The components of the *desired* eigenvector $\underline{v}_{K_i}$ can be used for decoupling. Indeed we can see from the modal decomposition of the transfer function provided in (5.231) that:

- Mode λ_{K_i} will not appear in the jth component of output vector $\underline{y}(t)$ if the following relation holds where $\underline{c}_j^T$ represents the jth row of $\mathbf{C}$:

$$\underline{c}_j^T\,\underline{v}_{K_i} = 0 \tag{5.233}$$

- Alternatively mode λ_{K_i} will not be excited by the jth component of control vector $\underline{u}(t)$ if the following relation holds where $\underline{b}_j$ represents the jth column of $\mathbf{B}$:

$$\underline{w}_{K_i}^T\,\underline{b}_j = 0 \tag{5.234}$$

5.9.2 Design procedure

We consider the following state-space representation of a controllable and observable system where $\underline{u}(t) \in \mathbb{R}^m$ denotes the control input, $\underline{x}(t) \in \mathbb{R}^n$ the state vector and $\underline{y}(t) \in \mathbb{R}^p$ the output vector:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.235}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$, $\mathbf{C} \in \mathbb{R}^{p\times n}$.
Assume the following state feedback control:

$$\underline{u}(t) = -\mathbf{K}\underline{x}(t) + \underline{r}(t) \tag{5.236}$$

Then the closed-loop system reads:

$$\begin{cases} \dot{\underline{x}}(t) = (\mathbf{A} - \mathbf{BK})\,\underline{x}(t) + \mathbf{B}\underline{r}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.237}$$

The following design procedure for mode decoupling and eigenstructure assignment has been suggested by *Chouaib and Pradin*[14]

- Assuming that state matrix $\mathbf{A}$ is of order n and that $(\mathbf{A}, \mathbf{B})$ is controllable, we define n closed-loop eigenvalues $\lambda_{K_1}, \cdots, \lambda_{K_n}$. Thus for each closed-loop eigenvalue λ_{K_i}, and denoting $\underline{v}_{K_i}$ the corresponding eigenvector, we shall have:

$$(\mathbf{A} - \mathbf{BK})\,\underline{v}_{K_i} = \lambda_{K_i}\underline{v}_{K_i} \Leftrightarrow \begin{bmatrix} \lambda_{K_i}\mathbb{I} - \mathbf{A} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \underline{v}_{K_i} \\ \mathbf{K}\,\underline{v}_{K_i} \end{bmatrix} = \underline{0} \tag{5.238}$$

- Furthermore assume that mode λ_{K_i} shall not appear in the jth component of output vector $\underline{y}(t)$. Then the following relation shall hold where $\underline{c}_j^T$ represents the jth row of $\mathbf{C}$:

$$\underline{c}_j^T\,\underline{v}_{K_i} = 0 \tag{5.239}$$

  Matrix $\mathbf{S}_i$ is then defined as follows:

$$\mathbf{S}_i = \begin{bmatrix} \lambda_{K_i}\mathbb{I} - \mathbf{A} & \mathbf{B} \\ \underline{c}_j^T & \mathbf{0} \end{bmatrix} \quad i = 1, \cdots, n \tag{5.240}$$

- Then compute matrix $\mathcal{N}_{S_i}$ whose columns constitute a basis of the right null-space of $\mathbf{S}_i$:

$$\mathbf{S}_i\,\mathcal{N}_{S_i} = \mathbf{0} \tag{5.241}$$

  It is worth noticing that matrix $\mathcal{N}_{S_i}$ can be obtained through a singular value decomposition of matrix $\mathbf{S}_i$[15]. Indeed singular value decomposition

[14] I. Chouaib and B. Pradin, On mode decoupling and minimum sensitivity by eigenstructure assignment, Electrotechnical Conference, 1994. Proceedings., 7th Mediterranean, Antalya, 1994, pp. 663-666 vol.2.

[15] P. Kocsis, R. Fonod, Eigenstructure Decoupling in State Feedback Control Design, ATP Journal plus, HMH s.r.o., 2012, ATP Journal plus, 2, pp.34-39. <hal-00847146>

of matrix $\mathbf{S}_i$ leads to the following decomposition where $\mathbf{U}_i\mathbf{U}_i^T = \mathbb{I}$ and $\mathbf{V}_i\mathbf{V}_i^T = \mathbb{I}$:

$$\mathbf{S}_i = \mathbf{U}_i \begin{bmatrix} \sigma_{i,1} & & \\ & \ddots & \mathbf{0} \\ & & \sigma_{i,n} \end{bmatrix} \mathbf{V}_i^T \tag{5.242}$$

Denoting $\mathbf{\Sigma}_i = \begin{bmatrix} \sigma_{i,1} & & \\ & \ddots & \\ & & \sigma_{i,n} \end{bmatrix}$ we get:

$$\begin{aligned} \mathbf{S}_i &= \mathbf{U}_i \begin{bmatrix} \mathbf{\Sigma}_i & \mathbf{0} \end{bmatrix} \mathbf{V}_i^T \\ \Leftrightarrow \mathbf{S}_i\mathbf{V}_i &= \mathbf{U}_i \begin{bmatrix} \mathbf{\Sigma}_i & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_i\mathbf{\Sigma}_i & \mathbf{0} \end{bmatrix} \end{aligned} \tag{5.243}$$

Let $\underline{v}_{i,1}, \underline{v}_{i,2}, \cdots, \underline{v}_{i,n+m}$ be the vectors which form matrix $\mathbf{V}_i$:

$$\mathbf{V}_i = \begin{bmatrix} \underline{v}_{i,1} & \cdots & \underline{v}_{i,n} & \underline{v}_{i,n+1} & \cdots & \underline{v}_{i,n+m} \end{bmatrix} \tag{5.244}$$

From (5.243) it is clear that the set of vectors $\underline{v}_{i,n+1}, \cdots, \underline{v}_{i,n+m}$ satisfy the following relation:

$$\mathbf{S}_i\, \underline{v}_{i,j} = \underline{0} \;\forall\; j = n+1, \cdots, n+m \tag{5.245}$$

Consequently matrix $\mathcal{N}_{S_i}$ can be defined as follows:

$$\mathcal{N}_{S_i} = \begin{bmatrix} \underline{v}_{i,n+1} & \cdots & \underline{v}_{i,n+m} \end{bmatrix} \tag{5.246}$$

– Matrix $\mathcal{N}_{S_i}$ is then compatibly partitioned between matrix $\mathbf{N}(\lambda_{K_i})$ and matrix $\mathbf{M}(\lambda_{K_i})$. Matrix $\mathbf{N}(\lambda_{K_i})$ is built from the n first rows of $\mathcal{N}_{S_i}$:

$$\mathcal{N}_{S_i} = \begin{bmatrix} \mathbf{N}(\lambda_{K_i}) \\ \mathbf{M}(\lambda_{K_i}) \end{bmatrix} \tag{5.247}$$

Each vector $\underline{\pi}_i$ which belongs to the kernel of $\mathbf{S}_i$ is characterized by a non zero parameter vector $\underline{z}_i$ such that:

$$\begin{cases} \underline{z}_i \neq \underline{0} \\ \underline{\pi}_i = \mathcal{N}_{S_i}\, \underline{z}_i \end{cases} \Rightarrow \mathbf{S}_i\mathcal{N}_{S_i}\, \underline{z}_i = \underline{0} \tag{5.248}$$

As a consequence the right eigenvector $\underline{v}_{K_i}$, which constitutes the ith column of matrix $\mathbf{P}$, can be written as follows:

$$\underline{v}_{K_i} = \mathbf{N}(\lambda_{K_i})\, \underline{z}_i \tag{5.249}$$

Parameter vector $\underline{z}_i$ may be used to minimize the sensitivity of the assigned eigenvalues as it will be seen in the next section.

– Finally decompose matrix $\mathbf{B}$ as follows where $\mathbf{Y}$ is a non-singular matrix and where $\mathbf{U}_i = \begin{bmatrix} \mathbf{U}_0 & \mathbf{U}_1 \end{bmatrix}$ is an orthogonal matrix such that:

$$\mathbf{B} = \begin{bmatrix} \mathbf{U}_0 & \mathbf{U}_1 \end{bmatrix} \begin{bmatrix} \mathbf{Y} \\ \mathbf{0} \end{bmatrix} \tag{5.250}$$

One possible way to derive this decomposition is to use the singular value decomposition of $\mathbf{B}$:

$$\mathbf{B} = \mathbf{U} \begin{bmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{bmatrix} \mathbf{V}^T \tag{5.251}$$

Where $\mathbf{\Sigma}$ is a diagonal matrix formed by the rank$(\mathbf{B}) = m$ singular values of $\mathbf{B}$ (matrix $\mathbf{B}$ is assumed to have full column rank) and where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices (i.e. $\mathbf{U}^T\mathbf{U} = \mathbb{I}$ and $\mathbf{V}^T\mathbf{V} = \mathbb{I}$).

Then we can define $\mathbf{Y} = \mathbf{\Sigma}\mathbf{V}^T$ and suitably split $\mathbf{U} = \begin{bmatrix} \mathbf{U}_0 & \mathbf{U}_1 \end{bmatrix}$ such that $\mathbf{U}_0$ has m columns:

$$\begin{cases} \mathbf{Y} = \mathbf{\Sigma}\,\mathbf{V}^T \\ \mathbf{U} = \begin{bmatrix} \mathbf{U}_0 & \mathbf{U}_1 \end{bmatrix} \end{cases} \tag{5.252}$$

Let $\mathbf{\Lambda}_{cl}$ be the diagonal matrix of the closed-loop eigenvalues:

$$\mathbf{\Lambda}_{cl} = \begin{bmatrix} \lambda_{K_1} & & \\ & \ddots & \\ & & \lambda_{K_n} \end{bmatrix} \tag{5.253}$$

As far as all eigenvectors $\underline{v}_{K_i}$ which compose each column of matrix $\mathbf{P}$ have been computed matrix $\mathbf{K}$ can be calculated by:

$$\mathbf{K} = \mathbf{Y}^{-1}\mathbf{U}_0^T \left(\mathbf{A} - \mathbf{P}\mathbf{\Lambda}_{cl}\mathbf{P}^{-1}\right) \tag{5.254}$$

Furthermore the sufficient condition for pole assignment reads as follows[16]:

$$\mathbf{U}_1^T\mathbf{P}\mathbf{\Lambda}_{cl}\mathbf{P}^{-1} = \mathbf{U}_1^T\mathbf{A} \Leftrightarrow \mathbf{U}_1^T\left(\mathbf{A}\mathbf{P} - \mathbf{P}\mathbf{\Lambda}_{cl}\right) = \mathbf{0} \tag{5.255}$$

To get those results we write the closed-loop state matrix as follows:

$$\mathbf{A} - \mathbf{B}\mathbf{K} = \mathbf{P}\mathbf{\Lambda}_{cl}\mathbf{P}^{-1} \tag{5.256}$$

Then pre-multiplying the preceding equation by $\mathbf{U}^T$, using the decomposition of $\mathbf{B}$ and the fact that $\mathbf{U}$ is an orthogonal matrix (i.e. $\mathbf{U}^T\mathbf{U} = \mathbb{I}$) we get:

$$\begin{aligned} \mathbf{U}^T\mathbf{P}\mathbf{\Lambda}_{cl}\mathbf{P}^{-1} &= \mathbf{U}^T\left(\mathbf{A} - \mathbf{B}\mathbf{K}\right) \\ &= \mathbf{U}^T\mathbf{A} - \mathbf{U}^T\mathbf{B}\mathbf{K} \\ &= \mathbf{U}^T\mathbf{A} - \begin{bmatrix} \mathbf{Y} \\ \mathbf{0} \end{bmatrix}\mathbf{K} \\ &= \begin{bmatrix} \mathbf{U}_0^T \\ \mathbf{U}_1^T \end{bmatrix}\mathbf{A} - \begin{bmatrix} \mathbf{Y} \\ \mathbf{0} \end{bmatrix}\mathbf{K} \end{aligned} \tag{5.257}$$

[16] J. Kautsky, N. K. Nichols, P. Van Dooren, Robust pole assignment in linear state feedback, International Journal of Control, Volume 41, 1985 - Issue 5

The *first* row of (5.257) reads:

$$\mathbf{U}_0^T\mathbf{P}\mathbf{\Lambda}_{cl}\mathbf{P}^{-1} = \mathbf{U}_0^T\mathbf{A} - \mathbf{Y}\mathbf{K} \Leftrightarrow \mathbf{K} = \mathbf{Y}^{-1}\mathbf{U}_0^T\left(\mathbf{A} - \mathbf{P}\mathbf{\Lambda}_{cl}\mathbf{P}^{-1}\right) \tag{5.258}$$

Furthermore the *second* row of (5.257) leads to the following necessary and sufficient condition for pole assignment:

$$\mathbf{U}_1^T\mathbf{P}\mathbf{\Lambda}_{cl}\mathbf{P}^{-1} = \mathbf{U}_1^T\mathbf{A} \Leftrightarrow \mathbf{U}_1^T\left(\mathbf{A}\mathbf{P} - \mathbf{P}\mathbf{\Lambda}_{cl}\right) = \mathbf{0} \tag{5.259}$$

This completes the proof. ■

It is worth noticing that if a closed-loop eigenvalue λ_{K_i} is complex, then its conjugate value, $\overline{\lambda}_{K_i}$ is also a closed-loop eigenvalue. In order to manipulate real matrices during the computations, eigenvectors $\underline{v}_{K_i}$ and $\overline{\underline{v}}_{K_i}$ corresponding to complex conjugate eigenvalues are replaced by $\mathrm{Re}(\underline{v}_{K_i})$ and $\mathrm{Im}(\underline{v}_{K_i})$ respectively:

$$\begin{bmatrix} \cdots & \underline{v}_{K_i} & \overline{\underline{v}}_{K_i} & \cdots \end{bmatrix} \rightarrow \begin{bmatrix} \cdots & \mathrm{Re}(\underline{v}_{K_i}) & \mathrm{Im}(\underline{v}_{K_i}) & \cdots \end{bmatrix} \tag{5.260}$$

Furthermore eigenvalues λ_{K_i} and $\overline{\lambda}_{K_i}$ in the diagonal matrix $\mathbf{\Lambda}_{cl}$ are replaced by $\mathrm{Re}(\lambda_{K_i})$ and $\mathrm{Im}(\lambda_{K_i})$ as follows:

$$\begin{bmatrix} \ddots & & & \\ & \lambda_{K_i} & & \\ & & \overline{\lambda}_{K_i} & \\ & & & \ddots \end{bmatrix} \rightarrow \begin{bmatrix} \ddots & & & \\ & \mathrm{Re}(\lambda_{K_i}) & \mathrm{Im}(\lambda_{K_i}) & \\ & -\mathrm{Im}(\lambda_{K_i}) & \mathrm{Re}(\lambda_{K_i}) & \\ & & & \ddots \end{bmatrix} \tag{5.261}$$

Indeed from the relations:

$$\begin{cases} \mathbf{A}_{cl}\underline{v}_{K_i} = \lambda_{K_i}\underline{v}_{K_i} \\ \mathbf{A}_{cl}\overline{\underline{v}}_{K_i} = \overline{\lambda}_{K_i}\overline{\underline{v}}_{K_i} \end{cases} \tag{5.262}$$

We get by adding and subtracting the preceding equations:

$$\begin{aligned} &\begin{cases} \mathbf{A}_{cl}\left(\underline{v}_{K_i} + \overline{\underline{v}}_{K_i}\right) = \lambda_{K_i}\underline{v}_{K_i} + \overline{\lambda}_{K_i}\overline{\underline{v}}_{K_i} \\ \mathbf{A}_{cl}\left(\underline{v}_{K_i} - \overline{\underline{v}}_{K_i}\right) = \lambda_{K_i}\underline{v}_{K_i} - \overline{\lambda}_{K_i}\overline{\underline{v}}_{K_i} \end{cases} \\ &\Rightarrow \begin{cases} \mathbf{A}_{cl}\,\mathrm{Re}(\underline{v}_{K_i}) = \mathrm{Re}(\lambda_{K_i}\underline{v}_{K_i}) \\ \mathbf{A}_{cl}\,\mathrm{Im}(\underline{v}_{K_i}) = \mathrm{Im}(\lambda_{K_i}\underline{v}_{K_i}) \end{cases} \\ &\Leftrightarrow \begin{cases} \mathbf{A}_{cl}\,\mathrm{Re}(\underline{v}_{K_i}) = \mathrm{Re}(\lambda_{K_i})\,\mathrm{Re}(\underline{v}_{K_i}) - \mathrm{Im}(\lambda_{K_i})\,\mathrm{Im}(\underline{v}_{K_i}) \\ \mathbf{A}_{cl}\,\mathrm{Im}(\underline{v}_{K_i}) = \mathrm{Re}(\lambda_{K_i})\,\mathrm{Im}(\underline{v}_{K_i}) + \mathrm{Im}(\lambda_{K_i})\,\mathrm{Re}(\underline{v}_{K_i}) \end{cases} \end{aligned} \tag{5.263}$$

That is:

$$\begin{aligned} &\mathbf{A}_{cl}\begin{bmatrix} \mathrm{Re}(\underline{v}_{K_i}) & \mathrm{Im}(\underline{v}_{K_i}) \end{bmatrix} \\ &\qquad = \begin{bmatrix} \mathrm{Re}(\underline{v}_{K_i}) & \mathrm{Im}(\underline{v}_{K_i}) \end{bmatrix}\begin{bmatrix} \mathrm{Re}(\lambda_{K_i}) & \mathrm{Im}(\lambda_{K_i}) \\ -\mathrm{Im}(\lambda_{K_i}) & \mathrm{Re}(\lambda_{K_i}) \end{bmatrix} \end{aligned} \tag{5.264}$$

5.9.3 Example

Following an example provided by A. Fossard[17] we consider the following system:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.265}$$

where:

$$\begin{cases} \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \mathbf{C} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix} \end{cases} \tag{5.266}$$

This system has $m = 2$ inputs, $n = 3$ states and $p = 2$ outputs and is both controllable and observable. We wish to find a state feedback matrix $\mathbf{K}$ such that the closed-loop eigenvalues are as follows:

$$\begin{cases} \lambda_{K_1} = -3 \\ \lambda_{K_2} = -3 + 4j \\ \lambda_{K_3} = \bar{\lambda}_{K_2} = -3 - 4j \end{cases} \tag{5.267}$$

Moreover it is desired that the first output $y_1(t)$ of $\underline{y}(t)$ is decoupled from the first mode λ_{K_1} whereas the second output $y_2(t)$ of $\underline{y}(t)$ is decoupled from the last two modes $\lambda_{K_2}, \lambda_{K_3}$.

The decoupling specifications leads to the following expression of the product $\mathbf{CP}$ where $*$ represents unspecified components:

$$\mathbf{CP} = \mathbf{C} \begin{bmatrix} \underline{v}_{K_1} & \underline{v}_{K_2} & \underline{v}_{K_3} \end{bmatrix} = \begin{bmatrix} 0 & * & * \\ * & 0 & 0 \end{bmatrix} \tag{5.268}$$

For the first row of matrix $\mathbf{C}$ we get:

$$\begin{bmatrix} 1 & 1 & -1 \end{bmatrix} \underline{v}_{K_1} = 0 \Leftrightarrow \underline{c}_1^T \underline{v}_{K_1} = 0 \tag{5.269}$$

And for the second row of matrix $\mathbf{C}$:

$$\begin{cases} \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \underline{v}_{K_2} = 0 \Leftrightarrow \underline{c}_2^T \underline{v}_{K_2} = 0 \\ \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \underline{v}_{K_3} = 0 \Leftrightarrow \underline{c}_2^T \underline{v}_{K_3} = 0 \end{cases} \tag{5.270}$$

[17] A. Fossard, Commande modale des systèmes dynamiques, notes de cours, Sup'Aéro, 1994

At the end of this step we define $n = 3$ matrices $\mathbf{S}_i$ as follows:

$$\begin{cases} \mathbf{S}(\lambda_{K_1}) = \begin{bmatrix} \mathbf{A} - \lambda_{K_1}\mathbb{I} & \mathbf{B} \\ \underline{c}_1^T & \mathbf{0} \end{bmatrix} = \left[\begin{array}{ccc|cc} 3 & 0 & 0 & 0 & 1 \\ 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & 3 & 0 & 1 \\ \hline 1 & 1 & -1 & 0 & 0 \end{array}\right] \\ \mathbf{S}(\lambda_{K_2}) = \begin{bmatrix} \mathbf{A} - \lambda_{K_2}\mathbb{I} & \mathbf{B} \\ \underline{c}_2^T & \mathbf{0} \end{bmatrix} = \left[\begin{array}{ccc|cc} 4-4j & 0 & 0 & 0 & 1 \\ 1 & 3-4j & 1 & 1 & 0 \\ 0 & 1 & 4-4j & 0 & 1 \\ \hline 1 & 1 & 0 & 0 & 0 \end{array}\right] \\ \mathbf{S}(\lambda_{K_3}) = \begin{bmatrix} \mathbf{A} - \lambda_{K_3}\mathbb{I} & \mathbf{B} \\ \underline{c}_2^T & \mathbf{0} \end{bmatrix} = \left[\begin{array}{ccc|cc} 4+4j & 0 & 0 & 0 & 1 \\ 1 & 3+4j & 1 & 1 & 0 \\ 0 & 1 & 4+4j & 0 & 1 \\ \hline 1 & 1 & 0 & 0 & 0 \end{array}\right] \end{cases} \tag{5.271}$$

Then we compute matrix $\mathcal{N}_{S_i}$ whose columns constitute a basis of the right kernel of $\mathbf{S}_i$:

$$\mathbf{S}_i \mathcal{N}_{S_i} = \mathbf{S}_i \begin{bmatrix} \mathbf{N}(\lambda_{K_i}) \\ \mathbf{M}(\lambda_{K_i}) \end{bmatrix} = \mathbf{0} \tag{5.272}$$

We get:

$$\begin{cases} \mathcal{N}_{S_1} = \begin{bmatrix} \mathbf{N}(\lambda_{K_1}) \\ \mathbf{M}(\lambda_{K_1}) \end{bmatrix} = \left[\begin{array}{c} -0.2581989 \\ 0 \\ -0.2581989 \\ \hline 0.5163978 \\ 0.7745967 \end{array}\right] \\ \mathcal{N}_{S_2} = \begin{bmatrix} \mathbf{N}(\lambda_{K_2}) \\ \mathbf{M}(\lambda_{K_2}) \end{bmatrix} = \left[\begin{array}{c} -0.0970284 - 0.0971143j \\ 0.0970284 + 0.0971143j \\ -0.0970177 - 0.1213821j \\ \hline -0.4854962 + 0.3152674j \\ 0.7765709 + 0.0003433j \end{array}\right] \\ \mathcal{N}_{S_3} = \begin{bmatrix} \mathbf{N}(\lambda_{K_3}) \\ \mathbf{M}(\lambda_{K_3}) \end{bmatrix} = \left[\begin{array}{c} -0.0970284 + 0.0971143j \\ 0.0970284 - 0.0971143j \\ -0.0970177 + 0.1213821j \\ \hline -0.4854962 - 0.3152674j \\ 0.7765709 - 0.0003433j \end{array}\right] \end{cases} \tag{5.273}$$

As far as each matrix $\mathcal{N}_{S_i}$ reduces here to be a column vector we set the non zero parameter vector $\underline{z}_i$ to 1; as a consequence vector $\underline{v}_{K_i} = \mathbf{N}(\lambda_{K_i})\underline{z}_i$ is set to $\mathbf{N}(\lambda_{K_i})$. Thus matrix eigenvector $\underline{v}_{K_i}$ is built from the $n = 3$ first rows of $\mathcal{N}_{S_i}$. In order to manipulate real matrices during the computations, eigenvectors $\underline{v}_{K_i}$ and $\overline{\underline{v}}_{K_i}$ corresponding to complex conjugate eigenvalues are

replaced by $\mathrm{Re}(\underline{v}_{K_i})$ and $\mathrm{Im}(\underline{v}_{K_i})$ respectively:

$$\begin{cases} \underline{v}_{K_1} = \mathbf{N}(\lambda_{K_1}) = \begin{bmatrix} -0.2581989 \\ 0 \\ -0.2581989 \end{bmatrix} \\ \underline{v}_{K_2} = \mathbf{N}(\lambda_{K_2}) = \begin{bmatrix} -0.0970284 \\ 0.0970284 \\ -0.0970177 \end{bmatrix} \\ \underline{v}_{K_3} = \mathbf{N}(\lambda_{K_3}) = \begin{bmatrix} -0.0971143 \\ 0.0971143 \\ -0.1213821 \end{bmatrix} \end{cases} \tag{5.274}$$

Eigenvectors $\underline{v}_{K_i}$ compose each column of matrix $\mathbf{P}$. We get:

$$\mathbf{P} = \begin{bmatrix} \underline{v}_{K_1} & \underline{v}_{K_2} & \underline{v}_{K_3} \end{bmatrix} = \begin{bmatrix} -0.2581989 & -0.0970284 & -0.0971143 \\ 0 & 0.0970284 & 0.0971143 \\ -0.2581989 & -0.0970177 & -0.1213821 \end{bmatrix} \tag{5.275}$$

Finally a singular value decomposition of $\mathbf{B}$ is performed:

$$\begin{aligned} \mathbf{B} &= \mathbf{U} \begin{bmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{bmatrix} \mathbf{V}^T \\ &= \begin{bmatrix} 0.7071068 & 0 & -0.7071068 \\ 0 & -1 & 0 \\ 0.7071068 & 0 & 0.7071068 \end{bmatrix} \begin{bmatrix} 1.4142136 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \end{aligned} \tag{5.276}$$

Then we define $\mathbf{Y} = \mathbf{\Sigma}\mathbf{V}^T$ and suitably split $\mathbf{U} = \begin{bmatrix} \mathbf{U}_0 & \mathbf{U}_1 \end{bmatrix}$ such that $\mathbf{U}_0$ has $m = 2$ columns:

$$\begin{cases} \mathbf{Y} = \mathbf{\Sigma}\mathbf{V}^T = \begin{bmatrix} 1.4142136 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1.4142136 \\ -1 & 0 \end{bmatrix} \\ \mathbf{U}_0 = \begin{bmatrix} 0.7071068 & 0 \\ 0 & -1 \\ 0.7071068 & 0 \end{bmatrix} \end{cases} \tag{5.277}$$

Let $\mathbf{\Lambda}_{cl}$ be the diagonal matrix of the closed-loop eigenvalues:

$$\mathbf{\Lambda}_{cl} = \begin{bmatrix} \lambda_{K_1} & & \\ & \lambda_{K_2} & \\ & & \lambda_{K_3} \end{bmatrix} = \begin{bmatrix} -3 & & \\ & -3+4j & \\ & & -3-4j \end{bmatrix} \tag{5.278}$$

Then state feedback gain $\mathbf{K}$ is calculated as follows:

$$\mathbf{K} = \mathbf{Y}^{-1}\mathbf{U}_0^T \left(\mathbf{A} - \mathbf{P}\mathbf{\Lambda}_{cl}\mathbf{P}^{-1}\right) = \begin{bmatrix} -31 & 7 & 33 \\ 35 & -5 & -32 \end{bmatrix} \tag{5.279}$$

We can check that the product $\mathbf{CP}$ satisfy the decoupling specifications. Indeed:

$$\begin{aligned} \mathbf{CP} &= \mathbf{C} \begin{bmatrix} \underline{v}_{K_1} & \underline{v}_{K_2} & \underline{v}_{K_3} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0.0970177 & 0.1213821 \\ -0.2581989 & 0 & 0 \end{bmatrix} := \begin{bmatrix} 0 & * & * \\ * & 0 & 0 \end{bmatrix} \end{aligned} \tag{5.280}$$

5.10 Sensitivity to additive uncertainties

5.10.1 Sensitivity analysis

Let $\mathbf{A}_{cl}$ be the closed-loop state matrix and assume that $\mathbf{A}_{cl}$ is a diagonalizable matrix:

$$\mathbf{A}_{cl} = \mathbf{A} - \mathbf{BK} \tag{5.281}$$

Let $\mathbf{P}$ be the matrix which is composed by the right eigenvectors $\underline{v}_{K_i}$ of $\mathbf{A}_{cl}$ corresponding to eigenvalue λ_{K_i} and $\mathbf{P}^{-1}$ be the matrix which is composed by the left eigenvectors $\underline{w}_{K_i}$ of $\mathbf{A}_{cl}$ corresponding to eigenvalue λ_{K_i}:

$$\begin{cases} \mathbf{P} = \begin{bmatrix} \underline{v}_{K_1} & \cdots & \underline{v}_{K_n} \end{bmatrix} \\ \\ \mathbf{P}^{-1} = \begin{bmatrix} \underline{w}_{K_1}^T \\ \vdots \\ \underline{w}_{K_n}^T \end{bmatrix} \end{cases} \tag{5.282}$$

Let $\mathbf{A}_{cl}+\Delta\mathbf{A}$ be the perturbed state matrix subject to additive uncertainties $\Delta\mathbf{A}$. Then the Bauer–Fike theorem[18] states that the variation $\Delta\lambda_{K_i}$ of the eigenvalues of the perturbed state matrix are bounded according to the following relation:

$$\max_i |\Delta\lambda_{K_i}| \leq \kappa(\mathbf{P}) \, \|\Delta\mathbf{A}\|_2 \tag{5.283}$$

where:

$$\mathbf{P}^{-1}\mathbf{A}_{cl}\mathbf{P} = \begin{bmatrix} \lambda_{K_1} & & \\ & \ddots & \\ & & \lambda_{K_n} \end{bmatrix} \tag{5.284}$$

Coefficient $\kappa(\mathbf{P})$ is called the condition number of matrix $\mathbf{P}$ and is defined as follows:

$$\kappa(\mathbf{P}) = \sqrt{\frac{\lambda_{max}\left(\mathbf{P}^T\mathbf{P}\right)}{\lambda_{min}\left(\mathbf{P}^T\mathbf{P}\right)}} \tag{5.285}$$

For a square invertible matrix $\mathbf{P}$ this reduces as follows:

$$\kappa(\mathbf{P}) = \|\mathbf{P}\|_2 \left\|\mathbf{P}^{-1}\right\|_2 \tag{5.286}$$

The induced matrix 2-norm $\|\mathbf{P}\|_2$ is defined as the largest singular value of $\mathbf{P}$, that is the root square of the largest eigenvalue of $\mathbf{P}^T\mathbf{P}$ (or $\mathbf{P}\mathbf{P}^T$).

According to (5.283), to guarantee a small variation of the assigned poles against possible perturbations $\Delta\mathbf{A}$, one has to achieve a small condition number $\kappa(\mathbf{P})$ of the eigenvector matrix.

To get this result we first rewrite the relation which links the eigenvalue λ_{K_i} and the corresponding right eigenvector $\underline{v}_{K_1}$:

$$\mathbf{A}_{cl}\underline{v}_{K_i} = \lambda_{K_i}\underline{v}_{K_i}, \quad i = 1, \cdots, n \tag{5.287}$$

[18] https://en.wikipedia.org/wiki/Bauer-Fike_theorem

Then the first order derivative of the preceding equation reads:

$$\Delta\mathbf{A}\underline{v}_{K_i} + \mathbf{A}_{cl}\Delta\underline{v}_{K_i} = \Delta\lambda_{K_i}\underline{v}_{K_i} + \lambda_{K_i}\Delta\underline{v}_{K_i} \tag{5.288}$$

On the other hand the relation which links the eigenvalue λ_{K_i} and the corresponding left eigenvector $\underline{v}_{K_1}$ is the following:

$$\underline{w}_{K_i}^T\mathbf{A}_{cl} = \lambda_{K_i}\underline{w}_{K_i}^T, \quad i = 1, \cdots, n \tag{5.289}$$

Pre-multiplying (5.288) by $\underline{w}_{K_i}^T$ and using (5.289) leads to the following expression of $d\lambda_{K_i}$:

$$\begin{aligned}
&\underline{w}_{K_i}^T\Delta\mathbf{A}\underline{v}_{K_i} + \underline{w}_{K_i}^T\mathbf{A}_{cl}\Delta\underline{v}_{K_i} = \Delta\lambda_{K_i}\underline{w}_{K_i}^T\underline{v}_{K_i} + \lambda_{K_i}\underline{w}_{K_i}^T\Delta\underline{v}_{K_i} \\
&\Leftrightarrow \underline{w}_{K_i}^T\Delta\mathbf{A}\underline{v}_{K_i} + \cancel{\lambda_{K_i}\underline{w}_{K_i}^T\Delta\underline{v}_{K_i}} = \Delta\lambda_{K_i}\underline{w}_{K_i}^T\underline{v}_{K_i} + \cancel{\lambda_{K_i}\underline{w}_{K_i}^T\Delta\underline{v}_{K_i}} \\
&\Leftrightarrow \Delta\lambda_{K_i} = \frac{\underline{w}_{K_i}^T\Delta\mathbf{A}\underline{v}_{K_i}}{\underline{w}_{K_i}^T\underline{v}_{K_i}}
\end{aligned} \tag{5.290}$$

As far as the left and right eigenvectors are normalized such that $\underline{w}_{K_i}^T\underline{v}_{K_i} = 1$ we get:

$$\Delta\lambda_{K_i} = \underline{w}_{K_i}^T\Delta\mathbf{A}\underline{v}_{K_i} \tag{5.291}$$

Be taking the norm of the preceding relation we finally obtain:

$$|\Delta\lambda_{K_i}| \leq \left\|\underline{v}_{K_i}\right\|_2 \left\|\underline{w}_{K_i}^T\right\|_2 \|\Delta\mathbf{A}\|_2 \tag{5.292}$$

From the fact that $\left\|\underline{v}_{K_i}\right\|_2 \left\|\underline{w}_{K_i}^T\right\|_2 \leq \|\mathbf{P}\|_2 \left\|\mathbf{P}^{-1}\right\|_2 \ \forall\ i$ we finally get:

$$\max_i |\Delta\lambda_{K_i}| \leq \kappa(\mathbf{P}) \|\Delta\mathbf{A}\|_2 \ \text{ where } \kappa(\mathbf{P}) = \|\mathbf{P}\|_2 \left\|\mathbf{P}^{-1}\right\|_2 \tag{5.293}$$

5.10.2 Robust placement method

We have seen that the variation of each closed-loop eigenvalue is bounded by the following relation:

$$|\Delta\lambda_{K_i}| \leq \left\|\underline{v}_{K_i}\right\|_2 \left\|\underline{w}_{K_i}^T\right\|_2 \|\Delta\mathbf{A}\|_2 \tag{5.294}$$

Thus in order to minimize the sensitivity of the assigned eigenvalues it could be worth to minimize the following criteria:

$$J = \sum_i J_i \ \text{ where } J_i = \left\|\underline{v}_{K_i}\right\|_2 \left\|\underline{w}_{K_i}^T\right\|_2 \tag{5.295}$$

As far as $\underline{w}_{K_i}^T\underline{v}_{K_i} = 1$ criteria J_i reads:

$$J_i = \left\|\underline{v}_{K_i}\right\|_2 \left\|\underline{w}_{K_i}^T\right\|_2 = \frac{1}{\cos(\underline{v}_{K_i}, \underline{w}_{K_i}^T)} \tag{5.296}$$

From the fact that $\underline{w}_{K_i}^T\underline{v}_{K_j} = 0\ \forall i \neq j$ we conclude that vector $\underline{w}_{K_i}^T$ is perpendicular to the subspace spanned by the vectors of matrix $\mathbf{V}_i$ where:

$$\mathbf{V}_i = \left[\ \underline{v}_{K_1} \quad \cdots \quad \underline{v}_{K_{i-1}} \quad \underline{v}_{K_{i+1}} \quad \cdots \quad \underline{v}_{K_n}\ \right] \tag{5.297}$$

Thus J_i can be interpreted as the inverse of the *sinus* of the angle between $\underline{v}_{K_i}$ and $\mathbf{V}_i$. Minimizing the sensitivity of the eigenvalues of $\mathbf{A}_{cl} = \mathbf{A} - \mathbf{BK}$ to perturbations can be done by choosing a set of eigenvectors $\underline{v}_{K_i}$ so that each is maximally orthogonal to the space spanned by the remaining vectors. In others words eigenvectors $\underline{v}_{K_i}$ are shaped such that they are as orthogonal as possible to the remaining eigenvectors, which consequently minimizes the condition number of $\kappa(\mathbf{P})$ where $\mathbf{P} = \begin{bmatrix} \underline{v}_{K_1} & \cdots & \underline{v}_{K_n} \end{bmatrix}$.

Unfortunately this method, known as *Kautsky et al.* method[19], cannot handle complex eigenvalues in its original form, due to the need to update two complex conjugate eigenvectors at the same time.

To overcome this difficulty we present hereafter the method proposed by *Byers et al.*[19]:

- Assuming that state matrix $\mathbf{A}$ is of dimension n and that $(\mathbf{A}, \mathbf{B})$ is controllable we define n closed-loop eigenvalues $\lambda_{K_1}, \cdots, \lambda_{K_n}$ and define n matrices $\mathbf{S}_i$ as follows:

$$\mathbf{S}_i = \begin{bmatrix} \mathbf{A} - \lambda_{K_i}\mathbb{I} & \mathbf{B} \end{bmatrix} \tag{5.298}$$

- Then compute matrix $\mathcal{N}_{S_i}$ whose columns constitute a basis of the right kernel of $\mathbf{S}_i$:

$$\mathbf{S}_i \mathcal{N}_{S_i} = \mathbf{0} \tag{5.299}$$

  For complex conjugate eigenvalues λ_{K_i} and $\overline{\lambda}_{K_i}$, the corresponding matrices $\mathcal{N}_{S_i}$ and $\overline{\mathcal{N}}_{S_i}$ are also complex conjugates. They are replaced by their real and imaginary part, $Re\,(\mathcal{N}_{S_i})$ and $Im\,(\mathcal{N}_{S_i})$, respectively.

  Matrix $\mathcal{N}_{S_i}$ is a $(n+m) \times s_i$ matrix, where $s_i = m$ unless λ_{K_i} is an uncontrollable mode of the pair $(\mathbf{A}, \mathbf{B})$ in which case $s_i > m$. In the following we will assume that none of the λ_{K_i}'s is an uncontrollable mode of the pair $(\mathbf{A}, \mathbf{B})$ and consequently $s_i = m$.

- We denote $\mathbf{Z}$ the following $n\,m \times n$ block diagonal free parameters matrix build from n blocks $\underline{z}(\lambda_{K_1})$ of size $m \times 1$.

$$\mathbf{Z} = \begin{bmatrix} \underline{z}(\lambda_{K_1}) & & \underline{0} \\ & \ddots & \\ \underline{0} & & \underline{z}(\lambda_{K_n}) \end{bmatrix} \tag{5.300}$$

  For complex conjugate eigenvalues λ_{K_i} and $\overline{\lambda}_{K_i}$, the corresponding free parameters matrices $\underline{z}(\lambda_{K_i})$ and $\underline{z}(\overline{\lambda}_{K_i})$ shall be chosen to be equal:

$$\underline{z}(\lambda_{K_i}) = \underline{z}(\overline{\lambda}_{K_i}) \tag{5.301}$$

[19] A. Pandey, R. Schmid, T. Nguyen, Y. Yang, V. Sima and A. L. Tits, Performance Survey of Robust Pole Placement Methods, 53rd IEEE Conference on Decision and Control, 2014. Los Angeles

– Let $\mathbf{R}(\mathbf{Z})$ be the following $(n+m)\times n$ matrix defined as the product between matrix $\left[\begin{array}{ccc}\mathcal{N}_{S_1} & \cdots & \mathcal{N}_{S_n}\end{array}\right]$ of size $(n+m)\times n\,m$ and the free parameters matrix $\mathbf{Z}$ of size $n\,m\times n$:

$$\mathbf{R}(\mathbf{Z}) = \left[\begin{array}{ccc}\mathcal{N}_{S_1} & \cdots & \mathcal{N}_{S_n}\end{array}\right]\times\mathbf{Z} \tag{5.302}$$

Matrix $\mathbf{R}(\mathbf{Z})$ is then compatibly partitioned between matrix $\mathbf{N}(\mathbf{Z})$ and matrix $\mathbf{M}(\mathbf{Z})$. Matrix $\mathbf{N}(\mathbf{Z})$ is a $n\times n$ matrix which is built from the n first rows of $\mathbf{R}(\mathbf{Z})$:

$$\mathbf{R}(\mathbf{Z}) = \left[\begin{array}{c}\mathbf{N}(\mathbf{Z})\\ \mathbf{M}(\mathbf{Z})\end{array}\right] \tag{5.303}$$

– Then *Schmid et al.*[20] have shown that for almost every choice of the parameter matrix $\mathbf{Z}$ the rank of matrix $\mathbf{N}(\mathbf{Z})$ is equal to n as well as the rank of matrix $\mathbf{Z}$. Furthermore the $m\times n$ gain matrix $\mathbf{K}$ such that the eigenvalues of $\mathbf{A}_{cl}=\mathbf{A}-\mathbf{B}\mathbf{K}$ read $(\lambda_{K_1},\cdots,\lambda_{K_n})$ is given by:

$$\mathbf{K} = -\mathbf{M}(\mathbf{Z})\mathbf{N}(\mathbf{Z})^{-1} \tag{5.304}$$

Last but not least we have:

$$\max_i |\Delta\lambda_{K_i}| \leq \kappa(\mathbf{N}(\mathbf{Z}))\,\|\Delta\mathbf{A}\|_2 \tag{5.305}$$

Coefficient $\kappa(\mathbf{N}(\mathbf{Z}))$ is called the condition number of matrix $\mathbf{N}(\mathbf{Z})$ and is defined as follows:

$$\kappa(\mathbf{N}(\mathbf{Z})) = \|\mathbf{N}(\mathbf{Z})\|_2\,\left\|\mathbf{N}(\mathbf{Z})^{-1}\right\|_2 \tag{5.306}$$

$\|\mathbf{N}(\mathbf{Z})\|_2$ is the induced matrix 2-norm and is defined as the largest singular value of $\mathbf{N}(\mathbf{Z})$; similarly $\|\Delta\mathbf{A}\|_2$ is the largest singular value of $\Delta\mathbf{A}$.

Consequently the free parameters matrix $\mathbf{Z}$ can be used to minimize the sensitivity of the closed-loop state matrix $\mathbf{A}_{cl}$ to additive uncertainties.

5.10.3 Lyapunov approach

Assume that the closed-loop dynamics reads as follows, where the state matrix $\mathbf{A}_{cl}$ is asymptotically stable:

$$\dot{\underline{x}}(t) = (\mathbf{A}_{cl}+\Delta\mathbf{A})\,\underline{x}(t) \tag{5.307}$$

Denoting $\sigma_{max}(\mathbf{X})$ the maximum singular value of $\mathbf{X}=\mathbf{X}^T\geq 0$ and $\sigma_{min}(\mathbf{X})$ the minimum singular value of $\mathbf{X}$, it is known[21] that the closed-loop remains asymptotically stable for all $\Delta\mathbf{A}$ such that:

$$\sigma_{max}(\Delta\mathbf{A}) \leq \frac{\sigma_{min}\left(\mathbf{Q}^{1/2}\right)}{\sigma_{max}\left(\mathbf{Q}^{-1/2}\mathbf{P}\right)} \tag{5.308}$$

[20] R. Schmid, P. Pandey, T. Nguyen, Robust Pole Placement With Moore's Algorithm, IEEE Trans. Automatic Control, 2014, 59(2), 500-505

[21] Horng-Giou Chen, Kuang-Wei Han, Improved quantitative measures of robustness for multivariable systems, IEEE Transactions on Automatic Control, 1994, Volume: 39, Issue: 4

where $\mathbf{Q} = \mathbf{Q}^T \geq 0$ and where $\mathbf{P} = \mathbf{P}^T > 0$ is the solution of the following Lyapunov equation:

$$\mathbf{A}_{cl}^T\mathbf{P} + \mathbf{P}\mathbf{A}_{cl} + 2\mathbf{Q} = \mathbf{0} \tag{5.309}$$

Thus, the smaller the two-norm of $\mathbf{P}$ is, the more robustly stable the closed-loop system will be to unstructured parametric variations.

Moreover, the bound is maximum when matrix $\mathbf{Q} = \mathbb{I}$ is assigned in the Lyapunov equation.

To get this result, let $V(\underline{x})$ the following candidate Lyapunov function where $\mathbf{P} = \mathbf{P}^T > 0$:

$$V(\underline{x}) = \underline{x}^T\mathbf{P}\underline{x} > 0 \;\forall\; \underline{x} \neq \underline{0} \tag{5.310}$$

The *sufficient* condition for stability reads:

$$\frac{dV(\underline{x})}{dt} \leq 0 \tag{5.311}$$

where:

$$\begin{aligned}
\tfrac{dV(\underline{x})}{dt} &= \tfrac{d\underline{x}}{dt}^T\mathbf{P}\underline{x} + \underline{x}^T\mathbf{P}\tfrac{d\underline{x}}{dt} \\
&= \left(\left(\mathbf{A}_{cl} + \Delta\mathbf{A}\right)\underline{x}\right)^T\mathbf{P}\underline{x} + \underline{x}^T\mathbf{P}\left(\left(\mathbf{A}_{cl} + \Delta\mathbf{A}\right)\underline{x}\right) \\
&= \underline{x}^T\left(\mathbf{A}_{cl} + \Delta\mathbf{A}\right)^T\mathbf{P}\underline{x} + \underline{x}^T\mathbf{P}\left(\mathbf{A}_{cl} + \Delta\mathbf{A}\right)\underline{x} \\
&= \underline{x}^T\left(\mathbf{A}_{cl}^T\mathbf{P} + \mathbf{P}\mathbf{A}_{cl}\right)\underline{x} + \underline{x}^T\left(\Delta\mathbf{A}^T\mathbf{P} + \mathbf{P}\Delta\mathbf{A}\right)\underline{x} \\
&= \underline{x}^T\left(\mathbf{A}_{cl}^T\mathbf{P} + \mathbf{P}\mathbf{A}_{cl}\right)\underline{x} + 2\,\underline{x}^T\Delta\mathbf{A}^T\mathbf{P}\underline{x}
\end{aligned} \tag{5.312}$$

Because $\mathbf{A}_{cl}$ is asymptotically stable, we have $\underline{x}^T\left(\mathbf{A}_{cl}^T\mathbf{P} + \mathbf{P}\mathbf{A}_{cl}\right)\underline{x} \leq 0$. This inequality is rewritten as equality (5.309) as follows:

$$\mathbf{A}_{cl}^T\mathbf{P} + \mathbf{P}\mathbf{A}_{cl} = -2\mathbf{Q} \tag{5.313}$$

Thus the sufficient condition for robust stability becomes:

$$\begin{aligned}
&-2\,\underline{x}^T\mathbf{Q}\underline{x} + 2\,\underline{x}^T\Delta\mathbf{A}^T\mathbf{P}\underline{x} \leq 0 \\
&\Leftrightarrow \underline{x}^T\Delta\mathbf{A}^T\mathbf{P}\underline{x} \leq \underline{x}^T\mathbf{Q}\underline{x}
\end{aligned} \tag{5.314}$$

Because $\mathbf{Q} = \mathbf{Q}^T \geq 0$, we have $\mathbf{Q} = \mathbf{Q}^{1/2}\mathbf{Q}^{1/2}$ where $\mathbf{Q}^{1/2} = \left(\mathbf{Q}^{1/2}\right)^T \geq 0$. Making the slight modification $\underline{x} = \left(\mathbf{Q}^{-1/2}\mathbf{Q}^{1/2}\right)\underline{x}$, the preceding inequality is equivalently rewritten as follows:

$$\underline{x}^T\Delta\mathbf{A}^T\mathbf{P}\underline{x} = \underline{x}^T\Delta\mathbf{A}^T\mathbf{P}\left(\mathbf{Q}^{-1/2}\mathbf{Q}^{1/2}\right)\underline{x} \leq \left(\mathbf{Q}^{1/2}\underline{x}\right)^T\left(\mathbf{Q}^{1/2}\underline{x}\right) \tag{5.315}$$

The preceding inequality is equivalently given by:

$$\left(\mathbf{Q}^{-1/2}\mathbf{P}\Delta\mathbf{A}\,\underline{x}\right)^T\left(\mathbf{Q}^{1/2}\underline{x}\right) \leq \left(\mathbf{Q}^{1/2}\underline{x}\right)^T\left(\mathbf{Q}^{1/2}\underline{x}\right) \tag{5.316}$$

Then a sufficient condition for stability can be given as:

$$\left\|\mathbf{Q}^{-1/2}\mathbf{P}\Delta\mathbf{A}\,\underline{x}\right\| \leq \left\|\mathbf{Q}^{1/2}\underline{x}\right\| \tag{5.317}$$

Knowing that:

$$\left\|\mathbf{Q}^{-1/2}\mathbf{P}\Delta\mathbf{A}\,\underline{x}\right\| \leq \sigma_{max}\left(\mathbf{Q}^{-1/2}\mathbf{P}\right)\|\Delta\mathbf{A}\,\underline{x}\| \tag{5.318}$$

and

$$\left\|\mathbf{Q}^{1/2}\underline{x}\right\| \geq \sigma_{min}\left(\mathbf{Q}^{1/2}\right)\|\underline{x}\| \tag{5.319}$$

Then (5.317) is sufficiently justified when the following inequality holds, which complete the proof:

$$\sigma_{max}\left(\mathbf{Q}^{-1/2}\mathbf{P}\right)\|\Delta\mathbf{A}\,\underline{x}\| \leq \sigma_{min}\left(\mathbf{Q}^{1/2}\right)\|\underline{x}\| \tag{5.320}$$

5.10.4 Eigenvalues derivatives

Assume that the closed-loop dynamics reads as follows:

$$\dot{\underline{x}}(t) = (\mathbf{A} - \mathbf{BKC})\,\underline{x}(t) \tag{5.321}$$

The eigenvalues λ_{K_i} and corresponding right and left eigenvectors, $\underline{v}_{K_i}$ and $\underline{w}_{K_i}$ respectively, are given by the solutions of the following equations:

$$\begin{cases} (-\lambda_{K_i}\mathbb{I} + \mathbf{A} - \mathbf{BKC})\,\underline{v}_{K_i} = \underline{0} \\ \underline{w}_{K_i}^T\,(-\lambda_{K_i}\mathbb{I} + \mathbf{A} - \mathbf{BKC}) = \underline{0} \end{cases} \tag{5.322}$$

Differentiating the first equation of (5.322) with respect to $\mathbf{K}$ reads:

$$\left(-\frac{\partial\lambda_{K_i}}{\partial\mathbf{K}}\mathbb{I} + \frac{\partial\,(\mathbf{A} - \mathbf{BKC})}{\partial\mathbf{K}}\right)\underline{v}_{K_i} + (-\lambda_{K_i}\mathbb{I} + \mathbf{A} - \mathbf{BKC})\,\frac{\partial\underline{v}_{K_i}}{\partial\mathbf{K}} = \underline{0} \tag{5.323}$$

This equation is now multiplied by $\underline{w}_{K_i}^T$ and the second term is then zero because of the second equation of (5.322):

$$\underline{w}_{K_i}^T\left(-\frac{\partial\lambda_{K_i}}{\partial\mathbf{K}}\mathbb{I} + \frac{\partial\,(\mathbf{A} - \mathbf{BKC})}{\partial\mathbf{K}}\right)\underline{v}_{K_i} = \underline{0} \tag{5.324}$$

Finally if λ_{K_i} is an eigenvalue of $\mathbf{A} - \mathbf{BKC}$ then it can be shown that[22]:

$$\frac{\partial\lambda_{K_i}}{\partial\mathbf{K}} = -\frac{\mathbf{B}^T\underline{w}_{K_i}\underline{v}_{K_i}^T\mathbf{C}^T}{\underline{w}_{K_i}^T\underline{v}_{K_i}} \tag{5.325}$$

where $\underline{w}_{K_i}$ and $\underline{v}_{K_i}$ are the row and column eigenvectors of $\mathbf{A} - \mathbf{BKC}$ corresponding to λ_{K_i}.

More generally, for differential equations of the form:

$$\mathbf{M}\ddot{\underline{x}} + \mathbf{C}\dot{\underline{x}} + \mathbf{D}\underline{x} = \underline{u} \tag{5.326}$$

the eigenvalues λ_{K_i} and corresponding right and left eigenvectors, $\underline{v}_{K_i}$ and $\underline{w}_{K_i}$ respectively, are given by the solutions of the following equations[23]:

$$\begin{cases} \left(\lambda_{K_i}^2\mathbf{M} + \mathbf{C}\lambda_{K_i} + \mathbf{D}\right)\underline{v}_{K_i} = \underline{0} \\ \underline{w}_{K_i}^T\left(\lambda_{K_i}^2\mathbf{M} + \mathbf{C}\lambda_{K_i} + \mathbf{D}\right) = \underline{0} \end{cases} \tag{5.327}$$

[22] H. Sirisena, S. Choi, Pole placement in prescribed regions of the complex plane using output feedback, IEEE Transactions on Automatic Control, 1975, Page(s):810 - 812

[23] Michael Ian Friswell and Sondipon Adhikari, Derivatives of complex eigenvectors using Nelson's method, AIAA Journal 38:2355-2357, January 2000, DOI:10.2514/3.14688

5.11 Input-output decoupling control

5.11.1 Necessary and sufficient condition

We consider the following state-space representation of a controllable and observable square system where $\underline{u}(t) \in \mathbb{R}^m$ denotes the control input, $\underline{x}(t) \in \mathbb{R}^n$ the state vector and $\underline{y}(t) \in \mathbb{R}^p$, $p = m$ the output vector (which has the same size that control input $\underline{u}(t)$):

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \quad \text{where } dim\left(\underline{y}(t)\right) = dim\left(\underline{u}(t)\right) \tag{5.328}$$

Let's assume that $\underline{u}(t)$ can be split into $\left[\begin{array}{cc} \underline{u}_1(t) & \underline{u}_2(t) \end{array}\right]^T$; similarly we assume that $\underline{y}(t)$ can be split into $\left[\begin{array}{cc} \underline{y}_1(t) & \underline{y}_2(t) \end{array}\right]^T$. Thus the state-space representation reads:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \left[\begin{array}{cc} \mathbf{B}_1 & \mathbf{B}_2 \end{array}\right] \left[\begin{array}{c} \underline{u}_1(t) \\ \underline{u}_2(t) \end{array}\right] \\ \left[\begin{array}{c} \underline{y}_1(t) \\ \underline{y}_2(t) \end{array}\right] = \left[\begin{array}{c} \mathbf{C}_1 \\ \mathbf{C}_2 \end{array}\right] \underline{x}(t) \end{cases} \tag{5.329}$$

The transfer function from $\underline{u}_1(t)$ to $\underline{y}_2(t)$ is obtained by setting $\underline{u}_2(t) = \underline{0}$ and assuming no initial condition. We get:

$$\mathbf{F}_{u_1 y_2}(s) = \mathbf{C}_2 \left(s\mathbb{I} - \mathbf{A}\right)^{-1} \mathbf{B}_1 \tag{5.330}$$

Thus input $\underline{u}_1(t)$ and output $\underline{y}_2(t)$ will be decoupled as soon as transfer function $\mathbf{F}_{u_1 y_2}(s)$ is null:

$$\mathbf{F}_{u_1 y_2}(s) = \mathbf{C}_2 \left(s\mathbb{I} - \mathbf{A}\right)^{-1} \mathbf{B}_1 = \mathbf{0} \tag{5.331}$$

This condition is equivalent to the existence of an invertible change of basis matrix $\mathbf{P}$ which defines a new state vector $\underline{\xi}(t)$ as follows:

$$\underline{x}(t) = \mathbf{P}\underline{\xi}(t) \Leftrightarrow \underline{\xi}(t) = \mathbf{P}^{-1}\underline{x}(t) \Rightarrow \begin{cases} \dot{\underline{\xi}}(t) = \widetilde{\mathbf{A}}\underline{\xi}(t) + \widetilde{\mathbf{B}}\underline{u}(t) \\ \underline{y}(t) = \widetilde{\mathbf{C}}\underline{\xi}(t) \end{cases} \tag{5.332}$$

where:

$$\begin{cases} \widetilde{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} := \left[\begin{array}{cc} \widetilde{\mathbf{A}}_{11} & \mathbf{0} \\ \widetilde{\mathbf{A}}_{21} & \widetilde{\mathbf{A}}_{22} \end{array}\right] \\ \widetilde{\mathbf{B}} = \mathbf{P}^{-1}\left[\begin{array}{cc} \mathbf{B}_1 & \mathbf{B}_2 \end{array}\right] := \left[\begin{array}{cc} \widetilde{\mathbf{B}}_1 & \widetilde{\mathbf{B}}_2 \end{array}\right] \text{ where } \widetilde{\mathbf{B}}_1 = \mathbf{P}^{-1}\mathbf{B}_1 = \left[\begin{array}{c} \mathbf{0} \\ \widetilde{\mathbf{B}}_{21} \end{array}\right] \\ \widetilde{\mathbf{C}} = \left[\begin{array}{c} \mathbf{C}_1 \\ \mathbf{C}_2 \end{array}\right] \mathbf{P} := \left[\begin{array}{c} \widetilde{\mathbf{C}}_1 \\ \widetilde{\mathbf{C}}_2 \end{array}\right] \text{ where } \widetilde{\mathbf{C}}_2 = \mathbf{C}_2\mathbf{P} = \left[\begin{array}{cc} \widetilde{\mathbf{C}}_{21} & \mathbf{0} \end{array}\right] \end{cases} \tag{5.333}$$

Indeed:

$$\begin{aligned}
\mathbf{C}_2\left(s\mathbb{I}-\mathbf{A}\right)^{-1}\mathbf{B}_1 &= \mathbf{C}_2\left(s\mathbf{P}\mathbf{P}^{-1}-\mathbf{A}\right)^{-1}\mathbf{B}_1 \\
&= \mathbf{C}_2\left(\mathbf{P}\left(s\mathbb{I}-\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right)\mathbf{P}^{-1}\right)^{-1}\mathbf{B}_1 \\
&= \mathbf{C}_2\mathbf{P}\left(s\mathbb{I}-\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right)^{-1}\mathbf{P}^{-1}\mathbf{B}_1 \\
&= \begin{bmatrix} \widetilde{\mathbf{C}}_{21} & \mathbf{0} \end{bmatrix}\left(s\mathbb{I}-\widetilde{\mathbf{A}}\right)^{-1}\begin{bmatrix} \mathbf{0} \\ \widetilde{\mathbf{B}}_{21} \end{bmatrix} \\
&= \begin{bmatrix} \widetilde{\mathbf{C}}_{21} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{0} \\ \left(s\mathbb{I}-\widetilde{\mathbf{A}}\right)^{-1}\widetilde{\mathbf{B}}_{21} \end{bmatrix} \\
&= \mathbf{0}
\end{aligned} \tag{5.334}$$

This completes the proof. ■

Moreover, from Neumann's theory, it is known that the inverse $\mathbf{M}^{-1}$ of a nonsingular matrix $\mathbf{M}$ has the following series expansion as soon as the spectral radius of the square matrix $\mathbb{I}-\mathbf{X}\mathbf{M}$, that is the maximum moduli of its eigenvalues, is lower than 1[24]:

$$\mathbf{M}^{-1} = \sum_{k=0}^{\infty}\left(\mathbb{I}-\mathbf{X}\mathbf{M}\right)^k\mathbf{X} \tag{5.335}$$

Setting $\mathbf{X}=\mathbb{I}$ and $\mathbf{M}=\mathbb{I}-\mathbf{A}$ we get:

$$\left(\mathbb{I}-\mathbf{A}\right)^{-1} = \sum_{k=0}^{\infty}\mathbf{A}^k \tag{5.336}$$

This relation can be related to the series expansion of $\left(s\mathbb{I}-\mathbf{A}\right)^{-1}$ as follows:

$$\left(s\mathbb{I}-\mathbf{A}\right)^{-1} = \frac{1}{s}\left(\mathbb{I}-\frac{\mathbf{A}}{s}\right)^{-1} = \frac{1}{s}\sum_{k=0}^{\infty}\left(\frac{\mathbf{A}}{s}\right)^k = \sum_{k=1}^{\infty}\frac{\mathbf{A}^{k-1}}{s^k} \tag{5.337}$$

Thus transfer function $\mathbf{F}_{u_1y_2}(s)$ reads:

$$\mathbf{F}_{u_1y_2}(s) = \mathbf{C}_2\left(s\mathbb{I}-\mathbf{A}\right)^{-1}\mathbf{B}_1 = \sum_{k=1}^{\infty}\mathbf{C}_2\frac{\mathbf{A}^{k-1}}{s^k}\mathbf{B}_1 \tag{5.338}$$

We conclude that transfer function $\mathbf{F}_{u_1y_2}(s)$ is null as soon as the following relation holds:

$$\mathbf{F}_{u_1y_2}(s) = \mathbf{0} \Leftrightarrow \mathbf{C}_2\mathbf{A}^k\mathbf{B}_1 = \mathbf{0} \qquad \forall k \geq 0 \tag{5.339}$$

Let $\mathbf{Q}_{c1}$ be the following controllability matrix:

$$\mathbf{Q}_{c1} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{A}\mathbf{B}_1 & \cdots & \mathbf{A}^{n-1}\mathbf{B}_1 \end{bmatrix} \tag{5.340}$$

Let $\mathbf{Q}_{o2}$ be the following observability matrix:

$$\mathbf{Q}_{o2} = \begin{bmatrix} \mathbf{C}_2 \\ \mathbf{C}_2\mathbf{A} \\ \vdots \\ \mathbf{C}_2\mathbf{A}^{n-1} \end{bmatrix} \tag{5.341}$$

[24] Joan-Josep Climent, Néstor Thome, Yimin Wei, A geometrical approach on generalized inverses by Neumann-type series, Linear Algebra and its Applications 332–334 (2001) 533–540

Then relation (5.339) with Cayley-Hamilton theorem indicates that a necessary and sufficient condition for $\underline{u}_1(t)$ to be decoupled from $\underline{y}_2(t)$ is that the controllable subspace of $(\mathbf{A}, \mathbf{B}_1)$ is contained in the unobservable subspace of $(\mathbf{A}, \mathbf{C}_2)$. Denoting $\ker(\mathbf{Q}_{o2})$ the kernel of $\mathbf{Q}_{o2}$ and by $\text{range}(\mathbf{Q}_{c1})$ the span of $\mathbf{Q}_{c1}$, we shall have[25]:

$$\text{range}(\mathbf{Q}_{c1}) \subset \ker(\mathbf{Q}_{o2}) \tag{5.342}$$

5.11.2 Relative degree

In this section, we consider the input-output decoupling of the following *square* system:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \quad \text{where } dim\left(\underline{y}(t)\right) = dim\left(\underline{u}(t)\right) \tag{5.343}$$

We will assume that the input vector $\underline{u}(t)$ and the output vector $\underline{y}(t)$ are of dimension m and we denote $\underline{c}_i^T$ the ith row of output matrix $\mathbf{C}$:

$$dim\left(\underline{y}(t)\right) = dim\left(\underline{u}(t)\right) = m \Rightarrow \mathbf{C} := \begin{bmatrix} \underline{c}_1^T \\ \vdots \\ \underline{c}_m^T \end{bmatrix} \tag{5.344}$$

Let $y_1(t)$ be the first component of $\underline{y}(t)$:

$$y_1(t) = \underline{c}_1^T \underline{x}(t) \tag{5.345}$$

The first time derivative of $y_1(t)$ reads:

$$\frac{d}{dt} y_1(t) = \underline{c}_1^T \dot{\underline{x}}(t) = \underline{c}_1^T \mathbf{A}\underline{x}(t) + \underline{c}_1^T \mathbf{B}\underline{u}(t) \tag{5.346}$$

We compute the time derivative of $y_1(t)$ until some components of the input vector $\underline{u}(t)$ explicitly appears in the time derivative. Let $r_1 \geq 1$ be the lowest order of the derivative such that $\underline{c}_1^T \mathbf{A}^{r_1-1}\mathbf{B} \neq \underline{0}$:

$$\frac{d^{r_1}}{dt^{r_1}} y_1(t) = \underline{c}_1^T \mathbf{A}^{r_1} \underline{x}(t) + \underline{c}_1^T \mathbf{A}^{r_1-1}\mathbf{B}\underline{u}(t) \text{ where } \underline{c}_1^T \mathbf{A}^{r_1-1}\mathbf{B} \neq \underline{0},\ r_1 \geq 1 \tag{5.347}$$

More generally, let $r_i \geq 1$ be the lowest order of the derivative of the ith component of the output vector $\underline{y}(t)$ such that some components of the input vector $\underline{u}(t)$ explicitly appears in the time derivative. Then for $i = 1, \cdots, m$, the corresponding r_i is such that:

$$\frac{d^{r_i}}{dt^{r_i}} y_i(t) = \underline{c}_i^T \mathbf{A}^{r_i} \underline{x}(t) + \underline{c}_i^T \mathbf{A}^{r_i-1}\mathbf{B}\underline{u}(t) \text{ where } \underline{c}_i^T \mathbf{A}^{r_i-1}\mathbf{B} \neq \underline{0},\ r_i \geq 1 \tag{5.348}$$

[25] Luigi Glielmo and Martin Corless, On output feedback control of singularly perturbed systems, Applied Mathematics and Computation Volume 217, Issue 3, 1 October 2010, Pages 1053-1070

Equations (5.347) and (5.348) are written in vector form for $i = 1, \cdots, m$ as follows:

$$\begin{bmatrix} y_1^{(r_1)}(t) \\ \vdots \\ y_m^{(r_m)}(t) \end{bmatrix} = \begin{bmatrix} \underline{c}_1^T \mathbf{A}^{r_1} \\ \vdots \\ \underline{c}_m^T \mathbf{A}^{r_m} \end{bmatrix} \underline{x}(t) + \begin{bmatrix} \underline{c}_1^T \mathbf{A}^{r_1-1} \\ \vdots \\ \underline{c}_m^T \mathbf{A}^{r_m-1} \end{bmatrix} \mathbf{B}\underline{u}(t) \tag{5.349}$$

or, equivalently:

$$\boxed{\begin{bmatrix} y_1^{(r_1)}(t) \\ \vdots \\ y_m^{(r_m)}(t) \end{bmatrix} = \mathbf{A}^* \underline{x}(t) + \mathbf{B}^* \underline{u}(t) \text{ where } \begin{cases} \mathbf{A}^* := \begin{bmatrix} \underline{c}_1^T \mathbf{A}^{r_1} \\ \vdots \\ \underline{c}_m^T \mathbf{A}^{r_m} \end{bmatrix} \\ \mathbf{B}^* := \begin{bmatrix} \underline{c}_1^T \mathbf{A}^{r_1-1} \\ \vdots \\ \underline{c}_m^T \mathbf{A}^{r_m-1} \end{bmatrix} \mathbf{B} \end{cases}} \tag{5.350}$$

Coefficients r_i are called relative degrees of output $y_i(t)$. For SISO systems, the relative degree corresponds to the difference between the degree of the denominator and the degree of the numerator of the transfer function.

5.11.3 Falb-Wolovich decoupling method

Falb-Wolovich decoupling method was published in 1967[26]. Assume that we desire to set the transfer function between $Y_i(s)$ and its commanded value $Y_{c_i}(s)$ as follows where λ_{i,r_i}, $i = 1, \cdots, r_i$ represents the r_i desired closed-loop poles of the transfer function:

$$\frac{Y_i(s)}{Y_{c_i}(s)} := \frac{(-\lambda_{i,1}) \cdots (-\lambda_{i,r_i})}{(s - \lambda_{i,1}) \cdots (s - \lambda_{i,r_i})} \quad \text{where } i = 1, \cdots, m \tag{5.351}$$

Once the closed-loop transfer function $\frac{Y_i(s)}{Y_{c_i}(s)}$ is developed we get:

$$\boxed{\frac{Y_i(s)}{Y_{c_i}(s)} = \frac{a_{i,0}}{s^{r_i} + a_{i,r_i-1}\, s^{r_i-1} + \cdots + a_{i,1}\, s + a_{i,0}} := \frac{a_{i,0}}{\phi_i(s)}} \tag{5.352}$$

Note that closed-loop transfer function $\frac{Y_i(s)}{Y_{c_i}(s)}$ has a unity static feedback gain;

$$\left.\frac{Y_i(s)}{Y_{c_i}(s)}\right|_{s=0} = 1 \tag{5.353}$$

Then we define matrix $\widetilde{\mathbf{A}}^*$ as follows:

$$\boxed{\begin{gathered} \widetilde{\mathbf{A}}^* = \begin{bmatrix} \underline{c}_1^T \phi_1(\mathbf{A}) \\ \vdots \\ \underline{c}_m^T \phi_m(\mathbf{A}) \end{bmatrix} \\ \text{where } \phi_i(\mathbf{A}) := \mathbf{A}^{r_i} + a_{i,r_i-1}\, \mathbf{A}^{r_i-1} + \cdots + a_{i,1}\, \mathbf{A} + a_{i,0}\, \mathbb{I} \end{gathered}} \tag{5.354}$$

[26] P. Falb, W. Wolovich, Decoupling in the design and synthesis of multivariable control systems, IEEE Transactions on Automatic Control, December 1967

Assuming that $\mathbf{B}^*$ in (5.350) is non-singular, control vector $\underline{u}(t)$ is chosen as follows:

$$\boxed{\underline{u}(t) = -\mathbf{K}\underline{x}(t) + \mathbf{H}\underline{r}(t) \text{ where } \begin{cases} \mathbf{K} := (\mathbf{B}^*)^{-1}\,\widetilde{\mathbf{A}}^* \\ \mathbf{H} := (\mathbf{B}^*)^{-1} \begin{bmatrix} a_{1,0} & & \\ & \ddots & \\ & & a_{m,0} \end{bmatrix} \end{cases}} \tag{5.355}$$

Then the closed-loop transfer function reads:

$$\boxed{\mathbf{G}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{BH} := \begin{bmatrix} \frac{a_{1,0}}{\phi_1(s)} & & \\ & \ddots & \\ & & \frac{a_{m,0}}{\phi_m(s)} \end{bmatrix}} \tag{5.356}$$

Indeed, coming back in the time domain, (5.352) reads:

$$\begin{aligned} & y_i^{(r_i)}(t) + a_{i,r_i-1}\, y_i^{(r_i-1)}(t) + \cdots + a_{i,1}\,\dot{y}_i(t) + a_{i,0}\, y_i(t) = a_{i,0}\, y_{i_c}(t) \\ \Leftrightarrow\ & y_i^{(r_i)}(t) = -a_{i,r_i-1}\, y_i^{(r_i-1)}(t) - \cdots - a_{i,1}\,\dot{y}_i(t) - a_{i,0}\, y_i(t) + a_{i,0}\, y_{i_c}(t) \end{aligned} \tag{5.357}$$

On the other hand, we have seen in (5.349) that the r_i^{th} time derivative of the ith output $y_i(t)$ reads:

$$y_i^{(r_i)}(t) = \underline{c}_i^T\mathbf{A}^{r_i}\underline{x}(t) + \underline{c}_i^T\mathbf{A}^{r_i-1}\mathbf{B}\underline{u}(t) \tag{5.358}$$

When the control $\underline{u}(t) = -\mathbf{K}\underline{x}(t) + \mathbf{H}\underline{r}(t)$ is applied we get:

$$y_i^{(r_i)}(t) = \left(\underline{c}_i^T\mathbf{A}^{r_i} - \underline{c}_i^T\mathbf{A}^{r_i-1}\mathbf{BK}\right)\underline{x}(t) + \underline{c}_i^T\mathbf{A}^{r_i-1}\mathbf{BH}\underline{r}(t) \tag{5.359}$$

Setting $\underline{r}(t) := \begin{bmatrix} y_{c_1}(t) \\ \vdots \\ y_{c_m}(t) \end{bmatrix}$, where $y_{c_i}(t)$ denotes the commanded value for the ith output $y_i(t)$, and comparing the targeted differential equation (5.357) with the actual differential equation (5.359), we conclude that decoupling control yields the following conditions[27]:

$$\begin{cases} a_{i,r_i-1}\, y_i^{(r_i-1)}(t) + \cdots + a_{i,1}\,\dot{y}_i(t) + a_{i,0}\, y_i(t) = \underline{c}_i^T\left(\mathbf{A}^{r_i-1}\mathbf{BK} - \mathbf{A}^{r_i}\right)\underline{x}(t) \\ \underline{c}_i^T\mathbf{A}^{r_i-1}\mathbf{BH} \;=\; \begin{bmatrix} 0 & \cdots & 0 & \underbrace{a_{i,0}}_{i^{th}\text{ column}} & 0 & \cdots & 0 \end{bmatrix} \end{cases} \tag{5.360}$$

Using the fact that $\underline{c}_i^T\mathbf{A}^{j-1}\mathbf{B} = \underline{0}\ \forall j < r_i$ we get:

$$\underline{c}_i^T\mathbf{A}^{j-1}\mathbf{B} = \underline{0}\ \forall j < r_i \Rightarrow y_i^{(j)}(t) = \underline{c}_i^T\mathbf{A}^j\underline{x}(t)\ \forall j < r_i \tag{5.361}$$

[27] Eric Ostertag, Mono- and Multivariable Control and Estimation, Springer, 2011

Thus the first equality of (5.360) reads:

$$\begin{aligned} a_{i,r_i-1}\,\underline{c}_i^T\mathbf{A}^{r_i-1}\underline{x}(t)+\cdots+a_{i,1}\,\underline{c}_i^T\mathbf{A}\underline{x}(t)+a_{i,0}\,\underline{c}_i^T\underline{x}(t)=\underline{c}_i^T\left(\mathbf{A}^{r_i-1}\mathbf{BK}-\mathbf{A}^{r_i}\right)\underline{x}(t)\\ \Leftrightarrow \underline{c}_i^T\left(a_{i,r_i-1}\mathbf{A}^{r_i-1}+\cdots+a_{i,1}\mathbf{A}+a_{i,0}\mathbb{I}\right)\underline{x}(t)=\underline{c}_i^T\left(\mathbf{A}^{r_i-1}\mathbf{BK}-\mathbf{A}^{r_i}\right)\underline{x}(t) \end{aligned} \tag{5.362}$$

This equality shall hold whatever the value of the state vector $\underline{x}(t)$. Thus we finally get:

$$\begin{aligned} &\underline{c}_i^T\left(a_{i,r_i-1}\mathbf{A}^{r_i-1}+\cdots+a_{i,1}\mathbf{A}+a_{i,0}\mathbb{I}\right)=\underline{c}_i^T\left(\mathbf{A}^{r_i-1}\mathbf{BK}-\mathbf{A}^{r_i}\right)\\ &\Leftrightarrow \underline{c}_i^T\left(\mathbf{A}^{r_i}+a_{i,r_i-1}\mathbf{A}^{r_i-1}+\cdots+a_{i,1}\mathbf{A}+a_{i,0}\mathbb{I}\right)=\underline{c}_i^T\mathbf{A}^{r_i-1}\mathbf{BK}\\ &\Leftrightarrow \underline{c}_i^T\phi_i(\mathbf{A})=\underline{c}_i^T\mathbf{A}^{r_i-1}\mathbf{BK}\\ &\text{where } \phi_i(\mathbf{A}):=\mathbf{A}^{r_i}+a_{i,r_i-1}\,\mathbf{A}^{r_i-1}+\cdots+a_{i,1}\,\mathbf{A}+a_{i,0}\,\mathbb{I} \end{aligned} \tag{5.363}$$

Applying the same process for each output component $y_i(t),\ i=1,\cdots,m$, and assuming that $\mathbf{B}^*:=\begin{bmatrix}\underline{c}_1^T\mathbf{A}^{r_1-1}\\ \vdots\\ \underline{c}_m^T\mathbf{A}^{r_m-1}\end{bmatrix}\mathbf{B}$ is non-singular, yields:

$$\begin{aligned} \widetilde{\mathbf{A}}^*:=\begin{bmatrix}\underline{c}_1^T\phi_1(\mathbf{A})\\ \vdots\\ \underline{c}_m^T\phi_m(\mathbf{A})\end{bmatrix}=\begin{bmatrix}\underline{c}_1^T\mathbf{A}^{r_1-1}\\ \vdots\\ \underline{c}_m^T\mathbf{A}^{r_m-1}\end{bmatrix}\mathbf{BK}:=\mathbf{B}^*\mathbf{K}\\ \Leftrightarrow \mathbf{K}:=(\mathbf{B}^*)^{-1}\,\widetilde{\mathbf{A}}^* \end{aligned} \tag{5.364}$$

Finally, let us focus on the second of (5.360). Applying the same constraint for each output component $y_i(t),\ i=1,\cdots,m$, and assuming again that $\mathbf{B}^*$ is non-singular, yields:

$$\begin{aligned} &\begin{bmatrix}\underline{c}_1^T\mathbf{A}^{r_1-1}\\ \vdots\\ \underline{c}_m^T\mathbf{A}^{r_m-1}\end{bmatrix}\mathbf{BH}=\begin{bmatrix}a_{1,0} & & \\ & \ddots & \\ & & a_{m,0}\end{bmatrix}\\ &\Leftrightarrow \mathbf{H}:=(\mathbf{B}^*)^{-1}\begin{bmatrix}a_{1,0} & & \\ & \ddots & \\ & & a_{m,0}\end{bmatrix} \end{aligned} \tag{5.365}$$

This completes the proof. ■

5.11.4 Zeros and zero dynamics

Note that $\det(s\mathbb{I}-\mathbf{A}+\mathbf{BK})$ has n eigenvalues whereas $r=\sum_{i=1}^m r_i$ eigenvalues have been set through the choice of polynomials $\phi_i(s),\ i=1,\cdots,m$. Thus $n-r$ eigenvalues have disappear due to pole-zero cancellation in the closed-loop transfer function $\mathbf{G}(s)$.

In the general case, let $\mathbf{B}$ be a $n\times m$ matrix and $\mathbf{C}$ be a $p\times n$ matrix. Then:

- The values of s which drops the normal rank of $\begin{bmatrix} s\mathbb{I}-\mathbf{A} & -\mathbf{B}\end{bmatrix}$ are called the input-decoupling zeros. Then it can be seen that input-decoupling zeros are also eigenvalues of state matrix $\mathbf{A}$. In addition, if λ_i is such an

input-decoupling zero, then $\text{rank}\left(\left[\begin{array}{cc} \lambda_i \mathbb{I} - \mathbf{A} & -\mathbf{B} \end{array}\right]\right) < n$ and thus there exists an $1 \times n$ vector $\underline{w}_i^T \neq \underline{0}$ such that:

$$\text{rank}\left(\left[\begin{array}{cc} \lambda_i \mathbb{I} - \mathbf{A} & -\mathbf{B} \end{array}\right]\right) < n \Leftrightarrow \exists \underline{w}_i \neq \underline{0} \; s.t. \; \underline{w}_i^T \left[\begin{array}{cc} \lambda_i \mathbb{I} - \mathbf{A} & -\mathbf{B} \end{array}\right] = \underline{0} \tag{5.366}$$

This however implies that λ_i is an uncontrollable eigenvalue of state matrix $\mathbf{A}$, and $\underline{w}_i^T$ the corresponding left eigenvector.

– Conversely, the values of s which drops the normal rank of $\left[\begin{array}{c} s\mathbb{I} - \mathbf{A} \\ \mathbf{C} \end{array}\right]$ are called the output-decoupling zeros. Then it can be seen that output-decoupling zeros are also eigenvalues of state matrix $\mathbf{A}$. In addition, if λ_i is such an output-decoupling zero, then $\text{rank}\left(\left[\begin{array}{c} \lambda_i \mathbb{I} - \mathbf{A} \\ \mathbf{C} \end{array}\right]\right) < n$ and thus there exists an $n \times 1$ vector $\underline{v}_i \neq \underline{0}$ such that:

$$\text{rank}\left(\left[\begin{array}{c} \lambda_i \mathbb{I} - \mathbf{A} \\ \mathbf{C} \end{array}\right]\right) < n \Leftrightarrow \exists \underline{v}_i \neq \underline{0} \; s.t. \; \left[\begin{array}{c} \lambda_i \mathbb{I} - \mathbf{A} \\ \mathbf{C} \end{array}\right] \underline{v}_i^T = \underline{0}, \tag{5.367}$$

This however implies that λ_i is an unobservable eigenvalue of state matrix $\mathbf{A}$, and $\underline{v}_i$ the corresponding right eigenvector.

There are eigenvalues of state matrix $\mathbf{A}$ which are both uncontrollable and unobservable. These uncontrollable and unobservable eigenvalues of state matrix $\mathbf{A}$ are called input-output-decoupling zeros.

In order to exhibit the dynamics of zeros, we consider the following change of coordinates:

$$\xi_{i,j} = \underline{c}_i^T \mathbf{A}^{j-1} \underline{x} \text{ where } \left\{ \begin{array}{l} i = 1, \cdots, m \\ j = 1, \cdots, r_i \end{array} \right. \tag{5.368}$$

Those $r = \sum_{i=1}^{m} r_i$ coordinates are completed by choosing $n - r$ coordinates z_k such that:

$$z_k = \underline{p}_k^T \underline{x} \text{ such that } \underline{p}_k^T \mathbf{B} = \underline{0} \text{ for } k = 1, \cdots, n - r \tag{5.369}$$

From the preceding relation, it is clean that $\underline{p}_k$ belongs to the kernel of $\mathbf{B}^T$:

$$\mathbf{B}^T \underline{p}_k = \underline{0} \text{ for } k = 1, \cdots, n - r \tag{5.370}$$

Then we get the following state vector $\left[\begin{array}{c} \underline{\xi} \\ \underline{z} \end{array}\right]$ of size n, which linearly depends on state vector $\underline{x}$ as follows where the coordinate transformation matrix $\mathbf{P}$ is

square and non-singular:

$$\begin{bmatrix} \underline{\xi} \\ \underline{z} \end{bmatrix} := \begin{bmatrix} \xi_{1,1} \\ \xi_{1,2} \\ \vdots \\ \xi_{1,r_1} \\ \hdashline \vdots \\ \hdashline \xi_{m,1} \\ \xi_{m,2} \\ \vdots \\ \xi_{m,r_m} \\ \hdashline z_1 \\ \vdots \\ z_{n-r} \end{bmatrix} = \begin{bmatrix} \underline{c}_1^T \\ \underline{c}_1^T \mathbf{A} \\ \vdots \\ \underline{c}_1^T \mathbf{A}^{r_1-1} \\ \hdashline \vdots \\ \hdashline \underline{c}_m^T \\ \underline{c}_m^T \mathbf{A} \\ \vdots \\ \underline{c}_m^T \mathbf{A}^{r_m-1} \\ \hdashline \underline{p}_1^T \\ \vdots \\ \underline{p}_{n-r}^T \end{bmatrix} \underline{x} := \mathbf{P}^{-1} \underline{x} \tag{5.371}$$

Then the state space representation (5.343) becomes :

$$\underline{x} = \mathbf{P} \begin{bmatrix} \underline{\xi} \\ \underline{z} \end{bmatrix} \Rightarrow \begin{cases} \begin{bmatrix} \dot{\underline{\xi}} \\ \dot{\underline{z}} \end{bmatrix} = \widetilde{\mathbf{A}} \begin{bmatrix} \underline{\xi} \\ \underline{z} \end{bmatrix} + \widetilde{\mathbf{B}} \underline{u} \\ \underline{y} = \widetilde{\mathbf{C}} \underline{\xi} \end{cases} \tag{5.372}$$

where:

$$\begin{cases} \widetilde{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} := \begin{bmatrix} \widetilde{\mathbf{A}}_{11} & \widetilde{\mathbf{A}}_{12} \\ \widetilde{\mathbf{A}}_{21} & \widetilde{\mathbf{A}}_{22} \end{bmatrix} \\ \widetilde{\mathbf{B}} = \mathbf{P}^{-1}\mathbf{B} = \begin{bmatrix} \widetilde{\mathbf{B}}_1 \\ \mathbf{0} \end{bmatrix} \\ \widetilde{\mathbf{C}} = \mathbf{C}\mathbf{P} = \begin{bmatrix} \widetilde{\mathbf{C}}_1 & \mathbf{0} \end{bmatrix} \end{cases} \tag{5.373}$$

Then we get:

$$\dot{\underline{z}} := \begin{bmatrix} \dot{z}_1 \\ \vdots \\ \dot{z}_{n-r} \end{bmatrix} = \widetilde{\mathbf{A}}_{21}\, \underline{\xi} + \widetilde{\mathbf{A}}_{22}\, \underline{z} \tag{5.374}$$

The zero dynamics is represented by the eigenvalues of the $(n-r) \times (n-r)$ square matrix $\widetilde{\mathbf{A}}_{22}$. It is worth noticing that the closed loop will be stable as soon as all the input-output-decoupling zeros (that are the eigenvalues of matrix $\widetilde{\mathbf{A}}_{22}$) have negative real part.

5.12 Frequency domain approach

5.12.1 Hsu-Chen theorem

We consider the following state-space representation of a controllable and observable system where $\underline{u}(t) \in \mathbb{R}^m$ denotes the control input, $\underline{x}(t) \in \mathbb{R}^n$ the state vector and $\underline{y}(t) \in \mathbb{R}^p$ the output vector:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.375}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$, $\mathbf{C} \in \mathbb{R}^{p\times n}$.
Assume the following state feedback control:

$$\underline{u}(t) = -\mathbf{K}\underline{x}(t) + \mathbf{H}\underline{r}(t) \tag{5.376}$$

Then the closed-loop system reads:

$$\begin{cases} \dot{\underline{x}}(t) = (\mathbf{A} - \mathbf{BK})\,\underline{x}(t) + \mathbf{BH}\underline{r}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) \end{cases} \tag{5.377}$$

The Hsu-Chen theorem[28] states that the following relation holds:

$$\boxed{\det(s\mathbb{I} - \mathbf{A} + \mathbf{BK}) = \det(s\mathbb{I} - \mathbf{A})\det(\mathbb{I} + \mathbf{K}\mathbf{\Phi}(s)\mathbf{B})} \tag{5.378}$$

where:

$$\mathbf{\Phi}(s) = (s\mathbb{I} - \mathbf{A})^{-1} \tag{5.379}$$

Hsu-Chen theorem is an application of the Sylvester's determinant theorem[29] which states that the following relation holds assuming that $\mathbf{M}_1$ is an $m \times n$ matrix and $\mathbf{M}_2$ an $n \times m$ matrix (so that $\mathbf{M}_1$ and $\mathbf{M}_2$ have dimensions allowing them to be multiplied in either order forming a square matrix):

$$\det(\mathbb{I}_m + \mathbf{M}_1\mathbf{M}_2) = \det(\mathbb{I}_n + \mathbf{M}_2\mathbf{M}_1) \tag{5.380}$$

Sylvester's determinant theorem may be proven using the Schur's formula, which is recalled hereafter:

$$\begin{aligned} \det\left(\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}\right) &= \det(\mathbf{A}_{22})\det(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}) \\ &= \det(\mathbf{A}_{11})\det(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}) \end{aligned} \tag{5.381}$$

Thus if $\mathbf{M} = \begin{bmatrix} \mathbb{I}_m & -\mathbf{M}_1 \\ \mathbf{M}_2 & \mathbb{I}_n \end{bmatrix}$, we get:

$$\begin{aligned} \det(\mathbf{M}) = \det\left(\begin{bmatrix} \mathbb{I}_m & -\mathbf{M}_1 \\ \mathbf{M}_2 & \mathbb{I}_n \end{bmatrix}\right) &= \det(\mathbb{I}_m + \mathbf{M}_1\mathbf{M}_2) \\ &= \det(\mathbb{I}_n + \mathbf{M}_2\mathbf{M}_1) \end{aligned} \tag{5.382}$$

In addition, for square matrices $\mathbf{M}_3$ and $\mathbf{M}_4$ of equal size, the determinant of the matrix product equals the product of their determinants:

$$\det(\mathbf{M}_3\mathbf{M}_4) = \det(\mathbf{M}_3)\det(\mathbf{M}_4) \tag{5.383}$$

Then we get:

$$\begin{aligned} \det(s\mathbb{I} - \mathbf{A} + \mathbf{BK}) &= \det\left((s\mathbb{I} - \mathbf{A})\left(\mathbb{I} + (s\mathbb{I} - \mathbf{A})^{-1}\mathbf{BK}\right)\right) \\ &= \det((s\mathbb{I} - \mathbf{A})(\mathbb{I} + \mathbf{\Phi}(s)\mathbf{BK})) \\ &= \det(s\mathbb{I} - \mathbf{A})\det(\mathbb{I} + \mathbf{\Phi}(s)\mathbf{BK}) \\ &= \det(s\mathbb{I} - \mathbf{A})\det(\mathbb{I} + \mathbf{K}\mathbf{\Phi}(s)\mathbf{B}) \end{aligned} \tag{5.384}$$

[28] Pole-shifting techniques for multivariable feedback systems, Retallack D.G., MacFarlane A.G.J., Proceedings of the Institution of Electrical Engineers, 1970
[29] https://en.wikipedia.org/wiki/Determinant

We finally get the Hsu-Chen theorem (5.378). ■

Furthermore, let $\mathbf{F}(s)$ be the transfer function of the plant and $\mathbf{G}(s)$ the closed-loop transfer function:

$$\begin{cases} \mathbf{F}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} \\ \mathbf{G}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{BH} \end{cases} \tag{5.385}$$

Then Wang[30] has shown that the closed-loop transfer function $\mathbf{G}(s)$ is related to the transfer function of the plant $\mathbf{F}(s)$ through the following relation:

$$\boxed{\mathbf{G}(s) = \mathbf{F}(s)\left(\mathbb{I} + \mathbf{K}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B}\right)^{-1}\mathbf{H}} \tag{5.386}$$

Indeed, we can write:

$$\begin{aligned} \mathbf{G}(s) &= \mathbf{C}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{BH} \\ &= \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\left(\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right) - \mathbf{BK}\right)\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{BH} \\ &= \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\left(\mathbb{I} - \mathbf{BK}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\right)\mathbf{BH} \\ &= \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\left(\mathbf{B} - \mathbf{BK}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{B}\right)\mathbf{H} \\ &= \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B}\left(\mathbb{I} - \mathbf{K}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{B}\right)\mathbf{H} \\ &= \mathbf{F}(s)\left(\mathbb{I} - \mathbf{K}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{B}\right)\mathbf{H} \end{aligned} \tag{5.387}$$

Then, notice that $\left(\mathbf{\Phi}(s) + \mathbf{BK}\right)^{-1}\left(\mathbb{I} + \mathbf{BK}\mathbf{\Phi}^{-1}(s)\right) = \mathbf{\Phi}^{-1}(s)$. Thus we get:

$$\left(\mathbb{I} - \mathbf{K}\left(\mathbf{\Phi}^{-1}(s) + \mathbf{BK}\right)^{-1}\mathbf{B}\right)\left(\mathbb{I} + \mathbf{K}\mathbf{\Phi}(s)\mathbf{B}\right) = \mathbb{I}$$
$$\text{where } \mathbf{\Phi}(s) = \left(s\mathbb{I} - \mathbf{A}\right)^{-1} \tag{5.388}$$

This completes the proof. ■

Furthermore from Equations (5.385) and (5.386) we can write:

$$\begin{cases} \mathbf{F}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} = \mathbf{C}\dfrac{\operatorname{adj}\left(s\mathbb{I} - \mathbf{A}\right)}{\det\left(s\mathbb{I} - \mathbf{A}\right)}\mathbf{B} \\ \mathbf{G}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{BH} = \mathbf{C}\dfrac{\operatorname{adj}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)}{\det\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)}\mathbf{BH} \\ \left(\mathbb{I} + \mathbf{K}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B}\right)^{-1} = \dfrac{\operatorname{adj}\left(\mathbb{I} + \mathbf{K}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B}\right)}{\det\left(\mathbb{I} + \mathbf{K}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B}\right)} \end{cases} \tag{5.389}$$

Thus Equation (5.386) reads as follows:

$$\mathbf{C}\frac{\operatorname{adj}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)}{\det\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)}\mathbf{BH} = \mathbf{C}\frac{\operatorname{adj}\left(s\mathbb{I} - \mathbf{A}\right)}{\det\left(s\mathbb{I} - \mathbf{A}\right)}\mathbf{B}\frac{\operatorname{adj}\left(\mathbb{I} + \mathbf{K}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B}\right)}{\det\left(\mathbb{I} + \mathbf{K}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B}\right)}\mathbf{H} \tag{5.390}$$

[30] Decoupling Control, Qing-Guo Wang, Springer, 2003

When identifying the denominator of this relation, we retrieve the Hsu-Chen theorem (5.378).

Finally Equation (5.387) indicates that closed-loop transfer function $\mathbf{G}(s)$ is equivalent to the transfer function of the plant $\mathbf{F}(s)$ serially driven by the following controller:

$$\mathbf{C}(s) = \left(\mathbb{I} - \mathbf{K}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{B}\right)\mathbf{H} \tag{5.391}$$

This controller can be represented by the following state space equations:

$$\begin{cases} \dot{\underline{x}}_c(t) = (\mathbf{A} - \mathbf{BK})\,\underline{x}_c(t) + \mathbf{BH}\underline{r}(t) \\ \underline{y}_c(t) = -\mathbf{K}\underline{x}_c(t) + \mathbf{H}\underline{r}(t) \end{cases} \tag{5.392}$$

5.12.2 Output-based controller

We consider the state-space representation (5.393) where the state vector $\underline{x}$ is of dimension n (that is the size of state matrix $\mathbf{A}$) and where $\underline{u}(t)$ denotes the input vector:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{5.393}$$

As in section 5.6.3, we will assume the following *output* feedback controller (5.134):

$$\begin{cases} \dot{\underline{x}}_c(t) = (\mathbf{A} - \mathbf{BK})\,\underline{x}_c(t) + \mathbf{L}\left(\underline{y}(t) - (\mathbf{C} - \mathbf{DK})\,\underline{x}_c(t)\right) \\ \underline{u}_o(t) = \mathbf{K}\underline{x}_c(t) \end{cases} \tag{5.394}$$

We have seen in section 5.6.3 that the characteristic polynomial $\chi_{A_{cl}}(s)$ of the closed-loop reads:

$$\begin{aligned} \chi_{A_{cl}}(s) &= \det\left(s\mathbb{I} - \mathbf{A}_{cl}\right) = \det\left(\begin{bmatrix} s\mathbb{I} - \mathbf{A} & \mathbf{BK} \\ -\mathbf{LC} & s\mathbb{I} - \mathbf{A} + \mathbf{BK} + \mathbf{LC} \end{bmatrix}\right) \\ &= \det\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)\det\left(s\mathbb{I} - \mathbf{A} + \mathbf{LC}\right) \end{aligned} \tag{5.395}$$

Furthermore from the Hsu-Chen theorem (5.378) the following relation holds:

$$\det\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right) = \det\left(s\mathbb{I} - \mathbf{A}\right)\det\left(\mathbb{I} + \mathbf{K}\mathbf{\Phi}(s)\mathbf{B}\right) \tag{5.396}$$

Let $D(s) = \det\left(s\mathbb{I} - \mathbf{A}\right)$ be the determinant of $\mathbf{\Phi}(s)$, that is the characteristic polynomial of the open-loop plant, and $\mathbf{N}_{ol}(s) = \operatorname{adj}\left(s\mathbb{I} - \mathbf{A}\right)\mathbf{B}$ be the adjugate matrix of $s\mathbb{I} - \mathbf{A}$ times matrix $\mathbf{B}$:

$$\mathbf{\Phi}(s)\mathbf{B} = \left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} = \frac{\operatorname{adj}\left(s\mathbb{I} - \mathbf{A}\right)\mathbf{B}}{\det\left(s\mathbb{I} - \mathbf{A}\right)} := \frac{\mathbf{N}_{ol}(s)}{D(s)} \tag{5.397}$$

Consequently (5.378) reads:

$$\det\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right) = \det\left(D(s)\mathbb{I} + \mathbf{K}\mathbf{N}_{ol}(s)\right) \tag{5.398}$$

As soon as λ_{K_i} is a desired closed-loop eigenvalue then the following relation holds:

$$\det\left(D(s)\mathbb{I}+\mathbf{K}\mathbf{N}_{ol}(s)\right)\big|_{s=\lambda_{K_i}}=0 \tag{5.399}$$

Consequently it is desired that matrix $D(s)\mathbb{I}+\mathbf{K}\mathbf{N}_{ol}(s)|_{s=\lambda_{K_i}}$ is singular. Following Shieh & al.[31], let $\underline{\omega}_i \neq \underline{0}$ be a vector of size $m \times 1$, where m is the number of columns of $\mathbf{B}$ (that is the size of the input vector $\underline{u}(t)$ of the plant), belonging to the kernel of matrix $D(s)\mathbb{I}+\mathbf{K}\mathbf{N}_{ol}(s)|_{s=\lambda_{K_i}}$. Thus changing s by λ_{K_i} we can write:

$$\left(D(\lambda_{K_i})\mathbb{I}+\mathbf{K}\mathbf{N}_{ol}(\lambda_{K_i})\right)\underline{\omega}_i=\underline{0} \tag{5.400}$$

Actually, vector $\underline{\omega}_i \neq \underline{0}$ can be used as a design parameter.

Alternatively, when making the parallel that λ_i is an eigenvalue of matrix $\mathbf{A}$ as soon as $\det\left(s\mathbb{I}-\mathbf{A}\right)|_{s=\lambda_i}=0$, we conclude that $D(\lambda_{Ki})$ is an eigenvalue of matrix $-\mathbf{K}\mathbf{N}_{ol}(\lambda_{Ki})$, and thus $\underline{\omega}_i$ is an eigenvector of $-\mathbf{K}\mathbf{N}_{ol}(\lambda_{Ki})$ corresponding to the eigenvalue $D(\lambda_{Ki})$. This remark can be extended to the output feedback case where $\mathbf{N}_{ol}(s)=\mathbf{C}\,\mathrm{adj}\left(s\mathbb{I}-\mathbf{A}\right)\mathbf{B}$.

In order to get gain $\mathbf{K}$ the preceding relation is rewritten as follows:

$$\mathbf{K}\mathbf{N}_{ol}(\lambda_{K_i})\underline{\omega}_i=-D(\lambda_{K_i})\underline{\omega}_i \tag{5.401}$$

This relation does not lead to the value of gain $\mathbf{K}$ as soon as $\mathbf{N}_{ol}(\lambda_{K_i})\underline{\omega}_i$ is a vector which is not invertible. Nevertheless assuming that n denotes the order of state matrix $\mathbf{A}$ we can apply this relation for the n desired closed-loop eigenvalues. We get:

$$\mathbf{K}\left[\begin{array}{ccc}\underline{v}_{K_1} & \cdots & \underline{v}_{K_n}\end{array}\right]=-\left[\begin{array}{ccc}\underline{p}_1 & \cdots & \underline{p}_n\end{array}\right] \tag{5.402}$$

Where vectors $\underline{v}_{K_i}$ and $\underline{p}_i$ are given by:

$$\left\{\begin{array}{l}\underline{v}_{K_i}=\mathbf{N}_{ol}(\lambda_{K_i})\,\underline{\omega}_i\\ \underline{p}_i=D(\lambda_{K_i})\,\underline{\omega}_i\end{array}\right. \tag{5.403}$$

We finally get the following static *state* feedback gain matrix $\mathbf{K}$:

$$\mathbf{K}=-\left[\begin{array}{ccc}\underline{p}_1 & \cdots & \underline{p}_n\end{array}\right]\left[\begin{array}{ccc}\underline{v}_{K_1} & \cdots & \underline{v}_{K_n}\end{array}\right]^{-1} \tag{5.404}$$

This is also the so-called Roppenecker's formula which has been developed in (5.74).

Note that for complex conjugate eigenvalues we have $\lambda_{K_i}=\bar{\lambda}_{K_{i+1}}$ and $\underline{v}_{K_i}=\bar{\underline{v}}_{K_{i+1}}$. Thus, assuming that $i=1$ and multiplying both of left sides of equation $\mathbf{K}\left[\begin{array}{cccc}\underline{v}_{K_1} & \underline{v}_{K_2} & \underline{v}_{K_3} & \cdots\end{array}\right]=-\left[\begin{array}{cccc}\underline{p}_1 & \underline{p}_2 & \underline{p}_3 & \cdots\end{array}\right]$ by the following nonsingular matrix $\mathbf{M}$ will not alter the calculation of $\mathbf{K}$:

$$\mathbf{M}=\left[\begin{array}{cc|c}0.5 & -0.5j & \mathbf{0}\\ 0.5 & 0.5j & \\ \hline \multicolumn{2}{c|}{\mathbf{0}} & \mathbb{I}\end{array}\right] \tag{5.405}$$

[31] L. S. Shieh, H. M. Dib and R. E. Yates, Sequential design of linear quadratic state regulators via the optimal root-locus techniques, IEE Proceedings D - Control Theory and Applications, vol. 135, no. 4, pp. 289-294, July 1988.

Then in this case (5.404) reads:

$$\mathbf{K} = -\left[\begin{array}{cccc} Re(\underline{p}_1) & Im(\underline{p}_1) & \underline{p}_3 & \cdots \end{array}\right]\left[\begin{array}{cccc} Re(\underline{v}_{K_1}) & Im(\underline{v}_{K_1}) & \underline{v}_{K_3} & \cdots \end{array}\right]^{-1} \tag{5.406}$$

A similar result can be achieved with $\det(s\mathbb{I} - \mathbf{A} + \mathbf{LC})$. Indeed from the Hsu-Chen theorem (5.378) we get:

$$\begin{aligned} \det(s\mathbb{I} - \mathbf{A} + \mathbf{LC}) &= \det\left((s\mathbb{I} - \mathbf{A} + \mathbf{LC})^T\right) \\ &= \det\left(s\mathbb{I} - \mathbf{A}^T + \mathbf{C}^T\mathbf{L}^T\right) \\ &= \det\left(s\mathbb{I} - \mathbf{A}^T\right)\det(\mathbb{I} + \mathbf{L}^T\mathbf{\Phi}'(s)\mathbf{C}^T) \end{aligned} \tag{5.407}$$

where:

$$\mathbf{\Phi}'(s) = \left(s\mathbb{I} - \mathbf{A}^T\right)^{-1} \tag{5.408}$$

Let $D(s) = \det\left(s\mathbb{I} - \mathbf{A}^T\right) = \det(s\mathbb{I} - \mathbf{A})$ be the determinant of $\mathbf{\Phi}'(s)$, that is the characteristic polynomial of the open-loop plant, and $\mathbf{N}'_{ol}(s) = \operatorname{adj}\left(s\mathbb{I} - \mathbf{A}^T\right)\mathbf{C}^T$ be the adjugate matrix of $s\mathbb{I} - \mathbf{A}^T$ times matrix $\mathbf{C}^T$:

$$\mathbf{\Phi}'(s)\mathbf{C}^T = \left(s\mathbb{I} - \mathbf{A}^T\right)^{-1}\mathbf{C}^T = \frac{\operatorname{adj}\left(s\mathbb{I} - \mathbf{A}^T\right)\mathbf{C}^T}{\det(s\mathbb{I} - \mathbf{A}^T)} := \frac{\mathbf{N}'_{ol}(s)}{D(s)} \tag{5.409}$$

Consequently (5.407) reads:

$$\det(s\mathbb{I} - \mathbf{A} + \mathbf{LC}) = \det\left(D(s)\mathbb{I} + \mathbf{L}^T\mathbf{N}'_{ol}(s)\right) \tag{5.410}$$

As soon as λ_{L_i} is a desired closed-loop eigenvalue then the following relation holds:

$$\det\left(D(s)\mathbb{I} + \mathbf{L}^T\mathbf{N}'_{ol}(s)\right)\Big|_{s=\lambda_{L_i}} = 0 \tag{5.411}$$

Consequently it is desired that matrix $D(s)\mathbb{I} + \mathbf{L}^T\mathbf{N}'_{ol}(s)\big|_{s=\lambda_{L_i}}$ is singular. Let $\underline{\omega}'_i \neq \underline{0}$ be a vector of size $p \times 1$, where p is the number of rows of $\mathbf{C}$ (that is the size of the output vector $\underline{y}(t)$ of the plant), belonging to the kernel of matrix $D(s)\mathbb{I} + \mathbf{L}^T\mathbf{N}'_{ol}(s)\big|_{s=\lambda_{L_i}}$. Thus changing s by λ_{L_i} we can write:

$$\left(D(\lambda_{L_i})\mathbb{I} + \mathbf{L}^T\mathbf{N}'_{ol}(\lambda_{L_i})\right)\underline{\omega}'_i = \underline{0} \tag{5.412}$$

Actually, vector $\underline{\omega}'_i \neq \underline{0}$ can be used as a design parameter.

In order to get gain $\mathbf{L}$ the preceding relation is rewritten as follows:

$$\mathbf{L}^T\mathbf{N}'_{ol}(\lambda_{L_i})\underline{\omega}'_i = -D(\lambda_{L_i})\underline{\omega}'_i \tag{5.413}$$

This relation does not lead to the value of gain $\mathbf{L}$ as soon as $\mathbf{N}_{ol}(\lambda_{L_i})\underline{\omega}'_i$ is a vector which is not invertible. Nevertheless assuming that n denotes the order of state matrix $\mathbf{A}$ we can apply this relation for the n desired closed-loop eigenvalues. We get:

$$\mathbf{L}^T\left[\begin{array}{ccc} \underline{v}_{L_1} & \cdots & \underline{v}_{L_n} \end{array}\right] = -\left[\begin{array}{ccc} \underline{p}'_1 & \cdots & \underline{p}'_n \end{array}\right] \tag{5.414}$$

Where vectors $\underline{v}_{L_i}$ and $\underline{p}'_i$ are given by:

$$\begin{cases} \underline{v}_{L_i} = \mathbf{N}_{ol}(\lambda_{L_i})\,\underline{\omega}'_i \\ \underline{p}'_i = D(\lambda_{L_i})\,\underline{\omega}'_i \end{cases} \tag{5.415}$$

We finally get the following static observer gain matrix $\mathbf{L}$, which is equivalent to (5.404) but dedicated to observer design:

$$\mathbf{L}^T = -\left[\ \underline{p}'_1 \quad \cdots \quad \underline{p}'_n\ \right]\left[\ \underline{v}_{L_1} \quad \cdots \quad \underline{v}_{L_n}\ \right]^{-1} \tag{5.416}$$

5.12.3 Invariance of (transmission) zeros under state feedback

We consider the following system:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{5.417}$$

We have seen in section 1.5 that the (transmission) zeros of the open-loop transfer function $\mathbf{F}(s) = \mathbf{C}\,(s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B} + \mathbf{D}$ are defined as the values of s such that the rank of the Rosenbrock's system matrix $\mathbf{R}(s) = \begin{bmatrix} s\mathbb{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$ is lower than its *normal* rank, meaning that the rank of $\mathbf{R}(s)$ drops.

Now, assume that we apply the following feedback on the plant:

$$\underline{u}(t) = -\mathbf{K}\underline{x}(t) + \mathbf{H}\underline{r}(t) \tag{5.418}$$

Thus the closed-loop state space realization reads:

$$\begin{cases} \dot{\underline{x}}(t) = (\mathbf{A} - \mathbf{BK})\,\underline{x}(t) + \mathbf{BH}\underline{r}(t) \\ \underline{y}(t) = (\mathbf{C} - \mathbf{DK})\,\underline{x}(t) + \mathbf{DH}\underline{r}(t) \end{cases} \tag{5.419}$$

Thus the closed-loop transfer function $\mathbf{G}(s)$ reads:

$$\mathbf{G}(s) = (\mathbf{C} - \mathbf{DK})\,(s\mathbb{I} - (\mathbf{A} - \mathbf{BK}))^{-1}\,\mathbf{BH} + \mathbf{DH} \tag{5.420}$$

The (transmission) zeros of the closed-loop transfer function $\mathbf{G}(s)$ are defined as the values of s such that the rank of the Rosenbrock's system matrix $\mathbf{R}_{cl}(s)$ is lower than its *normal* rank, meaning that the rank of $\mathbf{R}(s)$ drops, where $\mathbf{R}_{cl}(s)$ is defined as follows:

$$\mathbf{R}_{cl}(s) = \begin{bmatrix} s\mathbb{I} - (\mathbf{A} - \mathbf{BK}) & -\mathbf{BH} \\ (\mathbf{C} - \mathbf{DK}) & \mathbf{DH} \end{bmatrix} \tag{5.421}$$

The Rosenbrock's system matrix $\mathbf{R}_{cl}(s)$ can be re-written as follows:

$$\mathbf{R}_{cl}(s) = \begin{bmatrix} s\mathbb{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}\begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{K} & \mathbf{H} \end{bmatrix} = \mathbf{R}(s)\begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{K} & \mathbf{H} \end{bmatrix} \tag{5.422}$$

Thus, assuming that $\mathbf{R}(s)$ is a square matrix, we can write $\det(\mathbf{R}_{cl}(s)) = \det(\mathbf{R}(s))\det(\mathbf{H})$, from which it follows that the (transmission) zeros of a plant are invariant under state feedback.

5.12.4 Matrix fraction description of a transfer function

Right matrix fraction description

Given any transfer function $\mathbf{F}(s)$, the *right* matrix fraction description of $\mathbf{F}(s)$ reads as follows where $\mathbf{F}_{RN}(s)$ and $\mathbf{F}_{RD}(s)$ are *stable* transfer functions:

$$\mathbf{F}(s) = \mathbf{F}_{RN}(s)\mathbf{F}_{RD}^{-1}(s) \tag{5.423}$$

Stable transfer functions $\mathbf{F}_{RN}(s)$ and $\mathbf{F}_{RD}(s)$ are such that there exist two *stable* transfer functions $\mathbf{X}_R(s)$ and $\mathbf{Y}_R(s)$ such that the following relation holds:

$$\mathbf{Y}_R(s)\mathbf{F}_{RD}(s) + \mathbf{X}_R(s)\mathbf{F}_{RN}(s) = \mathbb{I} \qquad \forall\, s \tag{5.424}$$

To get a matrix fraction description, consider a minimal realization of transfer function $\mathbf{F}(s)$:

$$\mathbf{F}(s) = \mathbf{C}\,(s\mathbb{I} - \mathbf{A})^{-1}\,\mathbf{B} + \mathbf{D} := \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) \tag{5.425}$$

Then *stable* transfer functions $\mathbf{F}_{RN}(s)$ and $\mathbf{F}_{RD}(s)$ reads as follows where $\mathbf{K}$ is chosen such that $\mathbf{A} - \mathbf{BK}$ has all its eigenvalues with negative real part[32]:

$$\begin{cases} \mathbf{F}_{RN}(s) := \left(\begin{array}{c|c} \mathbf{A} - \mathbf{BK} & \mathbf{B} \\ \hline \mathbf{C} - \mathbf{DK} & \mathbf{D} \end{array}\right) \\ \mathbf{F}_{RD}(s) := \left(\begin{array}{c|c} \mathbf{A} - \mathbf{BK} & \mathbf{B} \\ \hline -\mathbf{K} & \mathbb{I} \end{array}\right) \end{cases} \tag{5.426}$$

Furthermore, *stable* transfer functions $\mathbf{X}_R(s)$ and $\mathbf{Y}_R(s)$ reads as follows where $\mathbf{L}$ is chosen such that $\mathbf{A} - \mathbf{LC}$ has all its eigenvalues with negative real part:

$$\begin{cases} \mathbf{X}_R(s) := \left(\begin{array}{c|c} \mathbf{A} - \mathbf{LC} & \mathbf{L} \\ \hline -\mathbf{K} & \mathbf{0} \end{array}\right) \\ \mathbf{Y}_R(s) := \left(\begin{array}{c|c} \mathbf{A} - \mathbf{LC} & \mathbf{B} - \mathbf{LD} \\ \hline \mathbf{K} & \mathbb{I} \end{array}\right) \end{cases} \tag{5.427}$$

If $\mathbf{A}$ is a $n \times n$ matrix, $\mathbf{B}$ a $n \times m$ matrix and $\mathbf{C}$ a $p \times n$ matrix, then $\mathbf{F}_{RN}(s)$ is a $p \times m$ matrix and $\mathbf{F}_{RD}(s)$ is a *square* $m \times m$ matrix. In addition, $\mathbf{X}_R(s)$ is a $m \times p$ matrix and $\mathbf{Y}_R(s)$ is a *square* $m \times m$ matrix.

Left matrix fraction description

Similarly the *left* matrix fraction description of $\mathbf{F}(s)$ reads as follows where $\mathbf{F}_{LN}(s)$ and $\mathbf{F}_{LD}(s)$ are *stable* transfer functions:

$$\mathbf{F}(s) = \mathbf{F}_{LD}^{-1}(s)\mathbf{F}_{LN}(s) \tag{5.428}$$

Stable transfer functions $\mathbf{F}_{LN}(s)$ and $\mathbf{F}_{LD}(s)$ are such that there exist two *stable* transfer functions $\mathbf{X}_L(s)$ and $\mathbf{Y}_L(s)$ such that the following relation holds:

$$\mathbf{F}_{LD}(s)\mathbf{Y}_L(s) - \mathbf{F}_{LN}(s)\mathbf{X}_L(s) = \mathbb{I} \qquad \forall\, s \tag{5.429}$$

[32] Fortuna L., Frasca M., Optimal and Robust Control, CRC Press, 2012

It can be shown that *stable* transfer functions $\mathbf{F}_{LN}(s)$ and $\mathbf{F}_{LD}(s)$ reads as follows where $\mathbf{K}$ is chosen such that $\mathbf{A}-\mathbf{BK}$ has all its eigenvalues with negative real part:

$$\begin{cases} \mathbf{F}_{LN}(s) := \left(\begin{array}{c|c} \mathbf{A}-\mathbf{LC} & \mathbf{B}-\mathbf{LD} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) \\ \mathbf{F}_{LD}(s) := \left(\begin{array}{c|c} \mathbf{A}-\mathbf{LC} & \mathbf{L} \\ \hline -\mathbf{C} & \mathbb{I} \end{array}\right) \end{cases} \tag{5.430}$$

Furthermore, *stable* transfer functions $\mathbf{X}_L(s)$ and $\mathbf{Y}_L(s)$ reads as follows where $\mathbf{L}$ is chosen such that $\mathbf{A}-\mathbf{LC}$ has all its eigenvalues with negative real part:

$$\begin{cases} \mathbf{X}_L(s) := \left(\begin{array}{c|c} \mathbf{A}-\mathbf{BK} & \mathbf{L} \\ \hline -\mathbf{K} & \mathbf{0} \end{array}\right) \\ \mathbf{Y}_L(s) := \left(\begin{array}{c|c} \mathbf{A}-\mathbf{BK} & \mathbf{L} \\ \hline \mathbf{C}-\mathbf{DK} & \mathbb{I} \end{array}\right) \end{cases} \tag{5.431}$$

The preceding relations can be rewritten in a shorter form as follows[33]:

$$\begin{bmatrix} \mathbf{F}_{RD}(s) & \mathbf{X}_L(s) \\ \mathbf{F}_{RN}(s) & \mathbf{Y}_L(s) \end{bmatrix} = \begin{bmatrix} -\mathbf{K} \\ \mathbf{C}_K \end{bmatrix} (s\mathbb{I}-\mathbf{A}_K)^{-1} \begin{bmatrix} \mathbf{B} & \mathbf{L} \end{bmatrix} + \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ \mathbf{D} & \mathbb{I} \end{bmatrix} \tag{5.432}$$

and:

$$\begin{bmatrix} \mathbf{Y}_R(s) & -\mathbf{X}_R(s) \\ -\mathbf{F}_{LN}(s) & \mathbf{F}_{LD}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{K} \\ -\mathbf{C} \end{bmatrix} (s\mathbb{I}-\mathbf{A}_L)^{-1} \begin{bmatrix} \mathbf{B}_L & \mathbf{L} \end{bmatrix} + \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{D} & \mathbb{I} \end{bmatrix} \tag{5.433}$$

where:

$$\begin{cases} \mathbf{C}_K = \mathbf{C}-\mathbf{DK} \\ \mathbf{A}_K = \mathbf{A}-\mathbf{BK} \\ \mathbf{A}_L = \mathbf{A}-\mathbf{LC} \\ \mathbf{B}_L = \mathbf{B}-\mathbf{LD} \end{cases} \tag{5.434}$$

The above functions constitute a doubly coprime factorization of $\mathbf{F}(s)$, meaning that:

$$\begin{aligned} &\begin{bmatrix} \mathbf{Y}_R(s) & -\mathbf{X}_R(s) \\ -\mathbf{F}_{LN}(s) & \mathbf{F}_{LD}(s) \end{bmatrix} \begin{bmatrix} \mathbf{F}_{RD}(s) & \mathbf{X}_L(s) \\ \mathbf{F}_{RN}(s) & \mathbf{Y}_L(s) \end{bmatrix} = \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \\ \Leftrightarrow &\begin{bmatrix} \mathbf{F}_{RD}(s) & \mathbf{X}_L(s) \\ \mathbf{F}_{RN}(s) & \mathbf{Y}_L(s) \end{bmatrix} \begin{bmatrix} \mathbf{Y}_R(s) & -\mathbf{X}_R(s) \\ -\mathbf{F}_{LN}(s) & \mathbf{F}_{LD}(s) \end{bmatrix} = \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \end{aligned} \tag{5.435}$$

5.12.5 Youla parameterization of all stabilizing controllers

We consider the model reference feedback loop shown in Figure 5.8, where $\mathbf{F}(s)$, $\mathbf{C}(s)$ and $\mathbf{C}_{pf}(s)$ are the plant, feedback controller and prefilter, respectively[34].

Let:

$$\mathbf{F}(s) = \mathbf{F}_{RN}(s)\mathbf{F}_{RD}^{-1}(s) = \mathbf{F}_{LD}^{-1}(s)\mathbf{F}_{LN}(s) \tag{5.436}$$

[33] Ruth Curtain, George Weiss and Martin Weiss, Coprime Factorization for Regular Linear Systems, Automatica. Vol. 32. No. 11, pp. 1519-1531, 1996

[34] Two-degree-of-freedom Optimal Flight Control System Design, Chih - Min Lin & Jiann - Min Wu, Proceedings of the 35th Conference on Decision and Control, Kobe, Japan December 1996

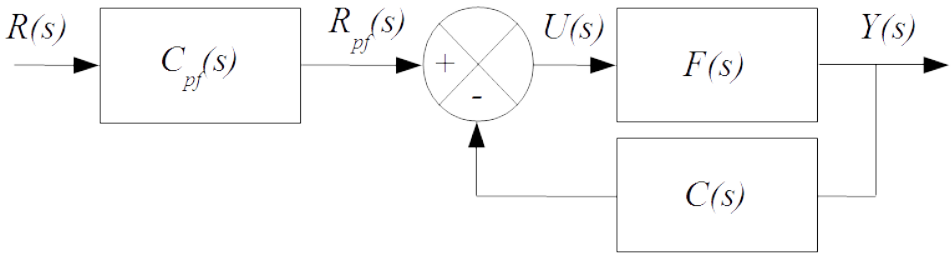

Figure 5.8: Feedback loop with two-degree of freedom

where the pairs $(\mathbf{F}_{RN}(s), \mathbf{F}_{RD}(s))$ and $(\mathbf{F}_{LN}(s), \mathbf{F}_{LD}(s))$ constitute any left and right coprime, proper, stable, rational decomposition of $\mathbf{F}(s)$. We have seen in (5.426) and (5.430) how to get $\mathbf{F}_{RN}(s)$, $\mathbf{F}_{RD}(s)$, $\mathbf{F}_{LN}(s)$, $\mathbf{F}_{LD}(s)$ from a minimal realization of transfer function $\mathbf{F}(s)$.

Then Youla & al.[35] have shown in 1976 that the class of controllers that make the closed-loop system internally stable is given as follows where $\mathbf{Q}(s)$ and $\mathbf{Q}_{pf}(s)$ are arbitrary compatibly dimensioned real rational matrices analytic in $Re(s) \geq 0$ (i.e. stable):

$$\begin{aligned}
\mathbf{C}(s) &= (\mathbf{Y}_R(s) + \mathbf{Q}(s)\mathbf{F}_{LN}(s))^{-1} (\mathbf{X}_R(s) + \mathbf{Q}(s)\mathbf{F}_{LD}(s)) \\
&= (\mathbf{X}_L(s) + \mathbf{F}_{RD}(s)\mathbf{Q}(s)) (\mathbf{Y}_L(s) + \mathbf{F}_{RN}(s)\mathbf{Q}(s))^{-1} \\
&\qquad \text{and} \\
\mathbf{C}_{pf}(s) &= (\mathbf{Y}_R(s) + \mathbf{Q}(s)\mathbf{F}_{LN}(s))^{-1} \mathbf{Q}_{pf}(s) \\
&= \mathbf{Q}_{pf}(s) (\mathbf{Y}_L(s) + \mathbf{F}_{RN}(s)\mathbf{Q}(s))^{-1}
\end{aligned} \tag{5.437}$$

5.12.6 Frequency domain representation of state feedback loop

The material of this section comes from the work of P. Hippe[36].

We consider the following linear, time-invariant MIMO systems where $\underline{x}(t) \in \mathbb{R}^n$, $\underline{y}(t) \in \mathbb{R}^p$ and $\underline{u}(t) \in \mathbb{R}^m$:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{z}(t) = \mathbf{N}\underline{x}(t) \end{cases} \tag{5.438}$$

Figure 5.9 shows the linear state feedback loop in the time domain where:

$$\underline{u}(t) = -\underline{u}_K(t) + \underline{r}(t) = -\mathbf{K}\underline{x}(t) + \underline{r}(t) \tag{5.439}$$

Reference input $\underline{r}(t) \in \mathbb{R}^p$ shall be multipliable by an appropriate pre-filter gain $\mathbf{H}$ for zero steady-state tracking error, but it will have no consequences on the following results.

Based on the right *polynomial* matrix fraction description (see Section C.1.6) of the open loop transfer function $\mathbf{F}(s)$ which reads $\mathbf{F}(s) = \mathbf{F}_{RN}(s)\mathbf{F}_{RD}^{-1}(s)$,

[35] Youla, D. C., Bongiomo, J. J., Jr., and Jabr, H. A., Modem Wiener-Hopf design of optimal controllers; Part 1: the single-input-output case, IEEE Transaction on Automatic Control, vol. AC-21, Feb. 1976, pp.3-13

[36] Hippe P., Parameterization of the full-order compensator in the frequency domain, International Journal of Control, Volume 48- Issue 4, 1988

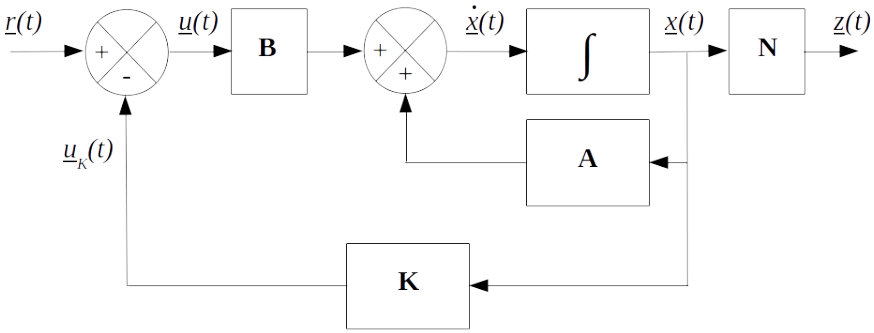

Figure 5.9: State feedback loop in the time domain

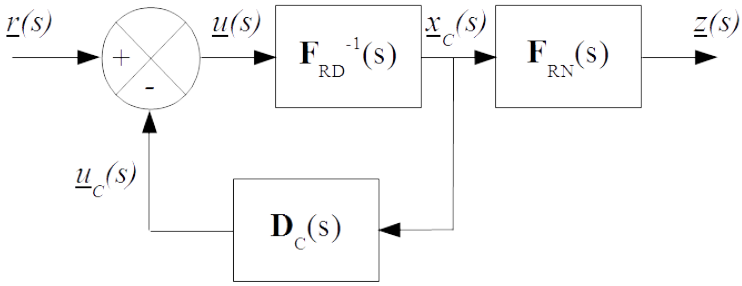

Figure 5.10: State feedback loop in the frequency domain

the state feedback loop in the time domain in Figure 5.9 can be equivalently represented by the state feedback loop in the frequency domain in Figure 5.10 where $\mathbf{D}_C(s)$ is a $m \times m$ *polynomial* matrix.

From Figure 5.9, we get the following relation between $\underline{u}_K(t)$ and $\underline{u}(t)$:

$$\underline{u}_K(s) = \mathbf{K}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B}\underline{u}(s) \tag{5.440}$$

On the other hand, from Figure 5.10, we get the following relation between $\underline{u}_C(t)$ and $\underline{u}(t)$:

$$\underline{u}_C(s) = \mathbf{D}_C(s)\mathbf{F}_{RD}^{-1}(s)\underline{u}(s) \tag{5.441}$$

Because both control loops exhibit the same input-output behavior, the following relation holds:

$$\underline{u}_K(s) = \underline{u}_C(s) \Rightarrow \mathbf{K}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} = \mathbf{D}_C(s)\mathbf{F}_{RD}^{-1}(s) \tag{5.442}$$

Similarly, from Figure 5.9 and Figure 5.10, we get the following expressions for $\underline{x}(s)$ and $\underline{x}_C(s)$:

$$\begin{aligned}
&s\underline{x}(s) = \mathbf{A}\underline{x}(s) + \mathbf{B}\left(\underline{r}(s) - \mathbf{K}\underline{x}(s)\right)\\
&\qquad \Leftrightarrow \underline{x}(s) = \left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{B}\underline{r}(s)\\
&\underline{x}_C(s) = \mathbf{F}_{RD}^{-1}(s)\left(\underline{r}(s) - \mathbf{D}_C(s)\underline{x}_C(s)\right)\\
&\qquad \Leftrightarrow \left(\mathbb{I} + \mathbf{F}_{RD}^{-1}(s)\mathbf{D}_C(s)\right)\underline{x}_C(s) = \mathbf{F}_{RD}^{-1}(s)\underline{r}(s)\\
&\qquad \Leftrightarrow \left(\mathbf{F}_{RD}(s) + \mathbf{D}_C(s)\right)\underline{x}_C(s) = \underline{r}(s)\\
&\qquad \Leftrightarrow \underline{x}_C(s) = \left(\mathbf{F}_{RD}(s) + \mathbf{D}_C(s)\right)^{-1}\underline{r}(s)
\end{aligned} \tag{5.443}$$

Then, by identifying the output $\underline{z}(s)$ in Figure 5.9 and Figure 5.10, we get:

$$\begin{cases} \underline{z}(s) = \mathbf{N}\underline{x}(s) = \mathbf{N}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{B}\underline{r}(s) \\ \underline{z}(s) = \mathbf{F}_{RN}(s)\underline{x}_C(s) = \mathbf{F}_{RN}(s)\left(\mathbf{F}_{RD}(s) + \mathbf{D}_C(s)\right)^{-1}\underline{r}(s) \end{cases} \tag{5.444}$$

The relation $\underline{z}(s) = \mathbf{F}_{RN}(s)\left(\mathbf{F}_{RD}(s) + \mathbf{D}_C(s)\right)^{-1}\underline{r}(s)$ indicates that the (transmission) zeros are invariant under state feedback.

Moreover, because both control loops exhibit the same input-output behavior, the following relation holds:

$$\mathbf{N}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{B} = \mathbf{F}_{RN}(s)\left(\mathbf{F}_{RD}(s) + \mathbf{D}_C(s)\right)^{-1} \tag{5.445}$$

Let the closed-loop polynomial matrix $\mathbf{D}_{cl}(s)$ be defined as follows:

$$\mathbf{D}_{cl}(s) = \mathbf{F}_{RD}(s) + \mathbf{D}_C(s) \tag{5.446}$$

Then the previous equation reads:

$$\boxed{\mathbf{N}\left(s\mathbb{I} - \mathbf{A} + \mathbf{BK}\right)^{-1}\mathbf{B} = \mathbf{F}_{RN}(s)\mathbf{D}_{cl}^{-1}(s)} \tag{5.447}$$

Furthermore, using this relation in (5.442) leads to the following relation:

$$\boxed{\mathbb{I} + \mathbf{K}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} = \mathbf{D}_{cl}(s)\mathbf{F}_{RD}^{-1}(s)} \tag{5.448}$$

Furthermore, the following relations hold[36] where $\mathbf{\Gamma}_c(\cdot)$ denotes the highest column-degree coefficient matrix and $\underline{\delta}_{c_i}(\cdot)$ denotes the ith column degree.

$$\begin{cases} \mathbf{\Gamma}_c\left(\mathbf{D}_{cl}(s)\right) = \mathbf{\Gamma}_c\left(\mathbf{F}_{RD}^{-1}(s)\right) \\ \underline{\delta}_{c_i}\left(\mathbf{D}_{cl}(s)\right) = \underline{\delta}_{c_i}\left(\mathbf{F}_{RD}^{-1}(s)\right) \qquad \forall\, i \end{cases} \tag{5.449}$$

Matrix $\mathbf{D}_{cl}(s)$ contains exactly the same number of free parameters than the state feedback matrix $\mathbf{K}$.

5.12.7 Pre-filtering applied to SISO plants

We will assume hereafter the following *state* feedback control of a SISO plant:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{u}(t) = -\mathbf{K}\underline{x}(t) + \underline{r}_{pf}(t) \\ z(t) = \mathbf{N}\underline{x}(t) \end{cases} \tag{5.450}$$

As shown in Figure 5.11, the pre-filter $C_{pf}(s)$ is a controller which is situated outside the feedback loop.

What is the purpose of the pre-filter ? Once the state feedback gain $\mathbf{K}$ is designed, the eigenvalues of closed-loop state matrix $\mathbf{A} - \mathbf{BK}$ are set, but not the zeros of the closed-loop transfer function $G(s)$:

$$G(s) = \frac{Z(s)}{R_{pf}(s)} = \mathbf{N}\left(s\mathbb{I} - (\mathbf{A} - \mathbf{BK})\right)^{-1}\mathbf{B} \tag{5.451}$$

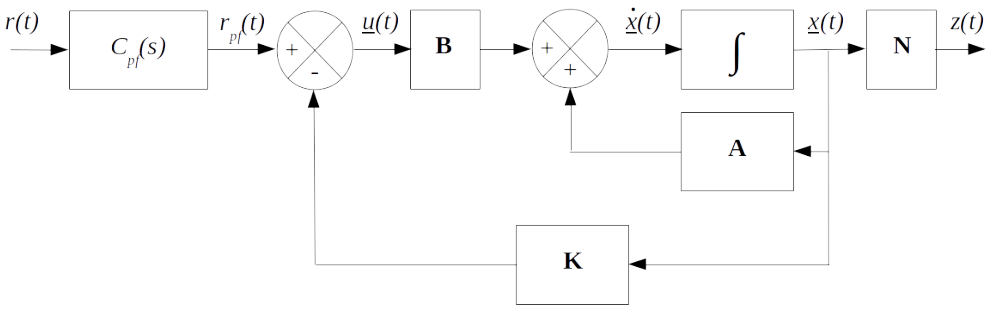

Figure 5.11: State feedback loop with pre-filter

These zeros may cause undesirable overshoots in the transient response of the closed-loop system. The purpose of the pre-filter $C_{pf}(s)$ is to reduce or eliminate such overshoots in the closed-loop system. Additionally the pre-filter may annihilate slow stable poles which sometimes cannot be shifted by the controller.

We focus in Figure 5.11. Let $N_{cl}(s)$ be the numerator of transfer function $G(s) = \frac{Z(s)}{R_{pf}(s)}$ and $D_{cl}(s)$ its denominator:

$$G(s) = \frac{Z(s)}{R_{pf}(s)} = \frac{N_{cl}(s)}{D_{cl}(s)} \tag{5.452}$$

We will assume that transfer function $\frac{Z(s)}{R_{pf}(s)}$ has all its zeros with negative real-parts, or equivalently that all the roots of $N_{cl}(s)$ are located in the left half plane.

Pre-filter $C_{pf}(s)$ is designed such that its poles cancel the zeros of the closed-loop system (i.e. the roots of $N_{cl}(s)$). If there is no pole of the closed-loop system to cancel, the numerator of the pre-filter is set to be a constant K_{pf}. In such a case the transfer function of the full system reads:

$$\frac{Z(s)}{R(s)} = \frac{K_{pf}}{D_{cl}(s)} \tag{5.453}$$

As a consequence the transfer function of the pre-filter reads:

$$C_{pf}(s) = \frac{K_{pf}}{N_{cl}(s)} \tag{5.454}$$

Note that this is only possible because the roots of $N_{cl}(s)$ have negative real-parts, meaning $C_{pf}(s)$ is stable.

Usually constant K_{pf} is set such that the static gain of $\frac{Z(s)}{R(s)}$ is unitary, meaning that the position error is zero:

$$\left.\frac{Y(s)}{R(s)}\right|_{s=0} = 1 \Rightarrow K_{pf} = D_{cl}(0) \tag{5.455}$$

Additionally the numerator of the pre-filter may also cancel some *slow* stable poles (poles in the left plane) of the closed-loop system when they are not placed

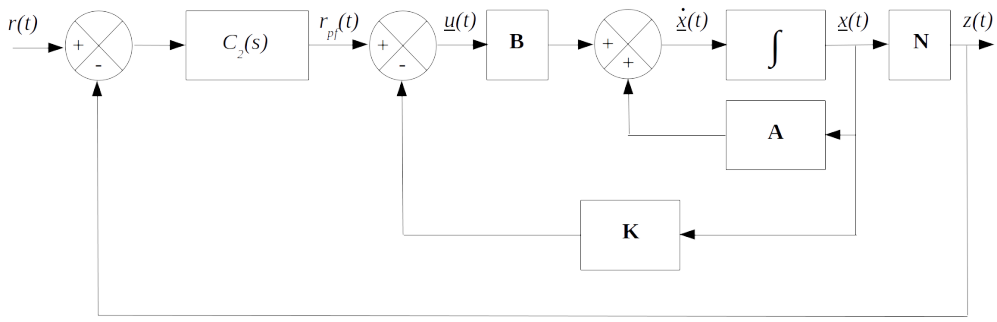

Figure 5.12: State feedback loop with pre-filter inside the closed-loop

by the controller $\mathbf{K}$. In this case, the numerator of the pre-filter $C_{pf}(s)$ is no more a constant.

Equivalently, the pre-filter may be inserted inside the closed-loop, as shown in Figure 5.12.

Figure 5.11 and 5.12 are equivalent as soon as the following relation holds:

$$C_{pf}(s)G(s) = \frac{C_2(s)G(s)}{1 + C_2(s)G(s)} \tag{5.456}$$

Finally, controller $C_2(s)$ can be computed from $C_{pf}(s)$ as follows:

$$C_{pf}(s) = \frac{C_2(s)}{1 + C_2(s)G(s)} \Leftrightarrow C_2(s) = \frac{C_{pf}(s)}{1 - G(s)C_{pf}(s)} \tag{5.457}$$

Appendices

Appendix A

Refresher on linear algebra

A.1 Section overview

The purpose of this chapter is to review the main results in elementary linear algebra. We will review vectors and matrices notions. This chapter ends with the presentation of the notions of eigenvalues and eigenvectors. The content of this chapter is mainly based on the material provided within the paper of Daniel S. Stutts[1] and Gregory J. Hakim[2].

A.2 Vectors

A.2.1 Definitions

A *column* vector, or simply a vector, is a set of numbers which are written in a column form:

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{A.1}$$

A *row* vector is a set of numbers which are written in a horizontal form. We denote a *row* vector by $\underline{x}^T$ where ^{T} denotes the transpose operation:

$$\underline{x}^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \tag{A.2}$$

Vectors can represent the coordinate of a point within a space of dimension n.

[1] https://www.researchgate.net/publication/242366881_Linear_Algebra_Primer
[2] https://atmos.washington.edu/ hakim/591/LA_primer.pdf

A.2.2 Vectors operations

- The sum (or subtraction) of two vectors proceed element-wise:

$$\begin{cases} \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ \underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \end{cases} \Rightarrow \underline{x} + \underline{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix} \tag{A.3}$$

- The sum (or subtraction) is:
  - Commutative:
$$\underline{x} + \underline{y} = \underline{y} + \underline{x} \tag{A.4}$$
  - Associative:
$$(\underline{x} + \underline{y}) + \underline{z} = \underline{x} + (\underline{y} + \underline{z}) \tag{A.5}$$
- The sum (or subtraction) of two vectors which are not of the same size is undefined.
- For vector subtraction, you have to replace $+$ by $-$ in the preceding expressions.
- Multiplication of a vector $\underline{x}$ by a scalar c is defined by the multiplication of each number of the vector by c:

$$c\underline{x} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix} \tag{A.6}$$

- The *inner product* (or *dot product*) $\underline{x}^T\underline{y}$ of two vectors $\underline{x}$ and $\underline{y}$ of the same size is obtained by multiplying each number element-wise:

$$\begin{cases} \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ \underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \end{cases} \Rightarrow \underline{x}^T\underline{y} = \sum_i^n x_i y_i \tag{A.7}$$

A.3 Matrices

A.3.1 Definitions

A $n \times m$ matrix is a rectangular array of numbers formed by n rows and m columns:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \tag{A.8}$$

Number a_{ij} refers to the number which is situated on the ith row and the jth column.

Matrix and vectors can be used to represent a system of equations in a compact form:

$$\begin{aligned} &\begin{cases} a_{11}x_1 + \cdots a_{1m}x_m & = b_1 \\ & \vdots \\ a_{n1}x_1 + \cdots a_{nm}x_m & = b_m \end{cases} \\ &\Leftrightarrow \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \\ &\Leftrightarrow \mathbf{A}\underline{x} = \underline{b} \end{aligned} \tag{A.9}$$

- A *square* matrix is a matrix with the same number of rows and columns;
- A *diagonal* matrix is a square matrix in which the numbers outside the main diagonal are all zero;
- The *identity* matrix $\mathbb{I}$ is a diagonal matrix having only ones along the main diagonal:

$$\mathbb{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ & \ddots & & \\ 0 & \cdots & 0 & 1 \end{bmatrix} \tag{A.10}$$

- The *transpose* of a matrix $\mathbf{A}$ has rows and columns which are interchanged: the first row becomes the first column, the second row becomes the second column and so on. The *transpose* of a matrix $\mathbf{A}$ is denoted $\mathbf{A}^T$:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \Rightarrow \mathbf{A}^T = \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1m} & \cdots & a_{nm} \end{bmatrix} \tag{A.11}$$

- A *symmetric* matrix is a square matrix that is equal to its transpose;
- The *trace* of a square matrix is the sum of its diagonal numbers:

$$\mathrm{tr}\,(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} \tag{A.12}$$

A.3.2 Matrix Operations

- The sum (or subtraction) of two matrices of the same size proceed element-wise:

$$\begin{cases} \mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nm} \end{bmatrix} \end{cases} \Rightarrow \mathbf{A}+\mathbf{B} = \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1m}+b_{1m} \\ \vdots & & \vdots \\ a_{n1}+b_{n1} & \cdots & a_{nm}+b_{nm} \end{bmatrix} \tag{A.13}$$

- The sum (or subtraction) of two matrices which are not of the same size is undefined.

- The sum (or subtraction) of a matrix with a scalar is defined as the sum (or subtraction) of each number of the matrix with the scalar:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \Rightarrow \mathbf{A}+c = \begin{bmatrix} a_{11}+c & \cdots & a_{1m}+c \\ \vdots & & \vdots \\ a_{n1}+c & \cdots & a_{nm}+c \end{bmatrix} \tag{A.14}$$

- Multiplication of a matrix $\mathbf{A}$ by a scalar c is defined by the multiplication of each number of the matrix by c:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \Rightarrow c\mathbf{A} = \begin{bmatrix} ca_{11} & \cdots & ca_{1m} \\ \vdots & & \vdots \\ ca_{n1} & \cdots & ca_{nm} \end{bmatrix} \tag{A.15}$$

- If $\mathbf{A}$ is an $n \times m$ matrix and $\mathbf{B}$ is an $m \times p$ matrix then the *matrix product* $\mathbf{AB}$ is defined to be the $n \times p$ matrix for which the number on the ith row and the jth column is obtained by taking the *dot product* of the corresponding i^{th} row of the left matrix with the jth column of the right matrix:

$$\begin{cases} \mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} = \begin{bmatrix} \underline{a}_1^T \\ \vdots \\ \underline{a}_n^T \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mp} \end{bmatrix} = \begin{bmatrix} \underline{b}_1 & \cdots & \underline{b}_p \end{bmatrix} \end{cases} \tag{A.16}$$

$$\Rightarrow \mathbf{AB} = \begin{bmatrix} \underline{a}_1^T \underline{b}_1 & \cdots & \underline{a}_1^T \underline{b}_p \\ \vdots & & \vdots \\ \underline{a}_n^T \underline{b}_1 & \cdots & \underline{a}_n^T \underline{b}_p \end{bmatrix}$$

- The kth *power* of a square matrix is obtained be multiplying k-times the matrix:

$$\mathbf{A}^k = \underbrace{\mathbf{A} \cdots \mathbf{A}}_{\text{k-times}} \tag{A.17}$$

A.3.3 Properties

For any matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ the following hold:

- $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
- $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
- $\mathbb{I}\mathbf{A} = \mathbf{A}\mathbb{I} = \mathbf{A}$
- $(\mathbf{AB})\,\mathbf{C} = \mathbf{A}\,(\mathbf{BC})$
- $\mathbf{A}\,(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
- $\mathbf{A}^0 = \mathbb{I}$
- $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$
- For any scalar c: $c\mathbf{A} = \mathbf{A}c$
- But be careful, in general $\mathbf{AB} \neq \mathbf{BA}$

A.3.4 Determinant and inverse

The determinant of a square matrix is a scalar. If the matrix is not square its determinant is undefined.

For of a 2×2 square matrix its determinant represents the area of the parallelogram obtained by the vectors in the rows of the matrix:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \Rightarrow \det(\mathbf{A}) = a_{11}a_{22} - a_{21}a_{12} \tag{A.18}$$

Let $\mathbf{A}$ be a square $n \times n$ matrix and $\mathbf{A}_{ij}$ be the square $(n-1) \times (n-1)$ submatrix obtained by removing the ith row and the jth column from $\mathbf{A}$. Then determinant of $\mathbf{A}$ may be obtained recursively by reduction to the 2×2 form as follows:

$$\det(\mathbf{A}) = \sum_{i=1}^{n} a_{ij}\left((-1)^{i+j}\det(\mathbf{A}_{ij})\right) = \sum_{j=1}^{n} a_{ij}\left((-1)^{i+j}\det(\mathbf{A}_{ij})\right) \tag{A.19}$$

The inverse of a square matrix $\mathbf{A}$ is the matrix denoted $\mathbf{A}^{-1}$ such that:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbb{I} \tag{A.20}$$

A matrix that has no inverse is called singular.

The inverse of a matrix can be expressed by the following formula where $\mathrm{adj}\,(\mathbf{A})$ is called the *adjoint* (or *adjugate*) matrix:

$$\mathbf{A}^{-1} = \frac{\mathrm{adj}\,(\mathbf{A})}{\det\,(\mathbf{A})} \tag{A.21}$$

The number on the ith and j^{th} column of the *adjoint* matrix $\mathrm{adj}\,(\mathbf{A})$ is the *cofactor* of a_{ij}. The *cofactor* of a_{ij} is the determinant of the submatrix $\mathbf{A}_{ij}$ obtained by removing the ith row and the jth column from $\mathbf{A}$ multiplied by $(-1)^{i+j}$.

For of a 2×2 square matrix we get:

$$\begin{aligned} &\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \Rightarrow \begin{cases} \det\,(\mathbf{A}) = a_{11}a_{22} - a_{21}a_{12} \\ \mathrm{adj}\,(\mathbf{A}) = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \end{cases} \\ &\Rightarrow \mathbf{A}^{-1} = \tfrac{1}{a_{11}a_{22}-a_{21}a_{12}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \end{aligned} \tag{A.22}$$

It can be shown that:

- If $\det\,(\mathbf{A}) \neq 0$ then $\mathbf{A}$ is nonsingular ;
- If any row or column of $\mathbf{A}$ is zero then $\det\,(\mathbf{A}) = 0$;
- If two rows or columns of $\mathbf{A}$ are proportional then $\det\,(\mathbf{A}) = 0$;
- $\det\,(\mathbf{AB}) = \det\,(\mathbf{A})\det\,(\mathbf{B})$;
- $\det\left(\mathbf{A}^T\right) = \det\,(\mathbf{A})$;
- $\det\left(\mathbf{A}^{-1}\right) = \frac{1}{\det(\mathbf{A})}$ provided that $\mathbf{A}^{-1}$ exists ;
- If $\mathbf{A}$ is a $n \times n$ matrix and c a scalar then $\det\,(c\mathbf{A}) = c^n \det\,(\mathbf{A})$;
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$;
- $\left(\mathbf{A}^{-1}\right)^T = \left(\mathbf{A}^T\right)^{-1}$;

Furthermore let $\mathbf{A}$ be of dimension $n \times n$, $\mathbf{B}$ of dimension $n \times k$, $\mathbf{D}$ of dimension $k \times k$ and $\mathbf{C}$ of dimension $k \times n$, then the Woodbury Formulae[3] reads:

$$\begin{cases} (\mathbf{A}+\mathbf{BDC})^{-1} &= \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}^{-1}+\mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1} \\ \mathbf{DC}(\mathbf{A}+\mathbf{BDC})^{-1} &= (\mathbf{D}^{-1}+\mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1} \end{cases} \tag{A.23}$$

[3]Linear Algebra Primer Gregory J. Hakim, University of Washington, 2 January 2009 v2.0

A.4 Eigenvalues and eigenvectors

A vector $\underline{x}$ is called a (right) eigenvector of matrix $\mathbf{A}$ if $\underline{x}$ is proportional to $\mathbf{A}\underline{x}$, or equivalently in the same direction than $\mathbf{A}\underline{x}$:

$$s\underline{x} = \mathbf{A}\underline{x} \tag{A.24}$$

In order for (A.24) to hold the following relation shall be satisfied:

$$(s\mathbb{I} - \mathbf{A})\,\underline{x} = \underline{0} \tag{A.25}$$

Relation (A.25) holds for $\underline{x} \neq \underline{0}$ as soon as the *resolvent* matrix $\Phi(s) = (s\mathbb{I} - \mathbf{A})^{-1}$ is singular. For the *resolvent* matrix $\Phi(s)$ to be singular we shall have:

$$\det(s\mathbb{I} - \mathbf{A}) = 0 \tag{A.26}$$

Assuming that $\mathbf{A}$ is of dimension $n \times n$ then the determinant $\det(s\mathbb{I} - \mathbf{A})$, which is called the *characteristic polynomial* of $\mathbf{A}$, is a polynomial of degree n. Furthermore its n roots are called the *eigenvalues* of $\mathbf{A}$ and are usually denoted $\lambda_1, \cdots, \lambda_n$.

Appendix B

Overview of Lagrangian Mechanics

B.1 Euler-Lagrange equations

Euler-Lagrange equations is a useful technique to obtain the equations of motion of mechanical systems. Euler-Lagrange equations rely on the expressions of the kinetic and potential energy of the system with respect to its *generalized coordinates* q_i, one for each degree of freedom of the system.

Euler-Lagrange equations read:

$$\boxed{\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) - \frac{\partial \mathcal{L}}{\partial q_i} = Q_i} \tag{B.1}$$

where:

- q_i denotes a *generalized coordinates* of the system. Generalized coordinates are composed by the set of minimum size of variables which allows to determine unambiguously the *configuration* of the system. They are either positions or angles. The number of *generalized coordinates* is equal to the number of degrees of freedom of the system. We will denote $\underline{q}$ the vector of generalized coordinates:

$$\underline{q} = [q_1, \cdots, q_n]^T \tag{B.2}$$

- The *Lagrangian* $\mathcal{L}$ denotes the difference between the kinetic energy, which is denoted $T(\underline{q}, \underline{\dot{q}})$, and the potential energy, which is denoted $V(\underline{q})$. The kinetic energy $T(\underline{q}, \underline{\dot{q}})$ depends on the generalized coordinates $\underline{q}$ and also on their derivatives $\underline{\dot{q}}$ whereas the potential energy $V(\underline{q})$ is a function of only the generalized coordinates $\underline{q}$:

$$\mathcal{L} = T(\underline{q}, \underline{\dot{q}}) - V(\underline{q}) \tag{B.3}$$

- The kinetic energy $T(\underline{q}, \underline{\dot{q}})$ of a rigid body is defined as follows, where P is any point of the rigid body $\mathcal{B}$:

$$T(\underline{q}, \underline{\dot{q}}) = \frac{1}{2} \iiint_{\mathcal{B}} \underline{v}(P)^T \underline{v}(P)\, dm \tag{B.4}$$

We recall that the velocity vector $\underline{v}(P)$ of any material point P of a rigid body can be expressed as the sum of the velocity $\underline{v}$ of its center of mass and the cross product between the angular velocity vector $\underline{\omega}$, which is independent of P, of the rigid body with respect to its center of mass and the relative position vector $\underline{\widetilde{x}}(P)$ of P with respect to the center of mass of the body:

$$\underline{v}(P) = \underline{v} + \underline{\omega} \times \underline{\widetilde{x}}(P) \tag{B.5}$$

Substituting for $\underline{v}(P)$ in the expression of $T(\underline{q}, \underline{\dot{q}})$ and expanding, we get:

$$\begin{aligned} T(\underline{q}, \underline{\dot{q}}) &= \tfrac{1}{2} \iiint_{\mathcal{B}} \left(\underline{v} + \underline{\omega} \times \underline{\widetilde{x}}(P)\right)^T \left(\underline{v} + \underline{\omega} \times \underline{\widetilde{x}}(P)\right) dm \\ &= \tfrac{1}{2} \iiint_{\mathcal{B}} \underline{v}^T \underline{v} + \left(\underline{\omega} \times \underline{\widetilde{x}}(P)\right)^T \left(\underline{\omega} \times \underline{\widetilde{x}}(P)\right) \\ &\qquad + 2\, \underline{v}^T \left(\underline{\omega} \times \underline{\widetilde{x}}(P)\right) dm \end{aligned} \tag{B.6}$$

However:

$$\frac{1}{2} \iiint_{\mathcal{B}} \underline{v}^T \underline{v}\, dm = \frac{1}{2} \underline{v}^T \underline{v} \iiint_{\mathcal{B}} dm = \frac{1}{2} m\, \underline{v}^T \underline{v} \tag{B.7}$$

and, by definition of the center of mass:

$$\begin{aligned} \tfrac{1}{2} \iiint_{\mathcal{B}} \underline{v}^T \left(\underline{\omega} \times \underline{\widetilde{x}}(P)\right) dm &= \tfrac{1}{2}\, \underline{v}^T \left(\underline{\omega} \times \iiint_{\mathcal{B}} \underline{\widetilde{x}}(P)\, dm\right) \\ &= \tfrac{1}{2}\, \underline{v}^T \left(\underline{\omega} \times \underline{0}\right) \\ &= 0 \end{aligned} \tag{B.8}$$

Moreover the following relation holds:

$$\left(\underline{\omega} \times \underline{\widetilde{x}}(P)\right)^T \left(\underline{\omega} \times \underline{\widetilde{x}}(P)\right) = \left(\left(\underline{\widetilde{x}}(P)^T \underline{\widetilde{x}}(P)\right) \underline{\omega} - \left(\underline{\widetilde{x}}(P)^T \underline{\omega}\right) \underline{\widetilde{x}}(P)\right)^T \underline{\omega} \tag{B.9}$$

We finally get:

$$\begin{aligned} T(\underline{q}, \underline{\dot{q}}) &= \tfrac{1}{2} m\, \underline{v}^T \underline{v} + \tfrac{1}{2} \left(\iiint_{\mathcal{B}} \left(\left(\underline{\widetilde{x}}(P)^T \underline{\widetilde{x}}(P)\right) \underline{\omega} - \left(\underline{\widetilde{x}}(P)^T \underline{\omega}\right) \underline{\widetilde{x}}(P)\right)^T dm\right) \underline{\omega} \\ &= \tfrac{1}{2} m\, \underline{v}^T \underline{v} + \tfrac{1}{2} \left(\iiint_{\mathcal{B}} \left(\left(\underline{\widetilde{x}}(P)^T \underline{\widetilde{x}}(P)\right) \underline{\omega} - \left(\underline{\widetilde{x}}(P)^T \underline{\omega}\right) \underline{\widetilde{x}}(P)\right) dm\right)^T \underline{\omega} \end{aligned} \tag{B.10}$$

Hence, we obtain the *König* decomposition:

$$T(\underline{q}, \underline{\dot{q}}) = \frac{1}{2} m\, \underline{v}^T \underline{v} + \frac{1}{2} \underline{H}^T \underline{\omega} \tag{B.11}$$

where $\underline{H}$ is the angular momentum of the rigid body relatively to its center of mass:

$$\begin{aligned} \underline{H} &:= \iiint_{\mathcal{B}} \left(\left(\underline{\widetilde{x}}(P)^T \underline{\widetilde{x}}(P)\right) \underline{\omega} - \left(\underline{\widetilde{x}}(P)^T \underline{\omega}\right) \underline{\widetilde{x}}(P)\right) dm \\ &= \iiint_{\mathcal{B}} \left(\underline{\widetilde{x}}(P) \times \underline{\omega} \times \underline{\widetilde{x}}(P)\right) dm \end{aligned} \tag{B.12}$$

Finally, the angular momentum $\underline{H}$ can expressed as the product between the angular velocity $\underline{\omega}$ and the moment of inertia $\mathbf{I} = \mathbf{I}^T$ of the rigid body relatively to its center of mass:

$$\underline{H} = \mathbf{I}\, \underline{\omega} \tag{B.13}$$

Thus, for a rigid body with mass m and moment of inertia $\mathbf{I}$ the kinetic energy $T(\underline{q}, \underline{\dot{q}})$ is obtained as the sum between the kinetic energy due to the velocity $\underline{v}$ of the center of mass of the body and its angular velocity $\underline{\omega}$, both velocities being expressed in an inertial frame:

$$\boxed{T(\underline{q}, \underline{\dot{q}}) = \frac{1}{2} m\, \underline{v}^T \underline{v} + \frac{1}{2} \underline{\omega}^T \mathbf{I}\, \underline{\omega}} \tag{B.14}$$

– Moment of inertia $\mathbf{I}$ is a positive definite matrix. Matrix $\mathbf{I}$ shall be computed with respect to the point O corresponding to the origin of the frame chosen for the *generalized coordinates* $\underline{q}$. Denoting by x, y and z the coordinates of an element of mass dm with respect to the point O, the moment of inertia $\mathbf{I}$ reads as follows:

$$\begin{aligned} \mathbf{I} &= \begin{bmatrix} I_x & -I_{xy} & -I_{xz} \\ -I_{xy} & I_y & -I_{yz} \\ -I_{xz} & -I_{yz} & I_z \end{bmatrix} \\ &= \begin{bmatrix} \int (y^2 + z^2)\, dm & -\int x\, y\, dm & -\int x\, z\, dm \\ -\int x\, y\, dm & \int (x^2 + z^2)\, dm & -\int y\, z\, dm \\ -\int x\, z\, dm & -\int y\, z\, dm & \int (x^2 + y^2)\, dm \end{bmatrix} \\ &= \mathbf{I}^T > 0 \end{aligned} \tag{B.15}$$

Quantities I_x, I_y and I_z are called moments of inertia with respect to the x, y and z axis,respectively whereas quantities I_{xy}, I_{xz} and I_{yz} are called products of inertia.

– Huygens-Steiner theorem gives the moment of inertia $\mathbf{I}_P$ at any point P of a solid of mass m given its moment of inertia $\mathbf{I}$ defined at a point O:

$$\underline{OP} = \begin{bmatrix} x_P \\ y_P \\ z_P \end{bmatrix} \Rightarrow \mathbf{I} = \mathbf{I}_P + m \begin{bmatrix} (y_P^2 + z_P^2) & -x_P\, y_P & -x_P\, z_P \\ -x_P\, y_P & (x_P^2 + z_P^2) & -y_P\, z_P \\ -x_P\, z_P & -y_P\, z_P & (x_P^2 + y_P^2) \end{bmatrix} \tag{B.16}$$

– It is worth noticing that the kinetic and the potential energy have to be evaluated in an inertial frame.

Assuming that a non-inertial frame is considered, let $\underline{v}^b$ be the linear velocity expressed in the non-inertial frame and $\underline{\nu}$ the angular velocity expressed in the non-inertial frame. In such a non-inertial frame the kinetic energy $T(\underline{q}, \underline{\dot{q}})$ reads:

$$T(\underline{q}, \underline{\dot{q}}) = \frac{1}{2} m \left(\underline{v}^b\right)^T \underline{v}^b + \frac{1}{2} \underline{\nu}^T \mathbf{I}\, \underline{\nu} \tag{B.17}$$

Denoting by $\underline{\eta}$ the vector of angles which allows to position the non-inertial frame (*body* frame) with respect to the inertial frame (those are *Euler* angles for example) and by $\mathbf{R}_b^i\left(\underline{\eta}\right)$ the rotation matrix from the non-inertial frame (body frame) to the inertial frame, we have:

$$\boxed{\underline{v} = \mathbf{R}_b^i\left(\underline{\eta}\right) \underline{v}^b} \tag{B.18}$$

Similarly, denoting by $\underline{\nu}$ the angular velocity in the non-inertial frame and by $\underline{\omega}$ the angular velocity in the inertial frame, we have:

$$\boxed{\underline{\nu} = \mathbf{W}(\underline{\eta})\,\underline{\omega}} \tag{B.19}$$

The relation between $\underline{\omega}$ and $\underline{\nu}$, that is matrix $\mathbf{W}(\underline{\eta})$, is obtained by developing each row of the following matrix equation:

$$\boxed{\frac{d}{dt}\mathbf{R}_b^i(\underline{\eta}) = \mathbf{R}_b^i(\underline{\eta})\,\mathbf{\Omega}(\underline{\nu}) \text{ where } \begin{cases} \underline{\omega} := \dot{\underline{\eta}} \\ \underline{\nu} := \begin{bmatrix} p & q & r \end{bmatrix}^T \\ \mathbf{\Omega}(\underline{\nu}) = \begin{bmatrix} 0 & -r & q \\ r & 0 & -p \\ -q & p & 0 \end{bmatrix} \end{cases}} \tag{B.20}$$

Let $\underline{w}$ be a vector with coordinates $\begin{bmatrix} a & b & c \end{bmatrix}^T$. Using the fact that $\mathbf{\Omega}(\underline{\nu})$ is an antisymmetric (or skew-symmetric) matrix, we get the following relation:

$$\begin{aligned} \underline{\nu} = \begin{bmatrix} p \\ q \\ r \end{bmatrix} \Rightarrow \mathbf{\Omega}(\underline{\nu})\,\underline{w} &= \begin{bmatrix} 0 & -r & q \\ r & 0 & -p \\ -q & p & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ &= -\begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix} \begin{bmatrix} p \\ q \\ r \end{bmatrix} \\ &= -\mathbf{\Omega}(\underline{w})\,\underline{\nu} \end{aligned} \tag{B.21}$$

Using the fact that a rotation matrix is an orthogonal matrix, meaning that $\left(\mathbf{R}_b^i\left(\underline{\eta}\right)\right)^{-1} = \left(\mathbf{R}_b^i\left(\underline{\eta}\right)\right)^T$, the kinetic energy $T(\underline{q}, \dot{\underline{q}})$ expressed in the inertial frame reads as follows:

$$\begin{aligned} T(\underline{q}, \dot{\underline{q}}) &= \tfrac{1}{2}m\left(\underline{v}^b\right)^T \underline{v} + \tfrac{1}{2}\left(\underline{\nu}\right)^T \mathbf{I}\,\underline{\nu} \\ &= \tfrac{1}{2}m\left(\left(\mathbf{R}_b^i\left(\underline{\eta}\right)\right)^T \underline{v}\right)^T \left(\left(\mathbf{R}_b^i\left(\underline{\eta}\right)\right)^T \underline{v}\right) + \tfrac{1}{2}\left(\mathbf{W}(\underline{\eta})\,\underline{\omega}\right)^T \mathbf{I}\,\left(\mathbf{W}(\underline{\eta})\,\underline{\omega}\right) \\ &= \tfrac{1}{2}m\underline{v}^T\underline{v} + \tfrac{1}{2}\underline{\omega}\,\mathbf{W}^T(\underline{\eta})\,\mathbf{I}\,\left(\mathbf{W}(\underline{\eta})\,\underline{\omega}\right) \end{aligned} \tag{B.22}$$

Thus, the kinetic energy $T(\underline{q}, \dot{\underline{q}})$ expressed in the inertial frame is:

$$T(\underline{q}, \dot{\underline{q}}) = \frac{1}{2}m\underline{v}^T\underline{v} + \frac{1}{2}\underline{\omega}^T\,\mathbf{J}\left(\underline{\eta}\right)\,\underline{\omega} \tag{B.23}$$

where the *fictitious* moment of inertia $\mathbf{J}\left(\underline{\eta}\right)$ is defined by:

$$\mathbf{J}\left(\underline{\eta}\right) = \mathbf{W}^T(\underline{\eta})\,\mathbf{I}\,\mathbf{W}(\underline{\eta}) \tag{B.24}$$

Finally, let $\mathbf{I}_P$ be the inertia matrix with respect to a point P of the rigid body, $\underline{v}_P^b$ the linear velocity of P expressed in the non-inertial frame, $\underline{\nu}_P$

its angular velocity expressed in the non-inertial frame and $\underline{r}_{PG}$ the vector between the rigid body centre of mass G and P. Then, denoting by $\times$ the cross product between two vectors, the kinetic energy $T(\underline{q}, \underline{\dot{q}})$ of P reads as follows[1]:

$$T(\underline{q}, \underline{\dot{q}}) = \frac{1}{2} m \left(\underline{v}_P^b\right)^T \underline{v}_P^b + \frac{1}{2} \left(\underline{\nu}_P\right)^T \mathbf{I}_P \, \underline{\nu}_P + m \left(\underline{v}_P^b\right)^T \left(\underline{\nu}_P \times \underline{r}_{PG}\right) \quad \text{(B.25)}$$

- Q_i represents the *non-conservative generalized forces* that are either forces or torques that are external to the system or which cannot be obtained as the derivative of a scalar potential function (that is a potential energy). Generalized force Q_i generates a movement, either a translation for a force or a rotation for a torque, according to the direction of the generalized coordinates q_i.

- The chain rule may be used to get the following relation:

$$\left(\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) - \frac{\partial \mathcal{L}}{\partial q_i}\right) q_i = \frac{\partial \mathcal{L}}{\partial t} - \frac{d}{dt}\left(\mathcal{L} - \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) \quad \text{(B.26)}$$

Consequently, assuming that *Lagrangian* $\mathcal{L}$ does not explicitly depends on time ($\frac{\partial \mathcal{L}}{\partial t} = 0$) and that no *non-conservative generalized forces* acts on the system ($Q_i = 0$), then the quantity $\mathcal{L} - \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$ is a first integral of the system:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial t} = 0 \\ Q_i = 0 \end{cases} \Rightarrow \mathcal{L} - \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \text{constant} \quad \text{(B.27)}$$

The kinetic energy $T(\underline{q}, \underline{\dot{q}})$ can always be expressed as follows where $\mathbf{M}(\underline{q}) = \mathbf{M}(\underline{q})^T > 0$ is a positive definite matrix:

$$\boxed{T(\underline{q}, \underline{\dot{q}}) = \frac{1}{2} \underline{\dot{q}}^T \mathbf{M}(\underline{q}) \, \underline{\dot{q}} \quad \text{where} \quad \mathbf{M}(\underline{q}) = \mathbf{M}^T(\underline{q}) > 0} \quad \text{(B.28)}$$

Using the preceding relation for the kinetic energy, it can be shown that:

$$\boxed{\frac{d}{dt}\left(\frac{\partial T(\underline{q}, \underline{\dot{q}})}{\partial \underline{\dot{q}}}\right) - \frac{\partial T(\underline{q}, \underline{\dot{q}})}{\partial \underline{q}} = \mathbf{M}(\underline{q}) \underline{\ddot{q}} + \mathbf{C}(\underline{q}, \underline{\dot{q}}) \, \underline{\dot{q}}} \quad \text{(B.29)}$$

Matrix $\mathbf{C}(\underline{q}, \underline{\dot{q}}) \underline{\dot{q}}$ is the so-called Coriolis (terms involving products $\dot{q}_i \, \dot{q}_j \; i \neq j$) and centrifugal (terms involving products $\dot{q}_i^2$) forces matrix. It is worth noticing that the kth row of matrix $\mathbf{C}(\underline{q}, \underline{\dot{q}})$, which will be denoted $\underline{c}_k^T(\underline{q}, \underline{\dot{q}})$, can be obtained thanks to the following relation:

$$\begin{cases} \underline{c}_k^T(\underline{q}, \underline{\dot{q}}) = \underline{\dot{q}}^T \mathbf{S}_k(\underline{q}) \\ \mathbf{S}_k(\underline{q}) = \frac{1}{2}\left(\frac{\partial \underline{M}_k(\underline{q})}{\partial \underline{q}} + \left(\frac{\partial \underline{M}_k(\underline{q})}{\partial \underline{q}}\right)^T - \frac{\partial \mathbf{M}(\underline{q})}{\partial q_k}\right) \end{cases} \quad \text{(B.30)}$$

[1] Complete dynamic model of the Twin Rotor MIMO System (TRMS) with experimental validation, Azamat Tastemirov, Andrea Lecchini-Visintini, Rafael M. Morales-Viviescas, Control Engineering Practice 66 (2017) 89–98

where $\underline{M}_k(\underline{q})$ is the kth column of matrix $\mathbf{M}(\underline{q})$ and q_k is the kth component of vector $\underline{q}$. Terms of matrix $\mathbf{C}(\underline{q}, \dot{\underline{q}})$ are called Christoffel symbols.

Once developed, Euler-Lagrange equations (B.1) can then be written as follows:

$$\boxed{\mathbf{M}(\underline{q})\ddot{\underline{q}} + \mathbf{C}(\underline{q}, \dot{\underline{q}})\dot{\underline{q}} + \underline{G}(\underline{q}) = \mathbf{D}(\underline{q}, \dot{\underline{q}}, \underline{u})} \tag{B.31}$$

- The term $\underline{G}(\underline{q})$ corresponds to conservative forces. We recall that *conservative generalized forces* can be obtained as $\frac{\partial V(q_i)}{\partial q_i}$ where $V(q_i)$ is a potential function. Such *conservative generalized forces* are taken into account within the Lagrangian $\mathcal{L}$ but not within the *generalized forces* Q_i.

- The term $\mathbf{D}(\underline{q}, \dot{\underline{q}}, \underline{u})$ corresponds to the *non-conservative generalized forces*. The term $\underline{u}$ corresponds to the control vector of the system.

- External forces due to dissipation are encompassed within the term $\mathbf{D}(\underline{q}, \dot{\underline{q}}, \underline{u})$. They are of the form $-\frac{\partial \mathcal{F}(\dot{\underline{q}})}{\partial \dot{\underline{q}}}$ where $\mathcal{F}(\dot{\underline{q}})$ is the Rayleigh dissipation function which by definition satisfies[2]:

$$\dot{\underline{q}}^T \frac{\partial \mathcal{F}(\dot{\underline{q}})}{\partial \dot{\underline{q}}} \geq 0 \tag{B.32}$$

  It is worth noticing that damping forces are always dissipative.

- The system is said to be passive when the following relation holds [2]:

$$\text{Passivity} \;\Leftrightarrow \frac{d}{dt}\mathbf{M}(\underline{q}) - 2\,\mathbf{C}(\underline{q}, \dot{\underline{q}}) = \mathbf{P} \text{ where } \mathbf{P} = -\mathbf{P}^T \tag{B.33}$$

Assume now that the generalized coordinates $\underline{q}$ are not all independent but subject to m *holonomic* constraints (*nonholonomic* constraints are of the form $g_j(\underline{q}, \dot{\underline{q}}) = 0$):

$$g_j(\underline{q}) = 0 \qquad j = 1, \cdots, m \tag{B.34}$$

Then the variations of δq_i are not free but must obey to the following relations:

$$\delta g_j(\underline{q}) = \sum_{i=1}^{n} \frac{\partial g_j(\underline{q})}{\partial q_i} \delta q_i = 0 \qquad j = 1, \cdots, m \tag{B.35}$$

In that situation the constraints (B.34) are associated with m Lagrange's multipliers λ_j and the Euler-Lagrange equations read:

$$\begin{cases} g_j(q_1, \cdots, q_n) = 0 \qquad j = 1, \cdots, m \\ \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) - \frac{\partial \mathcal{L}}{\partial q_i} = Q_i + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial q_i} \qquad i = 1, \cdots, n \end{cases} \tag{B.36}$$

[2]Passivity-based Control of Euler-Lagrange Systems, Romeo Ortega, Antonio Loria, Per Johan Nicklasson and Hebertt Sira-Ramírez, Springer, 1998

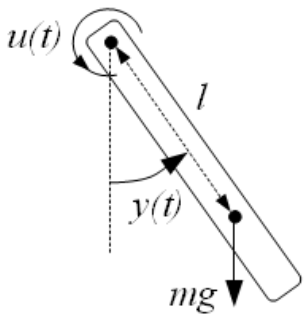

Figure B.1: Robot arm

The term $\sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial q_i}$ corresponds to the internal generalized force $\widetilde{Q}_i$ related to the constraints on component q_i of the generalized coordinates vector:

$$\widetilde{Q}_i = \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial q_i} \tag{B.37}$$

A generalization of the Euler-Lagrange equations taking into account the constraints on the components of the generalized coordinates vector are the Boltzmann-Hamel equations[3].

B.2 Robot arm

Let's consider Figure B.1 where a robot arm is depicted: $u(t)$ is the torque applied by a motor drive and $y(t)$ is the angular position of the arm. In addition we denote m the mass of the arm, l the distance between the axis of the motor and the centre of mass of the arm, b the viscous friction coefficient, I its inertia and g the acceleration of gravity.

The generalized coordinates is chosen to be the angle $q(t) = y(t)$. Indeed the knowledge of the value of $y(t)$ allows to determine unambiguously the configuration of the system. It is worth noticing that the knowledge of the coordinates of the centre of gravity of the arm also allows to determine unambiguously the configuration of the system. Nevertheless the coordinates of the centre of gravity form a vector of dimension 2 whereas the angle $q(t) = y(t)$ is a scalar of dimension 1. Consequently the coordinates of the centre of gravity don't constitute a set a minimum size.

The coordinates of the centre of gravity within the inertial frame read:

$$\vec{OG}(t) = \begin{bmatrix} x_G(t) \\ y_G(t) \end{bmatrix} = \begin{bmatrix} l\sin(\theta(t)) \\ -l\cos(\theta(t)) \end{bmatrix} \tag{B.38}$$

[3]Jonathan M. Cameron and Wayne Book, Modeling Mechanisms with Nonholonomic Joints Using the Botzmann-Hamel Equations, The International Journal of Robotics Research, February 1997 16(1):47-59, DOI:10.1177/027836499701600104

By taking the derivative we get the components of the velocity vector as well as the square of its norm:

$$\underline{v}(t) = \frac{d}{dt}\vec{OG}(t) = \begin{bmatrix} l\dot{\theta}\cos(\theta) \\ l\dot{\theta}\sin(\theta) \end{bmatrix} \Rightarrow \underline{v}(t)^T\underline{v}(t) = l^2\dot{\theta}^2 \tag{B.39}$$

The kinetic energy $T(q,\dot{q})$ and the potential energy $V(q)$ read:

$$\begin{cases} T(q,\dot{q}) = \frac{1}{2}m\underline{v}(t)^T\underline{v}(t) + \frac{1}{2}I\dot{\theta}^2 = \frac{1}{2}ml^2\dot{\theta}^2 + \frac{1}{2}I\dot{\theta}^2 \\ V(q) = -mgl\cos(\theta) \end{cases} \tag{B.40}$$

And the Lagrangian $\mathcal{L}$ reads:

$$\mathcal{L} = T(q,\dot{q}) - V(q) = \frac{1}{2}ml^2\dot{\theta}^2 + \frac{1}{2}I\dot{\theta}^2 + mgl\cos(\theta) \tag{B.41}$$

Consequently the partial derivatives have the following expression:

$$\begin{cases} \frac{\partial\mathcal{L}}{\partial\dot{\theta}} = (ml^2 + I)\dot{\theta} \\ \frac{\partial\mathcal{L}}{\partial\theta} = -mgl\sin(\theta) \end{cases} \tag{B.42}$$

The non-conservative generalized forces (forces and torques) are here the torque $u(t)$ applied by the motor as well as the friction torque $-k\dot{\theta}$ which is proportional to the angular velocity $\dot{\theta}$:

$$Q = u(t) - k\dot{\theta} \tag{B.43}$$

Applying the Euler-Lagrange equations (B.1) leads to the following dynamic model of the robot arm:

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\theta}}\right) - \frac{\partial\mathcal{L}}{\partial\theta} = Q \tag{B.44}$$

That is:

$$(ml^2 + \mathbf{I})\ddot{\theta} + mgl\sin(\theta) = u(t) - k\dot{\theta} \tag{B.45}$$

It is clear that the preceding equation can be written as $\mathbf{M}(\underline{q})\ddot{\underline{q}} + \mathbf{C}(\underline{q},\dot{\underline{q}})\dot{\underline{q}} + \underline{G}(\underline{q}) = \mathbf{D}(\underline{q},\dot{\underline{q}},\underline{u})$ (cf. (B.31)) where $\mathbf{C}(\underline{q},\dot{\underline{q}}) = 0$ and where the term $\mathbf{D}(\underline{q},\dot{\underline{q}},\underline{u}) := u(t) - k\dot{\theta}$ represents the generalized force encompassing control u and friction torque $k\dot{\theta}$, which is a non conservative force.

B.3 Quadrotor

The quadcopter structure is presented in Figure B.2. It shows angular velocities ω_i and forces f_i created by the four rotors, numbered from $i = 1$ to $i = 4$. Torque direction is opposite to velocities ω_i.

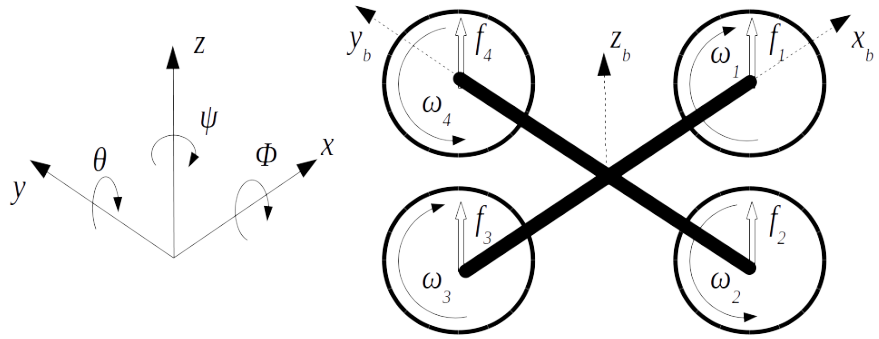

Figure B.2: Inertial and body frames of a quadcopter

B.3.1 Inertial frame and body frame

The following vectors will be used:

- $\underline{\xi} = \begin{bmatrix} x & y & z \end{bmatrix}^T$ is the vector whose components define the position of the centre of gravity of the quadcopter in the inertial frame;
- $\underline{\eta} = \begin{bmatrix} \phi & \theta & \psi \end{bmatrix}^T$ is the vector of so-called Euler angles whose components define the orientation (attitude) of the quadcopter in the inertial frame:
  - The roll angle ϕ determines the angular position of the quadcopter around the x-axis of the body frame;
  - The pitch angle θ determines the angular position of the quadcopter around the y-axis of the body frame;
  - The yaw angle ψ determines the angular position of the quadcopter around the z-axis of the body frame.
- $\underline{\nu} = \begin{bmatrix} p & q & r \end{bmatrix}^T$ is the vector whose components define the angular velocities of the quadcopter in the body frame.

B.3.2 Direct Cosine Matrix (DCM) and kinematic relations

Let $\underline{x}^i$ be a vector expressed in the inertial frame, $\underline{x}^b$ a vector expressed in the body frame and $\mathbf{R}_i^b(\underline{\eta})$ the rotation matrix, also called Direct Cosine Matrix (DCM), from the inertial frame to the body frame:

$$\underline{x}^b = \mathbf{R}_i^b(\underline{\eta})\underline{x}^i \tag{B.46}$$

Rotation matrix $\mathbf{R}_i^b(\underline{\eta})$ is obtained by the multiplication of the rotation matrices around Euler angles, namely yaw angle ψ, pitch angle θ and then roll angle ϕ, respectively. Denoting $c_x = \cos(x)$, $s_x = \sin(x)$ and $\mathbf{R}_y$ the rotation

matrix dedicated to angle y we get:

$$\begin{aligned}\mathbf{R}_i^b(\underline{\eta}) &= \mathbf{R}_\phi \mathbf{R}_\theta \mathbf{R}_\psi \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\phi & s_\phi \\ 0 & -s_\phi & c_\phi \end{bmatrix} \begin{bmatrix} c_\theta & 0 & -s_\theta \\ 0 & 1 & 0 \\ s_\theta & 0 & c_\theta \end{bmatrix} \begin{bmatrix} c_\psi & s_\psi & 0 \\ -s_\psi & c_\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} c_\theta c_\psi & c_\theta s_\psi & -s_\theta \\ (s_\phi s_\theta c_\psi - c_\phi s_\psi) & (s_\phi s_\theta s_\psi + c_\phi c_\psi) & s_\phi c_\theta \\ (c_\phi s_\theta c_\psi + s_\phi s_\psi) & (c_\phi s_\theta s_\psi - s_\phi c_\psi) & c_\phi c_\theta \end{bmatrix}\end{aligned} \tag{B.47}$$

It is worth noticing that $\mathbf{R}_i^b(\underline{\eta})$ is an orthogonal matrix. Consequently the rotation matrix $\mathbf{R}_b^i(\underline{\eta})$ from the body frame to the inertial frame is obtained as follows:

$$\begin{aligned}\mathbf{R}_b^i(\underline{\eta}) &:= \left(\mathbf{R}_i^b(\underline{\eta})\right)^{-1} = \left(\mathbf{R}_i^b(\underline{\eta})\right)^T \\ &= \begin{bmatrix} c_\theta c_\psi & (s_\phi s_\theta c_\psi - c_\phi s_\psi) & (c_\phi s_\theta c_\psi + s_\phi s_\psi) \\ c_\theta s_\psi & (s_\phi s_\theta s_\psi + c_\phi c_\psi) & (c_\phi s_\theta s_\psi - s_\phi c_\psi) \\ -s_\theta & s_\phi c_\theta & c_\phi c_\theta \end{bmatrix}\end{aligned} \tag{B.48}$$

The relation between the angular velocities (p, q, r) in the body frame and the time derivative of the Euler angles (ϕ, θ, ψ) is the following:

$$\underline{\nu} := \begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} \dot{\phi} \\ 0 \\ 0 \end{bmatrix} + \mathbf{R}_\phi \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix} + \mathbf{R}_\phi \mathbf{R}_\theta \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix} \tag{B.49}$$

We finally get:

$$\begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\sin(\theta) \\ 0 & \cos(\phi) & \sin(\phi)\cos\theta \\ 0 & -\sin(\phi) & \cos(\phi)\cos\theta \end{bmatrix} \begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} \tag{B.50}$$

That is:

$$\underline{\nu} = \mathbf{W}(\underline{\eta})\, \dot{\underline{\eta}} \tag{B.51}$$

where:

$$\underline{\eta} := \begin{bmatrix} \phi \\ \theta \\ \psi \end{bmatrix} \tag{B.52}$$

and:

$$\mathbf{W}(\underline{\eta}) = \begin{bmatrix} 1 & 0 & -\sin(\theta) \\ 0 & \cos(\phi) & \sin(\phi)\cos(\theta) \\ 0 & -\sin(\phi) & \cos(\phi)\cos(\theta) \end{bmatrix} \tag{B.53}$$

It is worth noticing that the preceding relation can be obtained from the following equality which simply states that the time derivative of matrix $\mathbf{R}_b^i(\underline{\eta})$ can be seen as matrix $\mathbf{\Omega}(\underline{\nu})$ of the angular velocities in the body frame expressed in the inertial frame:

$$\frac{d}{dt}\mathbf{R}_b^i(\underline{\eta}) = \mathbf{R}_b^i(\underline{\eta})\, \mathbf{\Omega}(\underline{\nu}) \text{ where } \mathbf{\Omega}(\underline{\nu}) = -\mathbf{\Omega}(\underline{\nu})^T = \begin{bmatrix} 0 & -r & q \\ r & 0 & -p \\ -q & p & 0 \end{bmatrix}] \tag{B.54}$$

Conversely we have:

$$\dot{\underline{\eta}} = \mathbf{W}(\underline{\eta})^{-1}\,\underline{\nu} \tag{B.55}$$

where:

$$\mathbf{W}(\underline{\eta})^{-1} = \begin{bmatrix} 1 & \sin(\phi)\tan(\theta) & \cos(\phi)\tan(\theta) \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \frac{\sin(\phi)}{\cos(\theta)} & \frac{\cos(\phi)}{\cos(\theta)} \end{bmatrix} \tag{B.56}$$

B.3.3 Forces and torques

We will use the following notation:

- I_r is the inertia moment of each rotor;
- d is the distance between the rotor and the centre of mass of the quadcopter, that is the arm length basically;
- ω_i the angular velocity of the ith rotor;
- f_i is the thrust force created by each rotor in the direction of the body z_b-axis;
- C_l is the lift coefficient;
- C_d is the drag coefficient.

The non-conservative generalized forces (forces and torques) are the following:

- Aerodynamic thrust force f_t in the direction of the body z_b-axis. This force is the sum of each force coming from each rotor:

$$f_t = \sum_{i=1}^{4} f_i = \sum_{i=1}^{4} C_l\,\omega_i^2 \tag{B.57}$$

  Let vector $\underline{f}_a^i$ be the thrust force created by all rotors in the inertial frame:

$$\underline{f}_a^i = \mathbf{R}_b^i(\underline{\eta}) \begin{bmatrix} 0 \\ 0 \\ f_t \end{bmatrix} = \mathbf{R}_b^i(\underline{\eta}) \begin{bmatrix} 0 \\ 0 \\ \sum_{i=1}^{4} C_l\,\omega_i^2 \end{bmatrix} \tag{B.58}$$

  Where $\mathbf{R}_b^i(\underline{\eta})$ denotes the rotation matrix from the body frame to the inertial frame.

- Aerodynamic torque $\underline{\tau}^b$ expressed in the body frame. Vector $\underline{\tau}^b$ is the sum of the following terms:
  - Torque $\underline{\tau}_a^b$ coming from the aerodynamics actions coming from propellers in the direction of the corresponding body frame angles:

$$\underline{\tau}_a^b = \begin{bmatrix} \tau_\phi \\ \tau_\theta \\ \tau_\psi \end{bmatrix} = \begin{bmatrix} d\,C_l\left(\omega_4^2 - \omega_2^2\right) \\ d\,C_l\left(\omega_3^2 - \omega_1^2\right) \\ C_d\left(-\omega_1^2 + \omega_2^2 - \omega_3^2 + \omega_4^2\right) \end{bmatrix} \tag{B.59}$$

– Torque $\underline{\tau}_g^b$ coming from the gyroscopic effect due to propeller rotation:

$$\begin{aligned}\underline{\tau}_g^b &= -\left(I_r \frac{d}{dt} \begin{bmatrix} 0 \\ 0 \\ \sum_{i=1}^4 sgn(\omega_i)\,\omega_i \end{bmatrix} + \begin{bmatrix} p \\ q \\ r \end{bmatrix} \times I_r \begin{bmatrix} 0 \\ 0 \\ \sum_{i=1}^4 sgn(\omega_i)\,\omega_i \end{bmatrix} \right) \\ &= \begin{bmatrix} I_r\, q\ (\omega_1 - \omega_2 + \omega_3 - \omega_4) \\ -I_r\, p\ (\omega_1 - \omega_2 + \omega_3 - \omega_4) \\ I_r\ (\dot{\omega}_1 - \dot{\omega}_2 + \dot{\omega}_3 - \dot{\omega}_4) \end{bmatrix}\end{aligned} \tag{B.60}$$

where $sgn(\omega_i) = +1$ for counterclockwise propeller rotation and $sgn(\omega_i) = -1$ for clockwise propeller rotation.

We finally get:

$$\begin{aligned}\underline{\tau}^b &= \underline{\tau}_a^b + \underline{\tau}_g^b \\ &= \begin{bmatrix} d\,C_l\left(\omega_4^2 - \omega_2^2\right) + I_r\, q\ (\omega_1 - \omega_2 + \omega_3 - \omega_4) \\ d\,C_l\left(\omega_3^2 - \omega_1^2\right) - I_r\, p\ (\omega_1 - \omega_2 + \omega_3 - \omega_4) \\ C_d\left(-\omega_1^2 + \omega_2^2 - \omega_3^2 + \omega_4^2\right) + I_r\ (\dot{\omega}_1 - \dot{\omega}_2 + \dot{\omega}_3 - \dot{\omega}_4) \end{bmatrix}\end{aligned} \tag{B.61}$$

It is worth noticing that terms which depends on I_r come from the gyroscopic effect due to propeller rotation and are usually omitted.

B.3.4 Generalized coordinates

The vector of generalized coordinates $\underline{q}$ which will determine the configuration of the quadcopter is a vector with six components which is defined as follows:

$$\underline{q} = \begin{bmatrix} \underline{\xi} \\ \underline{\eta} \end{bmatrix} \text{ where } \begin{cases} \underline{\xi} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \\ \underline{\eta} = \begin{bmatrix} \phi \\ \theta \\ \psi \end{bmatrix} \end{cases} \tag{B.62}$$

B.3.5 Inertia matrix

The quadcopter is assumed to have symmetric structure with four arms aligned with the body x and y axes. Thus the inertia matrix $\mathbf{I}$ is diagonal and $I_x = I_y$:

$$\mathbf{I} = \begin{bmatrix} I_x & 0 & 0 \\ 0 & I_y & 0 \\ 0 & 0 & I_z \end{bmatrix} \tag{B.63}$$

B.3.6 Kinetic energy

Because the inertia matrix is expressed in the body frame, vector $\underline{\nu}$ is naturally chosen to express the rotational kinetic energy. Nevertheless the rotational

kinetic energy shall be expressed as a function of the chosen generalized coordinates. Consequently we shall use the transformation matrix $\mathbf{W}(\underline{\eta})$ to get $\dot{\underline{\eta}}$ from $\underline{\nu}$ and express the rotational kinetic energy as a function of the chosen generalized coordinates. Consequently the kinetic energy $T\left(\underline{q},\dot{\underline{q}}\right)$ reads:

$$\begin{aligned} T\left(\underline{q},\dot{\underline{q}}\right) &= \tfrac{1}{2}m\dot{\underline{\xi}}^T\dot{\underline{\xi}} + \tfrac{1}{2}\underline{\nu}^T\,\mathbf{I}\,\underline{\nu} \\ &= \tfrac{1}{2}m\dot{\underline{\xi}}^T\dot{\underline{\xi}} + \tfrac{1}{2}\dot{\underline{\eta}}^T\mathbf{W}(\underline{\eta})^T\,\mathbf{I}\,\mathbf{W}(\underline{\eta})\dot{\underline{\eta}} \\ &= \tfrac{1}{2}m\dot{\underline{\xi}}^T\dot{\underline{\xi}} + \tfrac{1}{2}\dot{\underline{\eta}}^T\,\mathbf{J}(\underline{\eta})\,\dot{\underline{\eta}} \end{aligned} \tag{B.64}$$

where we use symmetric matrix $\mathbf{J}(\underline{\eta})$ defined as follows:

$$\mathbf{J}(\underline{\eta}) = \mathbf{W}(\underline{\eta})^T\,\mathbf{I}\,\mathbf{W}(\underline{\eta}) = \mathbf{J}(\underline{\eta})^T \tag{B.65}$$

From (B.53) and (B.63) matrix $\mathbf{J}(\underline{\eta})$ reads:

$$\begin{aligned} \mathbf{J}(\underline{\eta}) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\phi & -s_\phi \\ -s_\theta & s_\phi c_\theta & c_\phi c_\theta \end{bmatrix} \begin{bmatrix} I_x & 0 & 0 \\ 0 & I_y & 0 \\ 0 & 0 & I_z \end{bmatrix} \begin{bmatrix} 1 & 0 & -s_\theta \\ 0 & c_\phi & s_\phi c_\theta \\ 0 & -s_\phi & c_\phi c_\theta \end{bmatrix} \\ &= \begin{bmatrix} I_x & 0 & -I_x s_\theta \\ 0 & I_y c_\phi^2 + I_z s_\phi^2 & (I_y - I_z)\, c_\phi s_\phi c_\theta \\ -I_x s_\theta & (I_y - I_z)\, c_\phi s_\phi c_\theta & I_x s_\theta^2 + I_y s_\phi^2 c_\theta^2 + I_z c_\phi^2 c_\theta^2 \end{bmatrix} \end{aligned} \tag{B.66}$$

Thus:

$$\begin{aligned} \frac{1}{2}\dot{\underline{\eta}}^T\,\mathbf{J}(\underline{\eta})\,\dot{\underline{\eta}} = \frac{1}{2}I_x\left(\dot{\phi} - \dot{\psi}\sin\theta\right)^2 + \frac{1}{2}I_y\left(\dot{\theta}\cos\phi + \dot{\psi}\sin\phi\cos\theta\right)^2 \\ + \frac{1}{2}I_z\left(\dot{\theta}\sin\phi - \dot{\psi}\cos\phi\cos\theta\right)^2 \end{aligned} \tag{B.67}$$

Kinetic energy $T\left(\underline{q},\dot{\underline{q}}\right)$ as a function of the chosen generalized coordinates finally reads:

$$\begin{aligned} T\left(\underline{q},\dot{\underline{q}}\right) = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) + \frac{1}{2}I_x\left(\dot{\phi} - \dot{\psi}\sin\theta\right)^2 \\ + \frac{1}{2}I_y\left(\dot{\theta}\cos\phi + \dot{\psi}\sin\phi\cos\theta\right)^2 + \frac{1}{2}I_z\left(\dot{\theta}\sin\phi - \dot{\psi}\cos\phi\cos\theta\right)^2 \end{aligned} \tag{B.68}$$

It can be shown that the determinant of symmetric matrix $\mathbf{J}(\underline{\eta})$ reads as follows and that this is a positive definite matrix $\forall\,\theta \neq (2k{+}1)\,\pi/2,\ k = 1, 2, \cdots$:

$$\det\left(\mathbf{J}(\underline{\eta})\right) = I_x I_y I_z\,(\cos(\theta))^2 \tag{B.69}$$

B.3.7 Potential energy

Potential energy $V(\underline{q})$ a function of the chosen generalized coordinates reads:

$$V(\underline{q}) = m\,g\,z = m\,g\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}\underline{\xi} \tag{B.70}$$

B.3.8 Lagrangian

Consequently Lagrangian $\mathcal{L}$ reads:

$$\begin{aligned}\mathcal{L} &= T\left(\underline{q},\dot{\underline{q}}\right) - V(\underline{q}) \\ &= \tfrac{1}{2}m\dot{\underline{\xi}}^T\dot{\underline{\xi}} + \tfrac{1}{2}\dot{\underline{\eta}}^T\mathbf{J}(\underline{\eta})\,\dot{\underline{\eta}} - mg\begin{bmatrix}0 & 0 & 1\end{bmatrix}\underline{\xi} \\ &= \tfrac{1}{2}m\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right) + \tfrac{1}{2}I_x\left(\dot{\phi}-\dot{\psi}\sin\theta\right)^2 \\ &\quad + \tfrac{1}{2}I_y\left(\dot{\theta}\cos\phi+\dot{\psi}\sin\phi\cos\theta\right)^2 + \tfrac{1}{2}I_z\left(\dot{\theta}\sin\phi-\dot{\psi}\cos\phi\cos\theta\right)^2 \\ &\quad - m\,g\,z\end{aligned} \tag{B.71}$$

B.3.9 Euler-Lagrange equations

We have seen that the Lagrangian $\mathcal{L}$ reads:

$$\mathcal{L} = \frac{1}{2}m\dot{\underline{\xi}}^T\dot{\underline{\xi}} + \frac{1}{2}\dot{\underline{\eta}}^T\mathbf{J}(\underline{\eta})\,\dot{\underline{\eta}} - mg\begin{bmatrix}0 & 0 & 1\end{bmatrix}\underline{\xi} \tag{B.72}$$

Since the Lagrangian $\mathcal{L}$ contains no cross terms combining $\dot{\underline{\xi}}$ with $\dot{\underline{\eta}}$ the Euler-Lagrange equations can be partitioned as follows:

- Euler-Lagrange equations with respect to $\underline{\xi}$ lead to the translational equations of motion of the quadcopter. Applying the Euler-Lagrange equations (B.1) and denoting by $\underline{f}_a^i$ the thrust force created by all rotors in the inertial frame we get:

$$\begin{aligned}\underline{f}_a^i &= \tfrac{d}{dt}\left(\tfrac{\partial\mathcal{L}}{\partial\dot{\underline{\xi}}}\right) - \tfrac{\partial\mathcal{L}}{\partial\underline{\xi}} \\ \Leftrightarrow \underline{f}_a^i &= m\ddot{\underline{\xi}} + mg\begin{bmatrix}0 & 0 & 1\end{bmatrix}^T\end{aligned} \tag{B.73}$$

  From (B.58) we get the differential equations for the positional accelerations:

$$\ddot{\underline{\xi}} = \frac{1}{m}\mathbf{R}_b^i(\underline{\eta})\begin{bmatrix}0\\0\\\sum_{i=1}^4 f_i\end{bmatrix} - g\begin{bmatrix}0\\0\\1\end{bmatrix} \tag{B.74}$$

- Euler-Lagrange equations with respect to $\underline{\eta}$ lead to the rotational equations of motion of the quadcopter. Applying Euler-Lagrange equations (B.1) and denoting by $\underline{\tau}^i$ the torque created by all rotors in the inertial frame we get:

$$\begin{aligned}\underline{\tau}^i &= \tfrac{d}{dt}\left(\tfrac{\partial\mathcal{L}}{\partial\dot{\underline{\eta}}}\right) - \tfrac{\partial\mathcal{L}}{\partial\underline{\eta}} \\ &= \tfrac{d}{dt}\left(\mathbf{J}(\underline{\eta})\,\dot{\underline{\eta}}\right) - \tfrac{\partial}{\partial\underline{\eta}}\left(\tfrac{1}{2}\dot{\underline{\eta}}^T\mathbf{J}(\underline{\eta})\,\dot{\underline{\eta}}\right) \\ &= \mathbf{J}(\underline{\eta})\,\ddot{\underline{\eta}} + \tfrac{d\,\mathbf{J}(\underline{\eta})}{dt}\dot{\underline{\eta}} - \tfrac{1}{2}\tfrac{\partial\left(\dot{\underline{\eta}}^T\mathbf{J}(\underline{\eta})\right)}{\partial\underline{\eta}}\,\dot{\underline{\eta}} \\ &= \mathbf{J}(\underline{\eta})\,\ddot{\underline{\eta}} + \left(\tfrac{d\,\mathbf{J}(\underline{\eta})}{dt} - \tfrac{1}{2}\tfrac{\partial\left(\dot{\underline{\eta}}^T\mathbf{J}(\underline{\eta})\right)}{\partial\underline{\eta}}\right)\dot{\underline{\eta}}\end{aligned} \tag{B.75}$$

  The preceding equation can be rewritten as follows where $\mathbf{C}(\underline{\eta},\dot{\underline{\eta}})\,\dot{\underline{\eta}}$ is the Coriolis and centrifugal forces matrix:

$$\mathbf{J}(\underline{\eta})\ddot{\underline{\eta}} + \mathbf{C}(\underline{\eta},\dot{\underline{\eta}})\,\dot{\underline{\eta}} = \underline{\tau}^i \tag{B.76}$$

The expression of $\mathbf{J}(\underline{\eta})$ has been provided in (B.66) whereas the expression of coefficients C_{ij} of matrix $\mathbf{C}(\underline{\eta}, \underline{\dot{\eta}})$ are the following:

$$\mathbf{C}(\underline{\eta}, \underline{\dot{\eta}}) = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \tag{B.77}$$

where:

$$\begin{cases} C_{11} &= 0 \\ C_{12} &= (I_y - I_z)\dot{\theta} c_\phi s_\phi + \frac{1}{2}\dot{\psi} c_\theta \left((I_y - I_z)(s_\phi^2 - c_\phi^2) - I_x \right) \\ C_{13} &= (I_z - I_y)\dot{\psi} c_\phi s_\phi c_\theta^2 + \frac{1}{2}\dot{\theta} c_\theta \left((I_y - I_z)(s_\phi^2 - c_\phi^2) - I_x \right) \\ C_{21} &= (I_z - I_y)\dot{\theta} c_\phi s_\phi + \frac{1}{2}\dot{\psi} c_\theta \left((I_z - I_y)(s_\phi^2 - c_\phi^2) + I_x \right) \\ C_{22} &= (I_z - I_y)\dot{\phi} c_\phi s_\phi \\ C_{23} &= \left(I_y s_\phi^2 + I_z c_\phi^2 - I_x \right) \dot{\psi} s_\theta c_\theta \\ & \quad + \frac{1}{2}\dot{\phi} c_\theta \left((I_z - I_y)(s_\phi^2 - c_\phi^2) + I_x \right) \\ C_{31} &= (I_y - I_z)\dot{\psi} c_\theta^2 s_\phi c_\phi \\ & \quad + \frac{1}{2}\dot{\theta} c_\theta \left((I_y - I_z)(c_\phi^2 - s_\phi^2) - I_x \right) \\ C_{32} &= (I_z - I_y)\dot{\theta} c_\phi s_\phi s_\theta + \left(I_x - I_y s_\phi^2 - I_z c_\phi^2 \right) \dot{\psi} s_\theta c_\theta \\ & \quad + \frac{1}{2}\dot{\phi} c_\theta \left((I_y - I_z)(c_\phi^2 - s_\phi^2) - I_x \right) \\ C_{33} &= (I_y - I_z)\dot{\phi} c_\phi s_\phi c_\theta^2 + I_x \dot{\theta} c_\theta s_\theta - \dot{\psi} c_\theta s_\theta \left(I_y s_\phi^2 + I_z c_\phi^2 \right) \end{cases} \tag{B.78}$$

It is worth noticing that the kth row of matrix $\mathbf{C}(\underline{\eta}, \underline{\dot{\eta}})$, which will be denoted $\underline{c}_k^T(\underline{\eta}, \underline{\dot{\eta}})$, can be obtained thanks to the following relation:

$$\begin{cases} \underline{c}_k^T(\underline{\eta}, \underline{\dot{\eta}}) = \underline{\dot{\eta}}^T \mathbf{S}_k(\underline{\eta}) \\ \mathbf{S}_k(\underline{\eta}) = \frac{1}{2} \left(\frac{\partial \underline{J}_k(\underline{\eta})}{\partial \underline{\eta}} + \left(\frac{\partial \underline{J}_k(\underline{\eta})}{\partial \underline{\eta}} \right)^T - \frac{\partial \mathbf{J}(\underline{\eta})}{\partial \eta_k} \right) \end{cases} \tag{B.79}$$

where $\underline{J}_k(\underline{\eta})$ is the kth column of matrix $\mathbf{J}(\underline{\eta})$ and η_k is the kth component of vector $\underline{\eta}$.

From (B.76) we get the differential equations for the angular accelerations:

$$\underline{\ddot{\eta}} = \mathbf{J}(\underline{\eta})^{-1} \left(\underline{\tau}^i - \mathbf{C}(\underline{\eta}, \underline{\dot{\eta}})\, \underline{\dot{\eta}} \right) \tag{B.80}$$

B.3.10 Newton-Euler equations

Let $\underline{v}^b = \mathbf{R}_i^b(\underline{\eta})\underline{\dot{\xi}}$ be the translational velocity vector expressed in the body frame and $\underline{f}^b$ the external forces applied on the quadcopter expressed in the body frame:

$$\underline{f}^b = \begin{bmatrix} 0 \\ 0 \\ \sum_{i=1}^4 f_i \end{bmatrix} - \mathbf{R}_i^b(\underline{\eta})\, m\, g \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{B.81}$$

Applying Newton-Euler equations, the translational and angular dynamics in the body frame of the rigid body quadcopter reads as follows:

$$\begin{bmatrix} m\mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \dot{\underline{v}}^b \\ \dot{\underline{\nu}} \end{bmatrix} + \begin{bmatrix} m\,\mathbf{\Omega}(\underline{\nu}) & \mathbf{0} \\ \mathbf{0} & \mathbf{\Omega}(\underline{\nu})\,\mathbf{I} \end{bmatrix} \begin{bmatrix} \underline{v}^b \\ \underline{\nu} \end{bmatrix} = \begin{bmatrix} \underline{f}^b \\ \underline{\tau}^b \end{bmatrix} \tag{B.82}$$

Where:

$$\mathbf{\Omega}(\underline{\nu}) = -\mathbf{\Omega}(\underline{\nu})^T = \begin{bmatrix} 0 & -r & q \\ r & 0 & -p \\ -q & p & 0 \end{bmatrix} \tag{B.83}$$

The preceding Newton-Euler equations are equivalent to equations (B.74) and (B.80) obtained through the Euler-Lagrange formalism:

$$\begin{cases} \underline{v}^b = \mathbf{R}_i^b(\underline{\eta})\dot{\underline{\xi}} \\ m\,\dot{\underline{v}}^b + m\,\mathbf{\Omega}(\underline{\nu})\,\underline{v}^b = \underline{f}^b \\ \dot{\underline{\eta}} = \mathbf{W}(\underline{\eta})^{-1}\,\underline{\nu} \\ \dot{\underline{\nu}} = \mathbf{I}^{-1}\left(\underline{\tau}^b - \mathbf{\Omega}(\underline{\nu})\,\mathbf{I}\,\underline{\nu}\right) \end{cases} \Leftrightarrow \begin{cases} \ddot{\underline{\xi}} = \frac{1}{m}\mathbf{R}_b^i(\underline{\eta}) \begin{bmatrix} 0 \\ 0 \\ \sum_{i=1}^4 f_i \end{bmatrix} - g \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ \ddot{\underline{\eta}} = \mathbf{J}(\underline{\eta})^{-1}\left(\underline{\tau}^i - \mathbf{C}(\underline{\eta},\dot{\underline{\eta}})\,\dot{\underline{\eta}}\right) \end{cases} \tag{B.84}$$

Where:

$$\underline{\tau}^i = \mathbf{W}(\underline{\eta})^T\underline{\tau}^b \tag{B.85}$$

The equivalence of the translational equations of motion is easily verified thanks to the kinematics relations.

As far as the Newton-Euler equations related to the rotational equations of motion we get:

$$\begin{aligned} &\begin{cases} \underline{\nu} = \mathbf{W}(\underline{\eta})\,\dot{\underline{\eta}} \\ \mathbf{I}\,\dot{\underline{\nu}} + \mathbf{\Omega}(\underline{\nu})\,\mathbf{I}\,\underline{\nu} = \underline{\tau}^b \end{cases} \\ &\Rightarrow \begin{cases} \dot{\underline{\nu}} = \dot{\mathbf{W}}(\underline{\eta})\,\dot{\underline{\eta}} + \mathbf{W}(\underline{\eta})\,\ddot{\underline{\eta}} \\ \mathbf{I}\left(\dot{\mathbf{W}}(\underline{\eta})\,\dot{\underline{\eta}} + \mathbf{W}(\underline{\eta})\,\ddot{\underline{\eta}}\right) + \mathbf{\Omega}(\underline{\nu})\,\mathbf{I}\mathbf{W}(\underline{\eta})\,\dot{\underline{\eta}} = \underline{\tau}^b \end{cases} \\ &\Rightarrow \mathbf{I}\mathbf{W}(\underline{\eta})\,\ddot{\underline{\eta}} + \left(\mathbf{I}\,\dot{\mathbf{W}}(\underline{\eta}) + \mathbf{\Omega}(\underline{\nu})\,\mathbf{I}\mathbf{W}(\underline{\eta})\right)\,\dot{\underline{\eta}} = \underline{\tau}^b \end{aligned} \tag{B.86}$$

Multiplying both side by $\mathbf{W}(\underline{\eta})^T$ leads to the equation of the rotational equation of motion obtained through the Euler-Lagrange formalism:

$$\begin{aligned} &\mathbf{W}(\underline{\eta})^T\mathbf{I}\mathbf{W}(\underline{\eta})\,\ddot{\underline{\eta}} + \left(\mathbf{W}(\underline{\eta})^T\mathbf{I}\,\dot{\mathbf{W}}(\underline{\eta}) + \mathbf{W}(\underline{\eta})^T\mathbf{\Omega}(\underline{\nu})\,\mathbf{I}\mathbf{W}(\underline{\eta})\right)\,\dot{\underline{\eta}} = \mathbf{W}(\underline{\eta})^T\underline{\tau}^b \\ &\Leftrightarrow \mathbf{J}(\underline{\eta})\ddot{\underline{\eta}} + \mathbf{C}(\underline{\eta},\dot{\underline{\eta}})\,\dot{\underline{\eta}} = \underline{\tau}^i \end{aligned} \tag{B.87}$$

More generally the Newton-Euler equations for a point which is located to $(\Delta_x, \Delta_y, \Delta_z)$ with respect to the center of mass of the rigid body with velocity (u_b, v_b, w_b) in the body fixed axis (of course these components are the components of $\underline{v}^b$, that is the velocity of the center of mass when the center of mass is considered, that is when $\Delta_x = \Delta_y = \Delta_z = 0$) reads as follows[4]:

$$\begin{bmatrix} m\mathbb{I} & -\mathbf{\Delta} \\ \mathbf{\Delta} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \dot{\underline{v}}^b \\ \dot{\underline{\nu}} \end{bmatrix} + \begin{bmatrix} m\,\mathbf{\Omega}(\underline{\nu}) & -\mathbf{\Omega}(\underline{\nu})\mathbf{\Delta} \\ \mathbf{\Omega}(\underline{\nu})\mathbf{\Delta} & \mathbf{\Omega}(\underline{\nu})\mathbf{I} - \mathbf{V}\mathbf{\Delta} \end{bmatrix} \begin{bmatrix} \underline{v}^b \\ \underline{\nu} \end{bmatrix} = \begin{bmatrix} \underline{f}^b \\ \underline{\tau}^b \end{bmatrix} \tag{B.88}$$

[4]Barton J. Bacon and Irene M. Gregory, General Equations of Motion for a Damaged Asymmetric Aircraft, NASA Langley Research Center, Hampton, VA, 23681

Where:

$$\begin{cases} \mathbf{\Omega}(\underline{\nu}) = \begin{bmatrix} 0 & -r & q \\ r & 0 & -p \\ -q & p & 0 \end{bmatrix} \\ \mathbf{\Delta} = \begin{bmatrix} 0 & -m\,\Delta_z & m\,\Delta_y \\ m\,\Delta_z & 0 & -m\,\Delta_x \\ -m\,\Delta_y & m\,\Delta_x & 0 \end{bmatrix} \\ \mathbf{V} = \begin{bmatrix} 0 & -w_b & v_b \\ w_b & 0 & -u_b \\ -v_b & u_b & 0 \end{bmatrix} \end{cases} \tag{B.89}$$

B.3.11 Translational equations of motion with wind

The velocity of the quadcopter with respect to the inertial frame (the Earth) is the sum of the velocity of the quadcopter with respect to the wind, which is denoted $\mathbf{R}_b^i(\underline{\eta})\underline{v}^b$, and the wind velocity, which is denoted $\underline{w}$. Denoting by $\underline{\xi}$ the position of the drone, we have:

$$\dot{\underline{\xi}} := \underline{v}^i = \mathbf{R}_b^i(\underline{\eta})\underline{v}^b + \underline{w} \tag{B.90}$$

where:

$$\underline{v}^b = \begin{bmatrix} u \\ v \\ w \end{bmatrix} \tag{B.91}$$

Rotation matrix $\mathbf{R}_b^i(\underline{\eta})$ is given by (B.48).

Taking the time derivative of the velocity in the inertial frame, $\underline{v}^i$, we get:

$$\dot{\underline{v}}^i = \dot{\mathbf{R}}_b^i\underline{v}^b + \mathbf{R}_b^i(\underline{\eta})\dot{\underline{v}}^b + \dot{\underline{w}} \tag{B.92}$$

From Newton's translational equations of motion we have:

$$m\dot{\underline{v}}^i = \sum \underline{f}^i \tag{B.93}$$

Multiplying by $\mathbf{R}_i^b(\underline{\eta})$ leads to the following relation:

$$m\mathbf{R}_i^b(\underline{\eta})\dot{\underline{v}}^i = \mathbf{R}_i^b(\underline{\eta})\sum \underline{f}^i = \sum \underline{f}^b \tag{B.94}$$

We get:

$$\begin{aligned} \sum \underline{f}^b &= m(\mathbf{R}_i^b(\underline{\eta})\dot{\mathbf{R}}_b^i\underline{v}^b + \dot{\underline{v}}^b + \mathbf{R}_i^b(\underline{\eta})\dot{\underline{w}}) \\ \Leftrightarrow \dot{\underline{v}}^b &= \frac{\sum \underline{f}^b}{m} - \mathbf{R}_i^b(\underline{\eta})\dot{\mathbf{R}}_b^i\underline{v}^b - \mathbf{R}_i^b(\underline{\eta})\dot{\underline{w}} \end{aligned} \tag{B.95}$$

where $\mathbf{R}_i^b(\underline{\eta})\dot{\mathbf{R}}_b^i := \mathbf{\Omega}(\underline{\nu})$ has been seen previously. We finally get the following equation of motion taking into account the wind component $\underline{w}$ reads:

$$\dot{\underline{v}}^b = \frac{\sum \underline{f}^b}{m} - \mathbf{\Omega}(\underline{\nu})\underline{v}^b - \mathbf{R}_i^b(\underline{\eta})\dot{\underline{w}} \tag{B.96}$$

Furthermore the wind is assumed to be not a constant but dependent on time t as well as on the quadcopter location $\underline{\xi} := [x, y, z]^T$. So we have: $\underline{w} := \underline{w}(t, \underline{\xi})$. Taking into account the rule of chain derivative we have:

$$\dot{\underline{w}}(t, \underline{\xi}) = \frac{\partial \underline{w}(t, \underline{\xi})}{\partial t} + \frac{\partial \underline{w}(t, \underline{\xi})}{\partial \underline{\xi}} \frac{\partial \underline{\xi}}{\partial t} \tag{B.97}$$

Taking into account that the time derivative of the location of the quadcopter is its velocity expressed in the inertial we have:

$$\frac{\partial \underline{\xi}}{\partial t} = \underline{v}^i = \mathbf{R}_b^i(\underline{\eta})\underline{v}^b + \underline{w} \tag{B.98}$$

Thus (B.96) finally reads:

$$\dot{\underline{v}}^b = \frac{\sum \underline{f}^b}{m} - \left(\mathbf{\Omega}(\underline{\nu}) + \mathbf{\Omega}(\underline{w})\right) \underline{v}^b - \mathbf{R}_i^b(\underline{\eta}) \left(\frac{\partial \underline{w}(t, \underline{\xi})}{\partial t} + \frac{\partial \underline{w}(t, \underline{\xi})}{\partial \underline{\xi}} \underline{w} \right) \tag{B.99}$$

where:

$$\mathbf{\Omega}(\underline{w}) = \mathbf{R}_i^b(\underline{\eta}) \frac{\partial \underline{w}(t, \underline{\xi})}{\partial \underline{\xi}} \mathbf{R}_b^i(\underline{\eta}) \tag{B.100}$$

Of course, as soon as $\underline{w} = \underline{0}$, we have $\mathbf{\Omega}(\underline{w}) = \underline{0}$ and we retrieve equation of motion (B.82).

B.3.12 Small angle approximation of angular dynamics

The second equation of (B.82) represents the angular dynamics in the body frame of the rigid body quadcopter:

$$\begin{aligned} &\mathbf{I}\dot{\underline{\nu}} + \mathbf{\Omega}(\underline{\nu})\,\mathbf{I}\underline{\nu} = \underline{\tau}^b \\ &\Leftrightarrow \dot{\underline{\nu}} = \mathbf{I}^{-1}\left(\underline{\tau}^b - \mathbf{\Omega}(\underline{\nu})\,\mathbf{I}\underline{\nu}\right) \end{aligned} \tag{B.101}$$

We have seen that angular velocities in the inertial frame are expressed in the body frame through the transformation matrix $\mathbf{W}(\underline{\eta})^{-1}$:

$$\dot{\underline{\eta}} = \mathbf{W}(\underline{\eta})^{-1}\,\underline{\nu} \tag{B.102}$$

The derivative of (B.102) with respect to time of the preceding equation leads to the expression of $\ddot{\underline{\eta}}$:

$$\ddot{\underline{\eta}} = \frac{d\mathbf{W}(\underline{\eta})^{-1}}{dt}\,\underline{\nu} + \mathbf{W}(\underline{\eta})^{-1}\,\dot{\underline{\nu}} \tag{B.103}$$

According to (B.56) we have:

$$\begin{aligned} \mathbf{W}(\underline{\eta})^{-1} &= \begin{bmatrix} 1 & \sin(\phi)\tan(\theta) & \cos(\phi)\tan(\theta) \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \frac{\sin(\phi)}{\cos(\theta)} & \frac{\cos(\phi)}{\cos(\theta)} \end{bmatrix} \\ \Rightarrow \frac{d\mathbf{W}(\underline{\eta})^{-1}}{dt} &= \begin{bmatrix} 0 & \frac{\dot{\phi}c_\phi s_\theta}{c_\theta} + \frac{\dot{\theta}s_\phi}{c_\theta^2} & \frac{\dot{\theta}c_\phi}{c_\theta^2} - \frac{\dot{\phi}s_\phi s_\theta}{c_\theta} \\ 0 & -\dot{\phi}s_\phi & -\dot{\phi}c_\phi \\ 0 & \frac{\dot{\phi}c_\phi}{c_\theta} + \frac{\dot{\theta}s_\phi s_\theta}{c_\theta^2} & \frac{\dot{\theta}c_\phi s_\theta}{c_\theta^2} - \frac{\dot{\phi}s_\phi}{c_\theta} \end{bmatrix} \end{aligned} \tag{B.104}$$

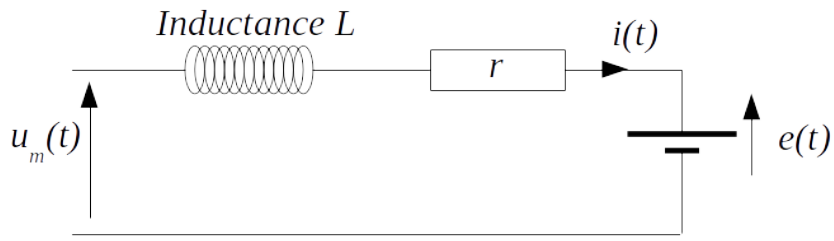

Figure B.3: Electrical model of a DC motor

Small angle approximation of angular dynamics is obtained by setting Euler angles vector $\underline{\eta}$ to zero within transformation matrix $\mathbf{W}(\underline{\eta})^{-1}$, that is by setting transformation matrix $\mathbf{W}(\underline{\eta})^{-1}$ to identity matrix. As a consequence the time derivative of the Euler angles $\dot{\underline{\eta}}$ is approximated by the angular velocities $\underline{\nu}$ of the quadcopter in the body frame:

$$\underline{\eta} = \underline{0} \Rightarrow \mathbf{W}(\underline{\eta})^{-1} \approx \mathbb{I} \Rightarrow \ddot{\underline{\eta}} \approx \dot{\underline{\nu}} \Rightarrow \dot{\underline{\eta}} \approx \underline{\nu} \tag{B.105}$$

In addition thanks to this approximation $\underline{\tau}^i = \mathbf{W}(\underline{\eta})^T \underline{\tau}^b \approx \underline{\tau}^b$ and (B.101) reads as follows, which is the small angle approximation of angular dynamics (B.80):

$$\ddot{\underline{\eta}} \approx \mathbf{I}^{-1} \left(\underline{\tau}^b - \mathbf{\Omega}(\dot{\underline{\eta}})\, \mathbf{I} \dot{\underline{\eta}} \right) \tag{B.106}$$

B.3.13 Synthesis model

The synthesis model is a simplified model when compared to the validation model (B.84). Synthesis model enables to design control laws in a linear time invariant frame.

As far as the angular dynamics is concerned, since term $\mathbf{\Omega}(\dot{\underline{\eta}})\, \mathbf{I} \dot{\underline{\eta}}$ is usually small when compared to $\underline{\tau}^b$, equation (B.106) can be reduced as follows, which is basically the dynamics of a double integrator:

$$\mathbf{\Omega}(\dot{\underline{\eta}})\, \mathbf{I} \dot{\underline{\eta}} \ll \underline{\tau}^b \Rightarrow \ddot{\underline{\eta}} \approx \mathbf{I}^{-1} \underline{\tau}^b \tag{B.107}$$

B.3.14 BLDC motor with ESC and propeller

Miniature BrushLess Direct Current (BLDC) motors are commonly used on small UAV. Such motor converts the electrical power to a torque on its output shaft at a specific angular speed ω_m[5].

A model of DC Motor Plant is shown in Figure B.3. Let:

- u_m be the motor input voltage;

[5] R. Martinez-Alvarado, F. J. Ruiz-Sanchez, A. Sanchez-Orta and O. Garcia-Salazar, Dynamic response of BLDC-thruster for small scale Quadrotors under aerodynamic load torque, 2014 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), 2014, pp. 1-6, doi: 10.1109/ROPEC.2014.7036341.

- i be the armature current;
- r be the motor internal resistance;
- L be the motor internal inductance;
- e be the back ElectroMotive Force (EMF).

From Figure B.3, the electrical model of a DC motor reads:

$$L\,\frac{d\,i}{dt} = u_m - r\,i - e \tag{B.108}$$

On the other hand, the mechanical model of a DC motor reads:

$$J\,\frac{d\,\omega_m}{dt} = \tau_m - \tau_d \tag{B.109}$$

where:

- ω_m is motor angular speed;
- τ_m is the motor torque;
- τ_d is the motor load;
- J is the propeller inertia.

The motor load τ_d is described as follows, where k_D is the aerodynamic torque coefficient which depends on the drag of the propeller:

$$\tau_d = k_D\,\omega_m\,|\omega_m| \tag{B.110}$$

Assuming that the BLDC motor is powered by an Electronic Speed Control (ESC) device, the coupling between the electrical model and the mechanical model appears through the following expression of angular speed ω_m and motor torque τ_m as a function of the ElectroMotive Force (EMF) e and the *constant* armature current i:

$$\boxed{\begin{cases} \omega_m = k_V\,e \\ \tau_m = k_T\,i \end{cases}} \tag{B.111}$$

The proportionality constant k_V is the motor's speed constant while k_T is the torque constant.

Finally, when all those relations are merged, we get the following dynamics model for a DC motor equipped with a propeller:

$$\boxed{\begin{cases} L\,\frac{d\,i}{dt} = u_m - r\,i - \frac{\omega_m}{k_V} \\ J\,\frac{d\,\omega_m}{dt} = k_T\,i - k_D\,\omega_m\,|\omega_m| \end{cases}} \tag{B.112}$$

When the motor internal inductance L is neglected, the first equation simplifies as follows:

$$L \approx 0 \Rightarrow u_m \approx r\,i + \frac{\omega_m}{k_V} \tag{B.113}$$

With this simplification, the second equation becomes:

$$i \approx \frac{u_m - \frac{\omega_m}{k_V}}{r} \Rightarrow J\,\frac{d\,\omega_m}{dt} \approx k_T\,\frac{u_m - \frac{\omega_m}{k_V}}{r} - k_D\,\omega_m\,|\omega_m| \tag{B.114}$$

We finally get:

$$\boxed{\frac{d\,\omega_m}{dt} \approx -\frac{1}{\tau}\,\omega_m - \frac{k_D}{J}\,\omega_m\,|\omega_m| + \frac{k_V}{\tau}\,u_m} \tag{B.115}$$

where:

$$\frac{1}{\tau} = \frac{k_T}{k_V\,J\,r} \tag{B.116}$$

Multiplying by τ the nonlinear differential equation (B.115), and assuming that time constant τ is *small*, we get the fact that the motor angular speed ω_m is proportional to the motor input voltage u_m:

$$\tau \approx 0 \Rightarrow \omega_m \approx k_V\,u_m \tag{B.117}$$

Alternatively nonlinear differential equation (B.115) can be linearized around an equilibrium point (ω_{me}, u_{me}):

$$\delta\dot{\omega}_m = A\,\delta\omega_m + B\,\delta u_m \text{ where } \begin{cases} 0 = -\frac{1}{\tau}\,\omega_{me} - \frac{k_D}{J}\,\omega_{me}\,|\omega_{me}| + \frac{k_V}{\tau}\,u_{me} \\ \delta\omega_m = \omega_m - \omega_{me} \\ \delta u_m = u_m - u_{me} \end{cases} \tag{B.118}$$

Finally, assuming that the electrical power $e\,i$ related to the back ElectroMotive Force (EMF) is fully converted into mechanical power $\tau_m\,\omega_m$, we have:

$$e\,i = \tau_m\,\omega_m \tag{B.119}$$

Using the coupling equations (B.111) we get $k_T = \frac{1}{k_V}$. Indeed:

$$\begin{cases} \omega_m = k_V\,e \\ \tau_m = k_T\,i \end{cases} \Rightarrow e\,i = \tau_m\,\omega_m \Leftrightarrow \frac{\cancel{\omega_m}\,i}{k_V} = k_T\,i\,\cancel{\omega_m} \Rightarrow k_T = \frac{1}{k_V} \tag{B.120}$$

Thanks to the ESC, the motor input voltage u_m is related to the commanded Pulse Width Modulation (PWM) $\delta_P \in [0, 1]$ by the following relation, where V_s is a constant supply voltage:

$$u_m = \delta_P\,V_s \tag{B.121}$$

To summarize, a BLDC motor with its ESC is characterized by the following 4 constants: the motor's speed constant k_V, the supply voltage V_s, the constant current passing through the windings i and the winding resistance r.

Appendix C

Linear Algebraic Method for Control System Design

C.1 Matrix polynomial fraction

C.1.1 Greatest Common Divisor of two polynomials

The greatest common divisor (frequently abbreviated as GCD) of two polynomials $N(s)$ and $\Psi(s)$ is a monic polynomial (a monic polynomial is a polynomial in which the nonzero coefficient of highest degree (also called the leading coefficient) is equal to 1) of the highest possible degree that is a factor of both $N(s)$ and $\Psi(s)$.

The greatest common divisor (GCD) of two polynomials $N(s)$ and $\Psi(s)$ can be obtained through the Euclidean algorithm: this consists in making repeated Euclidean division (according to the descending power of s) and to use the following results: the Euclidean division of $N(s)$ by $\Psi(s)$ such that $\deg(\Psi(s)) \leq \deg(N(s))$ provides a quotient $Q(s)$ and a remainder $R(s)$ such that:

$$\deg(\Psi(s)) \leq \deg(N(s)) \Rightarrow \begin{cases} N(s) = Q(s)\Psi(s) + R(s) \\ \deg(R(s)) < \deg(\Psi(s)) \end{cases} \tag{C.1}$$

Then:

$$\begin{cases} GCD(N(s), \Psi(s)) = GCD(\Psi(s), R(s)) \\ GCD(\Psi(s), R(s)) = GCD(\Psi(s), \lambda R(s)) \end{cases} \tag{C.2}$$

where λ is a non-zero scalar.

The greatest common divisor (GCD) is then the last non-zero remainder divided by the its leading coefficient (we recall that GCD is a monic polynomial by convention).

Example C.1. *Let $N(s)$ and $\Psi(s)$ be the following polynomials:*

$$\begin{cases} N(s) = s^3 + 6s^2 + 11s + 6 \\ \Psi(s) = s^3 + 4s^2 + s - 6 \end{cases} \tag{C.3}$$

Because $N(s)$ and $\Psi(s)$ have the same degree, we can write:

$$\begin{aligned} N(s) &= s^3 + 4s^2 + s - 6 + \left(2s^2 + 10s + 12\right) \\ &= \Psi(s) + 2s^2 + 10s + 12 = \Psi(s) + R(s) \end{aligned} \tag{C.4}$$

and:

$$\begin{aligned} GCD\left(N(s), \Psi(s)\right) &= GCD\left(\Psi(s), R(s)\right) \\ &= GCD\left(\Psi(s), 2s^2 + 10s + 12\right) \end{aligned} \tag{C.5}$$

Then we make the Euclidean division of $\Psi(s)$ *by* $2s^2 + 10s + 12$*:*

$$\begin{array}{rrrr|l} s^3 & +4s^2 & +s & -6 & 2s^2 + 10s + 12 \\ \cline{5-5} -(s^3 & +5s^2 & +6s) & & \frac{1}{2}s - \frac{1}{2} \\ \cline{1-3} & -s^2 & -5s & & \\ & -(-s^2 & -5s & -6) & \\ \cline{2-4} & & & 0 & \end{array} \tag{C.6}$$

Here the last non-zero remainder is s^2+5s+6*, which is a monic polynomial. We finally get:*

$$GCD\left(N(s), \Psi(s)\right) = s^2 + 5s + 6 \tag{C.7}$$

Finally we can write:

$$\begin{cases} N(s) & = s^3 + 6s^2 + 11s + 6 \\ & = (s+1)\left(s^2 + 5s + 6\right) \\ & = (s+1) \times GCD\left(N(s), \Psi(s)\right) \\ \Psi(s) & = s^3 + 4s^2 + s - 6 \\ & = (s-1)\left(s^2 + 5s + 6\right) \\ & = (s-1) \times GCD\left(N(s), \Psi(s)\right) \end{cases} \tag{C.8}$$

Those relations clearly show that $GCD\left(N(s), \Psi(s)\right)$ *is a factor of both* $N(s)$ *and* $\Psi(s)$*.*

■

C.1.2 Smith form of polynomial matrix

First, let's define unimodular polynomial matrix: a square polynomial matrix $\mathbf{U}(s)$ is unimodular as soon as its inverse is also a polynomial matrix. Equivalently a square polynomial matrix $\mathbf{U}(s)$ is unimodular if and only if its determinant is a nonzero constant independent of s.

Example C.2. *Consider the following square polynomial matrix* $\mathbf{U}(s)$*:*

$$\mathbf{U}(s) = \begin{bmatrix} s & 1 \\ 2s - 3 & 2 \end{bmatrix} \tag{C.9}$$

It is clear that $\det\left(\mathbf{U}(s)\right) = 3$*. Thus the determinant of* $\mathbf{U}(s)$ *is a nonzero constant independent of* s*. As a consequence* $\mathbf{U}(s)$ *as an unimodular polynomial matrix and its inverse is also a polynomial matrix:*

$$\mathbf{U}^{-1}(s) = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -2s + 3 & s \end{bmatrix} \tag{C.10}$$

■

Smith form is a canonical form for polynomial matrices. Let $\mathbf{N}(s)$ be a $p \times m$ polynomial matrix. Smith's theorem states that any polynomial matrix $\mathbf{N}(s)$ can be written as follows:

$$\mathbf{S}(s) = \mathbf{V}(s)\mathbf{N}(s)\mathbf{W}(s) \tag{C.11}$$

where:

- $\mathbf{S}(s)$ is a quasi-diagonal polynomial matrix;
- Square matrices $\mathbf{V}(s)$ and $\mathbf{W}(s)$ are unimodular polynomial matrices;
- The dimension of matrix $\mathbf{V}(s)$ is $p \times p$ whereas the dimension of matrix $\mathbf{W}(s)$ is $m \times m$.

Matrices $\mathbf{N}(s)$ and $\mathbf{S}(s)$ are said to be equivalent.

Quasi-diagonal matrix $\mathbf{S}(s)$ has the following expression, where r is the normal rank of $\mathbf{N}(s)$ (the normal rank of the polynomial matrix $\mathbf{N}(s)$ is the rank of the matrix for almost every value of s).

$$\mathbf{S}(s) = \begin{bmatrix} s_1(s) & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & s_r(s) & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{C.12}$$

Monic polynomials $s_k(s)$ are invariant factors of $\mathbf{N}(s)$. A monic polynomial is a polynomial in which the coefficient of highest degree (also called the leading coefficient) is equal to 1. Furthermore $s_k(s)$ divides $s_{k+1}(s) \, \forall \, k = 1, \ldots, r-1$:

$$s_k(s) \mid s_{k+1}(s) \, \forall \, k = 1, \ldots, r-1 \tag{C.13}$$

Elementary row and column operations are used to transform $\mathbf{N}(s)$ into its Smith form $\mathbf{S}(s)$. The three elementary operations for a polynomial matrix are:

- Multiplying a row or column by a constant;
- Interchanging two rows or two columns; and
- Adding a polynomial multiple of a row or column to another row or column.

The sequences, which are not unique, of elementary row and column operations are the following:

1. By performing suitable row and column swaps, put the polynomial with the lowest degree in first row and first column. Let $a_{11}(s)$ be this polynomial;
2. By Euclidean division expand each polynomial in the first row as follows where $\deg(r_j(s)) < \deg(a_{11}(s))$:

$$a_{1j}(s) = a_{11}(s)q_j(s) + r_j(s) \tag{C.14}$$

Then subtract $q_j(s)$ times the first column from column j and perform suitable row and column swaps to put the polynomial with the lowest degree in first row and first column. This is to be repeated until all elements of the first row are divisible by $a_{11}(s)$;

3. Subtract the appropriate multiple of the first column from columns j where $j > 1$ to get 0 in the first row and columns $j > 1$;

4. Proceed with the same type of elementary operations, but operating on rows rather than columns;

5. Once finalized, the process yields to the following matrix where $s_1(s) \neq 0$ is a monic polynomial:

$$\begin{bmatrix} s_1(s) & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{N}_1(s) & \\ 0 & & & \end{bmatrix} \tag{C.15}$$

If $\mathbf{N}_1(s) = \mathbf{0}$ the process stops. Otherwise the sequence of elementary row and column operations shall continue over $\mathbf{N}_1(s)$ to produce another diagonal element.

From an algorithmic point of view, matrices $\mathbf{V}(s)$ and $\mathbf{W}(s)$ are initialized to identity matrix with an appropriate dimension and are put left and right to $\mathbf{N}(s)$. Then the sequence of elementary operations on row (matrix $\mathbf{V}(s)$) and column (matrix $\mathbf{W}(s)$) are applied until matrix $\mathbf{S}(s)$ is achieved.

Alternatively the expression of $\mathbf{S}(s)$, that is the Smith canonical form of polynomial matrix $\mathbf{N}(s)$, can be obtained through all minors of size k of $\mathbf{N}(s)$ as follows: let $P_0(s) = 1$ and define $P_k(s)$, $k = 1, \ldots, r$ as the greatest common divisor (GCD) of all minors of size k of $\mathbf{N}(s)$. Then:

$$s_k(s) = \frac{P_k(s)}{P_{k-1}(s)}, \qquad k = 1, \ldots, r \tag{C.16}$$

Minors of size k of a matrix are the determinants of the square submatrices of size k. Let $\mathbf{N}(s)$ be a $p \times m$ matrix. Then minors of size k are the determinants of all the $k \times k$ square submatrices obtained from $\mathbf{N}(s)$ by deleting $p - k$ rows and $m - k$ columns. There are a total of $\begin{pmatrix} p \\ k \end{pmatrix} \times \begin{pmatrix} m \\ k \end{pmatrix} = \frac{p!}{k!(p-k)!} \times \frac{m!}{k!(m-k)!}$ minors of size k within $\mathbf{N}(s)$, where $k \leq \min(p, m)$.

Example C.3. *Let* $\mathbf{N}(s)$ *be the following polynomial matrix:*

$$\mathbf{N}(s) = \begin{bmatrix} 4 & -(s+2) \\ 2(s+2) & -0.5 \end{bmatrix} \tag{C.17}$$

The normal rank of $\mathbf{N}(s)$ *is* 2. *Thus the Smith form of* $\mathbf{N}(s)$ *reads:*

$$\mathbf{S}(s) = \begin{bmatrix} s_1(s) & 0 \\ 0 & s_2(s) \end{bmatrix} \tag{C.18}$$

In order to get the Smith form of $\mathbf{N}(s)$ *and express* $\mathbf{S}(s)$ *as* $\mathbf{V}(s)\mathbf{N}(s)\mathbf{W}(s)$ *we apply the sequences of elementary row and column operations (which are not unique) described in Table C.1.*

Sequence	$\mathbf{V}(s)$ (on rows)	$\mathbf{N}(s)$	$\mathbf{W}(s)$ (on columns)
	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 4 & -(s+2) \\ 2(s+2) & -0.5 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$r_1 \leftarrow r_1/4$	$\begin{bmatrix} 1/4 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -(s+2)/4 \\ 2(s+2) & -0.5 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$c_2 \leftarrow c_2 + (s+2)/4\, c_1$	$\begin{bmatrix} 1/4 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 2(s+2) & -0.5+(s+2)^2/2 \end{bmatrix}$	$\begin{bmatrix} 1 & (s+2)/4 \\ 0 & 1 \end{bmatrix}$
$r_2 \leftarrow r_2 - 2(s+2)\, r_1$	$\begin{bmatrix} 1/4 & 0 \\ -(s+2)/2 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & -0.5+(s+2)^2/2 = \frac{s^2+4s+3}{2} \end{bmatrix}$	$\begin{bmatrix} 1 & (s+2)/4 \\ 0 & 1 \end{bmatrix}$
$r_2 \leftarrow 2r_2$	$\begin{bmatrix} 1/4 & 0 \\ -(s+2) & 2 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & s^2+4s+3 \end{bmatrix}$	$\begin{bmatrix} 1 & (s+2)/4 \\ 0 & 1 \end{bmatrix}$

Table C.1: Smith form of $\mathbf{N}(s)$

We can check that $\mathbf{S}(s) = \mathbf{V}(s)\mathbf{N}(s)\mathbf{W}(s)$ *where* $\mathbf{V}(s)$ *and* $\mathbf{W}(s)$ *are unimodular polynomial matrices:*

$$\begin{cases} \mathbf{S}(s) = \begin{bmatrix} 1 & 0 \\ 0 & s^2+4s+3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & (s+1)(s+3) \end{bmatrix} \\ \mathbf{V}(s) = \begin{bmatrix} 1/4 & 0 \\ -(s+2) & 2 \end{bmatrix} \Rightarrow \det(\mathbf{V}(s)) = \frac{1}{2} \\ \mathbf{W}(s) = \begin{bmatrix} 1 & (s+2)/4 \\ 0 & 1 \end{bmatrix} \Rightarrow \det(\mathbf{W}(s)) = 1 \end{cases} \tag{C.19}$$

Alternatively the expression of $\mathbf{S}(s)$*, that is the Smith form of polynomial matrix* $\mathbf{N}(s)$*, can be obtained through all minors of size* k *of* $\mathbf{N}(s)$*.*

- *We start with:*

$$P_0(s) = 1 \tag{C.20}$$

- *Minors of size* 1 *of* $\mathbf{N}(s)$ *are the following:*

$$4, \quad -(s+2), \quad 2(s+2), \quad -0.5 \tag{C.21}$$

Thus the greatest common divisor (GCD) of all minors of size 1 *of* $\mathbf{N}(s)$ *is* -0.5*. Because this is not a monic polynomial, we divide it by its leading coefficient, that is* -0.5 *here. We get:*

$$P_1(s) = 1 \tag{C.22}$$

- *The minor of size* 2 *of* $\mathbf{N}(s)$ *is the following:*

$$\det\left(\begin{bmatrix} 4 & -(s+2) \\ 2(s+2) & -0.5 \end{bmatrix}\right) = -2 + 2(s+2)^2 = 2(s^2+4s+3) \tag{C.23}$$

Thus the greatest common divisor (GCD) of all minors of size 2 *of* $\mathbf{N}(s)$ *is* $2(s^2+4s+3)$*. Because this is not a monic polynomial, we divide it by its leading coefficient, that is* 2 *here. We get:*

$$P_2(s) = s^2 + 4s + 3 \tag{C.24}$$

We finally get:

$$s_k(s) = \frac{P_k(s)}{P_{k-1}(s)}, \qquad k = 1,2 \Rightarrow \begin{cases} s_1(s) = \frac{P_1(s)}{P_0(s)} = 1 \\ s_2(s) = \frac{P_2(s)}{P_1(s)} = s^2 + 4s + 3 \end{cases} \tag{C.25}$$

The Smith form of $\mathbf{N}(s)$ *is thus the following matrix* $\mathbf{S}(s)$*:*

$$\mathbf{S}(s) = \begin{bmatrix} s_1(s) & 0 \\ 0 & s_2(s) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & s^2+4s+3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & (s+1)(s+3) \end{bmatrix} \tag{C.26}$$

■

Example C.4. *Let* $\mathbf{N}(s)$ *be the following polynomial matrix*[1]*:*

$$\mathbf{N}(s) = \begin{bmatrix} 1 & -1 \\ s^2+s-4 & 2s^2-s-8 \\ s^2-4 & 2s^2-8 \end{bmatrix} \tag{C.27}$$

The normal rank of $\mathbf{N}(s)$ *is* 2*. Thus the Smith form of* $\mathbf{N}(s)$ *reads:*

$$\mathbf{S}(s) = \begin{bmatrix} s_1(s) & 0 \\ 0 & s_2(s) \\ 0 & 0 \end{bmatrix} \tag{C.28}$$

In order to get the Smith form of $\mathbf{N}(s)$ *and express* $\mathbf{S}(s)$ *as* $\mathbf{V}(s)\mathbf{N}(s)\mathbf{W}(s)$ *we apply the sequences of elementary row and column operations (which are not unique) described in Table C.2.*

We can check that $\mathbf{S}(s) = \mathbf{V}(s)\mathbf{N}(s)\mathbf{W}(s)$ *where* $\mathbf{V}(s)$ *and* $\mathbf{W}(s)$ *are unimodular polynomial matrices:*

$$\begin{cases} \mathbf{S}(s) = \begin{bmatrix} 1 & 0 \\ 0 & s^2-4 \\ 0 & 0 \end{bmatrix} \\ \mathbf{V}(s) = \begin{bmatrix} 1/3 & 0 & 0 \\ -s/3 & 1/3 & 0 \\ s & -1 & 1 \end{bmatrix} \Rightarrow \det(\mathbf{V}(s)) = \frac{1}{9} \\ \mathbf{W}(s) = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \Rightarrow \det(\mathbf{W}(s)) = 3 \end{cases} \tag{C.29}$$

Alternatively the expression of $\mathbf{S}(s)$*, that is the Smith form of polynomial matrix* $\mathbf{N}(s)$*, can be obtained through all minors of size* k *of* $\mathbf{N}(s)$*.*

- *We start with:*

$$P_0(s) = 1 \tag{C.30}$$

- *Minors of size* 1 *of* $\mathbf{N}(s)$ *are the following:*

$$1, \quad -1, \quad s^2+s-4, \quad 2s^2-s-8, \quad s^2-4, \quad 2s^2-8 \tag{C.31}$$

Thus the greatest common divisor (GCD) of all minors of size 1 *of* $\mathbf{N}(s)$ *is:*

$$P_1(s) = 1 \tag{C.32}$$

[1] Multivariable Feedback Design (Addison-Wesley) J.M. Maciejowski (1989)

Sequence	$\mathbf{V}(s)$ (on rows)	$\mathbf{N}(s)$	$\mathbf{W}(s)$ (on columns)
	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -1 \\ s^2+s-4 & 2s^2-s-8 \\ s^2-4 & 2s^2-8 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$c_2 \leftarrow c_2 + c_1$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ s^2+s-4 & 3s^2-12 \\ s^2-4 & 3s^2-12 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$
$c_1 \leftarrow 3c_1 - c_2$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 3 & 0 \\ 3s & 3s^2-12 \\ 0 & 3s^2-12 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$
$r_2 \leftarrow r_2 - sr_1$	$\begin{bmatrix} 1 & 0 & 0 \\ -s & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 3 & 0 \\ 0 & 3s^2-12 \\ 0 & 3s^2-12 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$
$r_1 \leftarrow r_1/3$	$\begin{bmatrix} 1/3 & 0 & 0 \\ -s & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 3s^2-12 \\ 0 & 3s^2-12 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$
$r_3 \leftarrow r_3 - r_2$	$\begin{bmatrix} 1/3 & 0 & 0 \\ -s & 1 & 0 \\ s & -1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 3s^2-12 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$
$r_2 \leftarrow r_2/3$	$\begin{bmatrix} 1/3 & 0 & 0 \\ -s/3 & 1/3 & 0 \\ s & -1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & s^2-4 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$

Table C.2: Smith form of $\mathbf{N}(s)$

– *Minors of size* 2 *of* $\mathbf{N}(s)$ *are the following:*

$$\begin{cases} \det\left(\begin{bmatrix} 1 & -1 \\ s^2+s-4 & 2s^2-s-8 \end{bmatrix}\right) = 3s^2-12 \\ \det\left(\begin{bmatrix} 1 & -1 \\ s^2-4 & 2s^2-8 \end{bmatrix}\right) = 3s^2-12 \\ \det\left(\begin{bmatrix} s^2+s-4 & 2s^2-s-8 \\ s^2-4 & 2s^2-8 \end{bmatrix}\right) = s\left(3s^2-12\right) \end{cases} \tag{C.33}$$

Thus the greatest common divisor (GCD) of all minors of size 2 *of* $\mathbf{N}(s)$ *is* $3s^2-12$. *Because this is not a monic polynomial, we divide it by its leading coefficient, that is* 3 *here. We get:*

$$P_2(s) = \frac{3s^2-12}{3} = s^2-4 \tag{C.34}$$

We finally get:

$$s_k(s) = \frac{P_k(s)}{P_{k-1}(s)}, \qquad k = 1,2 \Rightarrow \begin{cases} s_1(s) = \frac{P_1(s)}{P_0(s)} = 1 \\ s_2(s) = \frac{P_2(s)}{P_1(s)} = s^2-4 \end{cases} \tag{C.35}$$

The Smith form of $\mathbf{N}(s)$ *is thus the following matrix* $\mathbf{S}(s)$*:*

$$\mathbf{S}(s) = \begin{bmatrix} s_1(s) & 0 \\ 0 & s_2(s) \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & s^2-4 \\ 0 & 0 \end{bmatrix} \tag{C.36}$$

■

C.1.3 Smith-McMillan form of transfer matrix

The Smith-McMillan form of a transfer matrix can be used to generalize the notion of poles and zeros of a transfer matrix with multiple inputs and/or outputs.

We consider the following transfer matrix $\mathbf{F}(s)$ where $\mathbf{N}(s)$ is a polynomial matrix and $\Psi(s)$ is a polynomial which is actually the least common multiple of the denominators of all the elements in $\mathbf{F}(s)$):

$$\mathbf{F}(s) = \frac{\mathbf{N}(s)}{\Psi(s)} \tag{C.37}$$

We will assume that $\frac{\mathbf{N}(s)}{\Psi(s)}$ is proper, meaning that all the polynomials which appear in $\mathbf{N}(s)$ have a degree lower or equal to the degree of $\Psi(s)$:

$$\lim_{s\to\infty} \mathbf{F}(s) < \infty \tag{C.38}$$

Using the Smith form (C.12) of polynomial matrix $\mathbf{N}(s)$ we get:

$$\begin{aligned} \mathbf{S}(s) &= \begin{bmatrix} s_1(s) & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & s_r(s) & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix} \\ \Rightarrow \frac{\mathbf{S}(s)}{\Psi(s)} &= \begin{bmatrix} \frac{s_1(s)}{\Psi(s)} & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & \frac{s_r(s)}{\Psi(s)} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix} \end{aligned} \tag{C.39}$$

Matrix $\frac{\mathbf{S}(s)}{\Psi(s)}$ is called the Smith-McMillan form of transfer matrix $\mathbf{F}(s)$: this is a quasi-diagonal matrix of rational functions which may possibly simplify. Let $\frac{\varepsilon_i(s)}{\psi_i(s)}$ be the rational functions obtained after simplification of each element of quasi-diagonal matrix $\frac{\mathbf{S}(s)}{\Psi(s)}$. Then the Smith-McMillan form of transfer matrix $\mathbf{F}(s)$ reads:

$$\frac{\mathbf{S}(s)}{\Psi(s)} = \begin{bmatrix} \frac{\varepsilon_1(s)}{\psi_1(s)} & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & \frac{\varepsilon_r(s)}{\psi_r(s)} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{C.40}$$

Polynomials $\varepsilon_k(s)$ and $\psi_k(s)$ are coprime. Furthermore $\varepsilon_k(s)$ divides $\varepsilon_{k+1}(s)$ and $\psi_{k+1}(s)$ divides $\psi_k(s)\,\forall\, k = 1, \ldots, r-1$:

$$\left.\begin{array}{l} \varepsilon_k(s) \,|\, \varepsilon_{k+1}(s) \\ \psi_{k+1}(s) \,|\, \psi_k(s) \end{array}\right\} \forall\, k = 1, \ldots, r-1 \tag{C.41}$$

Example C.5. *Let* $\mathbf{F}(s)$ *be the following transfer matrix:*

$$\begin{aligned}\mathbf{F}(s) &= \begin{bmatrix} \frac{1}{(s+1)(s+2)} & \frac{-1}{(s+1)(s+2)} \\ \frac{s^2+s-4}{(s+1)(s+2)} & \frac{2s^2-s-8}{(s+1)(s+2)} \\ \frac{s-2}{s+1} & \frac{2s-4}{s+1} \end{bmatrix} \\ &= \frac{1}{(s+1)(s+2)} \begin{bmatrix} 1 & -1 \\ s^2+s-4 & 2s^2-s-8 \\ s^2-4 & 2s^2-8 \end{bmatrix} \\ &:= \frac{\mathbf{N}(s)}{\Psi(s)}\end{aligned} \tag{C.42}$$

The Smith form of polynomial matrix $\mathbf{N}(s)$ *which appears in the numerator of transfer matrix* $\mathbf{F}(s)$ *has been achieved in a previous example. Thus the Smith-McMillan form of transfer matrix* $\mathbf{F}(s)$ *reads as follows where we can notice some simplifications in the rational functions within the quasi-diagonal matrix* $\frac{\mathbf{S}(s)}{\Psi(s)}$*:*

$$\begin{aligned}\frac{\mathbf{S}(s)}{\Psi(s)} &= \frac{1}{(s+1)(s+2)} \begin{bmatrix} 1 & 0 \\ 0 & s^2-4 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{(s+1)(s+2)} & 0 \\ 0 & \frac{(s-2)(s+2)}{(s+1)(s+2)} \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{(s+1)(s+2)} & 0 \\ 0 & \frac{s-2}{s+1} \\ 0 & 0 \end{bmatrix}\end{aligned} \tag{C.43}$$

■

Example C.6. *Let* $\mathbf{F}(s) := \frac{\mathbf{N(s)}}{\Psi(s)}$ *be the following transfer matrix:*

$$\mathbf{F}(s) = \begin{bmatrix} \frac{2}{s+2} & \frac{s+1}{s+3} \\ \frac{1}{s+2} & \frac{5}{s+2} \end{bmatrix} = \frac{1}{(s+2)(s+3)} \begin{bmatrix} 2(s+3) & (s+1)(s+2) \\ s+3 & 5(s+3) \end{bmatrix} \tag{C.44}$$

Let $\mathbf{N}(s)$ *be the following polynomial matrix:*

$$\mathbf{N}(s) = \begin{bmatrix} 2(s+3) & (s+1)(s+2) \\ s+3 & 5(s+3) \end{bmatrix} = \begin{bmatrix} 2s+6 & s^2+3s+2 \\ s+3 & 5s+15 \end{bmatrix} \tag{C.45}$$

The normal rank of $\mathbf{N}(s)$ *is* 2*. Thus the Smith form of* $\mathbf{N}(s)$ *reads:*

$$\mathbf{S}(s) = \begin{bmatrix} s_1(s) & 0 \\ 0 & s_2(s) \end{bmatrix} \tag{C.46}$$

In order to get the Smith form of $\mathbf{N}(s)$ *and express* $\mathbf{S}(s)$ *as* $\mathbf{V}(s)\mathbf{N}(s)\mathbf{W}(s)$ *we apply the sequences of elementary row and column operations (which are not unique) described in Table C.3.*

Séquence	$\mathbf{V}(s)$ (lignes)	$\mathbf{N}(s)$	$\mathbf{W}(s)$ (colonnes)
	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 2s+6 & s^2+3s+2 \\ s+3 & 5s+15 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$r_1 \leftarrow r_1/2$	$\begin{bmatrix} 0.5 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} s+3 & s^2/2+1.5s+1 \\ s+3 & 5s+15 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$r_2 \leftarrow r_2 - r_1$	$\begin{bmatrix} 0.5 & 0 \\ -0.5 & 1 \end{bmatrix}$	$\begin{bmatrix} s+3 & s^2/2+1.5s+1 \\ 0 & -s^2/2+3.5s+14 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$c_2 \leftarrow c_2 - 0.5\, s\, c_1$	$\begin{bmatrix} 0.5 & 0 \\ -0.5 & 1 \end{bmatrix}$	$\begin{bmatrix} s+3 & 1 \\ 0 & -s^2/2+3.5s+14 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.5s \\ 0 & 1 \end{bmatrix}$
$c_2 \leftrightarrow c_1$	$\begin{bmatrix} 0.5 & 0 \\ -0.5 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & s+3 \\ -s^2/2+3.5s+14 & 0 \end{bmatrix}$	$\begin{bmatrix} -0.5s & 1 \\ 1 & 0 \end{bmatrix}$
$r_2 \leftarrow p(s)\, r_1 - r_2$ $p(s) = -s^2/2 + 3.5s + 14$	$\begin{bmatrix} 0.5 & 0 \\ -0.25s^2+1.75s+7.5 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & s+3 \\ 0 & (-s^2/2+3.5s+14)(s+3) \end{bmatrix}$	$\begin{bmatrix} -0.5s & 1 \\ 1 & 0 \end{bmatrix}$
$c_2 \leftarrow (s+3)c_1 - c_2$	$\begin{bmatrix} 0.5 & 0 \\ -0.25s^2+1.75s+7.5 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & (s^2/2-3.5s-14)(s+3) \end{bmatrix}$	$\begin{bmatrix} -0.5s & -s^2/2-1.5s-1 \\ 1 & s+3 \end{bmatrix}$
$c_2 \leftarrow 2\, c_2$	$\begin{bmatrix} 0.5 & 0 \\ -0.25s^2+1.75s+7.5 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & (s^2-7s-28)(s+3) \end{bmatrix}$	$\begin{bmatrix} -0.5s & -s^2-3s-2 \\ 1 & 2(s+3) \end{bmatrix}$

Table C.3: Forme de Smith de $\mathbf{N}(s)$

We can check that $\mathbf{S}(s) = \mathbf{V}(s)\mathbf{N}(s)\mathbf{W}(s)$ *where* $\mathbf{V}(s)$ *and* $\mathbf{W}(s)$ *are unimodular polynomial matrices:*

$$\begin{cases} \mathbf{S}(s) = \begin{bmatrix} 1 & 0 \\ 0 & (s^2-7s-28)(s+3) \end{bmatrix} \\ \mathbf{V}(s) = \begin{bmatrix} 0.5 & 0 \\ -0.25s^2+1.75s+7.5 & -1 \end{bmatrix} \Rightarrow \det(\mathbf{V}(s)) = -0.5 \\ \mathbf{W}(s) = \begin{bmatrix} -0.5s & -s^2-3s-2 \\ 1 & 2(s+3) \end{bmatrix} \Rightarrow \det(\mathbf{W}(s)) = 2 \end{cases} \tag{C.47}$$

Alternatively the expression of $\mathbf{S}(s)$*, that is the Smith form of polynomial matrix* $\mathbf{N}(s)$*, can be obtained through all minors of size* k *of* $\mathbf{N}(s)$*:*

- *We start with:*

$$P_0(s) = 1 \tag{C.48}$$

- *Minors of size* 1 *of* $\mathbf{N}(s)$ *are the following:*

$$2(s+3), \quad (s+1)(s+2), \quad s+3, \quad 5(s+3) \tag{C.49}$$

  Thus the greatest common divisor (GCD) of all minors of size 1 *of* $\mathbf{N}(s)$ *is:*

$$P_1(s) = 1 \tag{C.50}$$

- *The only minor of size* 2 *of* $\mathbf{N}(s)$ *is the following:*

$$\begin{aligned} \det(\mathbf{N}(s)) &= \det\left(\begin{bmatrix} 2(s+3) & (s+1)(s+2) \\ s+3 & 5(s+3) \end{bmatrix}\right) \\ &= 10(s+3)^2 - (s+1)(s+2)(s+3) \end{aligned} \tag{C.51}$$

  Thus the greatest common divisor (GCD) of all minors of size 2 *is obviously* $\det(\mathbf{N}(s)) = 10(s+3)^2 - (s+1)(s+2)(s+3)$ *is* $3s^2 - 12$*. Because this is not a monic polynomial, we divide it by its leading coefficient, that is* -1 *here. We get:*

$$P_2(s) = (s+1)(s+2)(s+3) - 10(s+3)^2 \tag{C.52}$$

We finally get:

$$\begin{aligned} & s_k(s) = \tfrac{P_k(s)}{P_{k-1}(s)}, \qquad k = 1, 2 \\ & \Rightarrow \begin{cases} s_1(s) = \frac{P_1(s)}{P_0(s)} = 1 \\ s_2(s) = \frac{P_2(s)}{P_1(s)} = (s+1)(s+2)(s+3) - 10(s+3)^2 \end{cases} \end{aligned} \tag{C.53}$$

Consequently the Smith form of polynomial matrix $\mathbf{N}(s)$ *is the following:*

$$\begin{aligned} \mathbf{S}(s) = \begin{bmatrix} s_1(s) & 0 \\ 0 & s_2(s) \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & (s+1)(s+2)(s+3) - 10(s+3)^2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & (s^2 - 7s - 28)(s+3) \end{bmatrix} \end{aligned} \tag{C.54}$$

Consequently the Smith-McMillan form of transfer matrix $\mathbf{F}(s)$ *is obtained when dividing polynomial matrix* $\mathbf{S}(s)$ *by polynomial* $\Psi(s) = (s+2)(s+3)$*:*

$$\begin{aligned} \frac{\mathbf{S}(s)}{\Psi(s)} &= \tfrac{1}{(s+2)(s+3)} \begin{bmatrix} 1 & 0 \\ 0 & (s^2 - 7s - 28)(s+3) \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{(s+2)(s+3)} & 0 \\ 0 & \frac{s^2-7s-28}{s+2} \end{bmatrix} \\ &:= \begin{bmatrix} \frac{\varepsilon_1(s)}{\psi_1(s)} & 0 \\ 0 & \frac{\varepsilon_2(s)}{\psi_2(s)} \end{bmatrix} \end{aligned} \tag{C.55}$$

■

C.1.4 System dimension, poles and zeros

As soon as the Smith-McMillan form (C.40) of transfer matrix $\mathbf{F}(s)$ is obtained, the following results can be applied:

- System dimension n is obtained as the sum of the degrees of invariant factors $\psi_k(s)$:

$$n = \sum_{k=1}^{r} \deg\left(\psi_k(s)\right) \tag{C.56}$$

- The characteristic polynomial of the system is defined as follows:

$$\chi_A(s) = \prod_{k=1}^{r} \psi_k(s) \tag{C.57}$$

- Poles of $\mathbf{F}(s)$ are the roots of polynomial $\chi_A(s) = \prod_{k=1}^{r} \psi_k(s)$;
- Transmission zeros of $\mathbf{F}(s)$ are the roots of polynomial $\prod_{k=1}^{r} \varepsilon_k(s)$;
- Poles and zeros at infinity are obtained by evaluating poles and zeros of the Smith-McMillan form of transfer matrix $\mathbf{F}\left(\frac{1}{s}\right) := \frac{\overline{\mathbf{N}}(s)}{d(s)}$ where $d(s)$ is

the least common multiple of all elements of $\mathbf{F}\left(\frac{1}{s}\right)$ and where $\mathbf{V}(s)$ and $\mathbf{W}(s)$ are some unimodular matrices[2]:

$$\mathbf{F}\left(\tfrac{1}{s}\right) := \frac{\overline{\mathbf{N}}(s)}{d(s)} \Rightarrow \begin{cases} \overline{\mathbf{N}}(s) = \mathbf{V}(s)\overline{\mathbf{S}}(s)\mathbf{W}(s) \\ \overline{\mathbf{S}}(s) = \begin{bmatrix} s_1(s) & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & s_r(s) & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix} \end{cases}$$

$$\Rightarrow \mathbf{F}\left(\tfrac{1}{s}\right) = \mathbf{V}(s)\frac{\overline{\mathbf{S}}(s)}{d(s)}\mathbf{W}(s) = \mathbf{V}(s)\begin{bmatrix} \frac{\psi_1(s)}{\varepsilon_1(s)} & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & \frac{\psi_r(s)}{\varepsilon_r(s)} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{W}(s) \tag{C.58}$$

Polynomials $\varepsilon_k(s)$ and $\psi_k(s)$ are coprime. Since $s_k(s)$ divides $s_{k+1}(s)$, $k = 1, \cdots, r-1$, polynomial $\psi_k(s)$ divides $\psi_{k+1}(s)$ and $\varepsilon_{k+1}(s)$ divides $\varepsilon_k(s)\,\forall\, k = 1, \ldots, r-1$:

$$\left.\begin{array}{c} \psi_k(s) \,|\, \psi_{k+1}(s) \\ \varepsilon_{k+1}(s) \,|\, \varepsilon_k(s) \end{array}\right\} \forall\, k = 1, \ldots, r-1 \tag{C.59}$$

Furthermore we said that $\mathbf{F}(s)$ has a pole at infinity with multiplicity ν_p if $s = 0$ is a root of $\prod_{k=1}^{r} \varepsilon_k(s)$ with multiplicity ν_p. Similarly, we said that $\mathbf{F}(s)$ has a zero at infinity with multiplicity ν_z if $s = 0$ is a root of $\prod_{k=1}^{r} \psi_k(s)$ with multiplicity ν_z.

Example C.7. *Let $\mathbf{F}(s)$ be the following transfer matrix:*

$$\mathbf{F}(s) = \begin{bmatrix} \frac{2}{s+2} & \frac{s+1}{s+3} \\ \frac{1}{s+2} & \frac{5}{s+2} \end{bmatrix} \tag{C.60}$$

We have already seen that the Smith-McMillan form of transfer matrix $\mathbf{F}(s)$ reads as follows:

$$\frac{\mathbf{S}(s)}{\Psi(s)} = \begin{bmatrix} \frac{\varepsilon_1(s)}{\psi_1(s)} & 0 \\ 0 & \frac{\varepsilon_2(s)}{\psi_2(s)} \end{bmatrix} = \begin{bmatrix} \frac{1}{(s+2)(s+3)} & 0 \\ 0 & \frac{s^2-7s-28}{s+2} \end{bmatrix} \tag{C.61}$$

Consequently we have the following results:

- *System dimension n is obtained as the sum of the degrees of invariant factors $\psi_i(s)$:*

$$n = \sum_{i=1}^{r} \deg\left(\psi_i(s)\right) = 2 + 1 = 3 \tag{C.62}$$

System dimension n is the dimension of a minimal state-space realization of $\mathbf{F}(s)$;

[2] Juan C. Zuniga–Anaya, Structural Properties of Polynomial and Rational Matrices, a Survey, Mathematica Aeterna, Vol. 1, 2011, no. 06, 361 - 403

– *The characteristic polynomial of the system is defined as follows:*

$$\chi_A(s) = \prod_{i=1}^{r} \psi_i(s) = (s+2)^2(s+3) \tag{C.63}$$

– *Poles of* $\mathbf{F}(s)$ *are the roots of polynomials* $\psi_i(s)$*, that are here* -2*, with multiplicity* 2*, and* -3*;*

– *Transmission of* $\mathbf{F}(s)$ *are the roots of polynomials* $\varepsilon_i(s)$*, , that are here the roots of* $s^2 - 7s - 28$*.*

■

C.1.5 Transfer matrix realization

We start with the following relation:

$$\underline{Y}(s) = \mathbf{F}(s)\underline{U}(s) \tag{C.64}$$

where:

$$\begin{cases} \mathbf{F}(s) = \frac{\mathbf{N}(s)}{\Psi(s)} \\ \mathbf{S}(s) = \mathbf{V}(s)\mathbf{N}(s)\mathbf{W}(s) \end{cases} \Rightarrow \mathbf{F}(s) = \mathbf{V}^{-1}(s)\frac{\mathbf{S}(s)}{\Psi(s)}\mathbf{W}^{-1}(s) \tag{C.65}$$

For a system with p outputs and m inputs, we split the square unimodular polynomial matrices $\mathbf{V}(s)$ (of dimension $p \times p$) and $\mathbf{W}(s)$ (of dimension $m \times m$) as follows:

$$\begin{cases} \mathbf{V}^{-1}(s) = \begin{bmatrix} \underline{v}_1(s) & \dots & \underline{v}_p(s) \end{bmatrix} \\ \mathbf{W}^{-1}(s) = \begin{bmatrix} \underline{w}_1^T(s) \\ \vdots \\ \underline{w}_m^T(s) \end{bmatrix} \end{cases} \tag{C.66}$$

Then we get:

$$\mathbf{F}(s) = \sum_{i=1}^{r} \frac{\varepsilon_i(s)}{\psi_i(s)} \underline{v}_i(s)\underline{w}_i^T(s) \tag{C.67}$$

Let $\mathbf{F}_i(s)$ be the following multiple-input multiple-output elementary transfer matrix of dimension $p \times m$:

$$\mathbf{F}_i(s) = \frac{\varepsilon_i(s)}{\psi_i(s)} \underline{v}_i(s)\underline{w}_i^T(s) \tag{C.68}$$

Let $\mathbf{A}_i$ be a companion matrix of polynomial $\psi_i(s)$. Then the characteristic polynomial of $\mathbf{A}_i$ is $\psi_i(s)$:

$$\det(s\mathbb{I} - \mathbf{A}_i) = \psi_i(s) \tag{C.69}$$

Using the results dedicated to state-space realizations of transfer functions, we can achieve the following realization of $\mathbf{F}_i(s)$:

$$\begin{aligned}
&\mathbf{F}_i(s) = \tfrac{\varepsilon_i(s)}{\psi_i(s)}\,\underline{v}_i(s)\underline{w}_i^T(s) = \mathbf{C}_i\,(s\mathbb{I}-\mathbf{A}_i)^{-1}\,\mathbf{B}_i + \mathbf{D}_i := \left(\begin{array}{c|c}\mathbf{A}_i & \mathbf{B}_i\\ \hline \mathbf{C}_i & \mathbf{D}_i\end{array}\right)\\
&\Leftrightarrow \begin{cases}\dot{\underline{x}}_i(t) = \mathbf{A}_i\underline{x}_i(t) + \mathbf{B}_i\underline{u}(t)\\ \underline{y}_i(t) = \mathbf{C}_i\underline{x}_i(t) + \mathbf{D}_i\underline{u}(t)\end{cases}
\end{aligned} \tag{C.70}$$

For example matrix $\mathbf{A}_i$ could be a block diagonal Jordan matrix. Then matrices $\mathbf{C}_i$ and $\mathbf{B}_i$ are constant matrices of dimension $p \times n_i$ and $n_i \times m$, respectively, where n_i is the degree of polynomial $\psi_i(s)$. Constant matrices $\mathbf{C}_i$ and $\mathbf{B}_i$ shall be computed such that:

$$(s\mathbb{I}-\mathbf{A}_i)^{-1} = \frac{\operatorname{adj}(s\mathbb{I}-\mathbf{A}_i)}{\psi_i(s)} \Rightarrow \mathbf{C}_i\operatorname{adj}(s\mathbb{I}-\mathbf{A}_i)\,\mathbf{B}_i = \varepsilon_i(s)\,\underline{v}_i(s)\underline{w}_i^T(s) \tag{C.71}$$

Furthermore $\mathbf{D}_i = \lim_{s\to\infty}\mathbf{F}_i(s)$. Other realizations will be presented in Section C.1.6.

A realization of $\mathbf{F}(s)$ is then obtained by aggregating the realizations of all elementary transfer matrices $\mathbf{F}_i(s)$:

$$\begin{cases}
\mathbf{\Lambda} = \begin{bmatrix}\mathbf{A}_1 & & & \mathbf{0}\\ & \mathbf{A}_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{A}_r\end{bmatrix}\\
\mathbf{B} = \begin{bmatrix}\mathbf{B}_1\\ \vdots\\ \mathbf{B}_r\end{bmatrix}\\
\mathbf{C} = \begin{bmatrix}\mathbf{C}_1 & \cdots & \mathbf{C}_r\end{bmatrix}\\
\mathbf{D} = \sum_i \mathbf{D}_i
\end{cases} \tag{C.72}$$

This minimal realization is a generalization of Gilbert's method.

Example C.8. *Let* $\mathbf{F}(s)$ *be the following transfer matrix:*

$$\begin{aligned}
\mathbf{F}(s) &= \begin{bmatrix}\frac{1}{(s+1)(s+2)} & \frac{-1}{(s+1)(s+2)}\\ \frac{s^2+s-4}{(s+1)(s+2)} & \frac{2s^2-s-8}{(s+1)(s+2)}\\ \frac{s-2}{s+1} & \frac{2s-4}{s+1}\end{bmatrix}\\
&= \tfrac{1}{(s+1)(s+2)}\begin{bmatrix}1 & -1\\ s^2+s-4 & 2s^2-s-8\\ s^2-4 & 2s^2-8\end{bmatrix}\\
&:= \tfrac{\mathbf{N}(s)}{\Psi(s)}
\end{aligned} \tag{C.73}$$

We have seen in a previous example that the Smith-McMillan form of transfer

matrix $\mathbf{F}(s)$ *reads as follows:*

$$\begin{aligned}
\mathbf{F}(s) &= \mathbf{V}^{-1}(s)\tfrac{\mathbf{S}(s)}{\Psi(s)}\mathbf{W}^{-1}(s) \\
&= \mathbf{V}^{-1}(s)\tfrac{1}{(s+1)(s+2)}\begin{bmatrix} 1 & 0 \\ 0 & s^2-4 \\ 0 & 0 \end{bmatrix}\mathbf{W}^{-1}(s) \\
&= \begin{bmatrix} \underline{v}_1(s) & \underline{v}_2(s) & \underline{v}_3(s) \end{bmatrix}\begin{bmatrix} \frac{1}{(s+1)(s+2)} & 0 \\ 0 & \frac{s-2}{s+1} \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \underline{w}_1^T(s) \\ \underline{w}_2^T(s) \end{bmatrix} \\
&= \tfrac{1}{(s+1)(s+2)}\,\underline{v}_1(s)\underline{w}_1^T(s) + \tfrac{s-2}{s+1}\,\underline{v}_2(s)\underline{w}_2^T(s)
\end{aligned} \tag{C.74}$$

where:

$$\begin{cases}
\mathbf{V}(s) = \begin{bmatrix} 1/3 & 0 & 0 \\ -s/3 & 1/3 & 0 \\ s & -1 & 1 \end{bmatrix} \Rightarrow \mathbf{V}^{-1}(s) := \begin{bmatrix} \underline{v}_1(s) & \underline{v}_2(s) & \underline{v}_3(s) \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 3s & 3 & 0 \\ 0 & 3 & 1 \end{bmatrix} \\
\mathbf{W}(s) = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \Rightarrow \mathbf{W}^{-1}(s) := \begin{bmatrix} \underline{w}_1^T(s) \\ \underline{w}_2^T(s) \end{bmatrix} = \begin{bmatrix} 1/3 & -1/3 \\ 1/3 & 2/3 \end{bmatrix}
\end{cases} \tag{C.75}$$

Then we get:

$$\mathbf{F}(s) = \frac{1}{(s+1)(s+2)}\begin{bmatrix} 1 & -1 \\ s & -s \\ 0 & 0 \end{bmatrix} + \frac{s-2}{s+1}\begin{bmatrix} 0 & 0 \\ 1 & 2 \\ 1 & 2 \end{bmatrix} \tag{C.76}$$

Let $\mathbf{F}(s) = \mathbf{F}_1(s) + \mathbf{F}_2(s)$ *where the elementary transfer matrices* $\mathbf{F}_1(s)$ *and* $\mathbf{F}_2(s)$ *read as follows:*

$$\begin{cases}
\mathbf{F}_1(s) = \frac{1}{(s+1)(s+2)}\begin{bmatrix} 1 & -1 \\ s & -s \\ 0 & 0 \end{bmatrix} = \frac{1}{s+1}\begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 0 & 0 \end{bmatrix} + \frac{1}{s+2}\begin{bmatrix} -1 & 1 \\ 2 & -2 \\ 0 & 0 \end{bmatrix} \\
\mathbf{F}_2(s) = \frac{s-2}{s+1}\begin{bmatrix} 0 & 0 \\ 1 & 2 \\ 1 & 2 \end{bmatrix} = \left(1 - \frac{3}{s+1}\right)\begin{bmatrix} 0 & 0 \\ 1 & 2 \\ 1 & 2 \end{bmatrix}
\end{cases} \tag{C.77}$$

Elementary transfer matrices $\mathbf{F}_1(s)$ *and* $\mathbf{F}_2(s)$ *are such that all the factors of* $\frac{1}{s-\lambda_i}$ *are* 3×2 *matrices with rank* $n_i = 1$. *Those matrices can then be written as the product between two constant matrices* $\mathbf{C}_i$ *and* $\mathbf{B}_i$ *of dimension* 3×1 *and* 1×2, *respectively:*

$$\begin{cases}
\mathbf{F}_1(s) = \frac{1}{s+1}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\begin{bmatrix} 1 & -1 \end{bmatrix} + \frac{1}{s+2}\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}\begin{bmatrix} 1 & -1 \end{bmatrix} \\
\mathbf{F}_2(s) = \frac{1}{s+1}\begin{bmatrix} 0 \\ -3 \\ -3 \end{bmatrix}\begin{bmatrix} 1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 2 \\ 1 & 2 \end{bmatrix}
\end{cases} \tag{C.78}$$

The realization of $\mathbf{F}_1(s)$ *and* $\mathbf{F}_2(s)$ *using block diagonal Jordan form is then*

the following:

$$\begin{cases} \mathbf{F}_1(s) = \left(\begin{array}{c|c} \mathbf{A}_1 & \mathbf{B}_1 \\ \hline \mathbf{C}_1 & \mathbf{D}_1 \end{array}\right) = \left(\begin{array}{cc|cc} -1 & 0 & 1 & -1 \\ 0 & -2 & 1 & -1 \\ \hline 1 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \\ \mathbf{F}_2(s) = \left(\begin{array}{c|cc} \mathbf{A}_2 & \mathbf{B}_2 \\ \hline \mathbf{C}_2 & \mathbf{D}_2 \end{array}\right) = \left(\begin{array}{c|cc} -1 & 1 & 2 \\ \hline 0 & 0 & 0 \\ -3 & 1 & 1 \\ -3 & 2 & 2 \end{array}\right) \end{cases} \tag{C.79}$$

A minimal realization of $\mathbf{F}(s)$ *is then obtained by aggregating the realizations of all elementary transfer matrices* $\mathbf{F}_i(s)$*:*

$$\begin{cases} \mathbf{\Lambda} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix} = \left[\begin{array}{cc:c} -1 & 0 & 0 \\ 0 & -2 & 0 \\ \hdashline 0 & 0 & -1 \end{array}\right] \\ \mathbf{B} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} = \left[\begin{array}{cc} 1 & -1 \\ 1 & -1 \\ \hdashline 1 & 2 \end{array}\right] \\ \mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix} = \left[\begin{array}{cc:c} 1 & -1 & 0 \\ -1 & 2 & -3 \\ 0 & 0 & -3 \end{array}\right] \\ \mathbf{D} = \sum_i \mathbf{D}_i = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 2 & 2 \end{bmatrix} \end{cases} \tag{C.80}$$

■

C.1.6 Polynomial matrix fraction description (P-MFD)

Polynomial Matrix Fraction Description (P-MFD) is an alternative method to state space representation. It generalizes the notion of numerator and denominator for MIMO (Multiple-Input Multiple-Output) transfer matrices.

A *right* polynomial matrix fraction description of transfer matrix $\mathbf{F}(s)$ of dimension $p \times m$ reads as follows where $\mathbf{F}_{RN}(s)$ is a polynomial matrix of dimension $p \times m$ and $\mathbf{F}_{RD}(s)$ is a square polynomial matrix of dimension $m \times m$:

$$\mathbf{F}(s) = \mathbf{F}_{RN}(s)\mathbf{F}_{RD}^{-1}(s) \tag{C.81}$$

Similarly, a *left* polynomial matrix fraction description of transfer matrix $\mathbf{F}(s)$ of dimension $p \times m$ reads as follows where $\mathbf{F}_{LD}(s)$ is a square polynomial matrix of dimension $p \times p$ and $\mathbf{F}_{LN}(s)$ is a polynomial matrix of dimension $p \times m$:

$$\mathbf{F}(s) = \mathbf{F}_{LD}^{-1}(s)\mathbf{F}_{LN}(s) \tag{C.82}$$

We have seen that transfer matrix $\mathbf{F}(s)$ can be decomposed as follows:

$$\begin{cases} \mathbf{F}(s) = \frac{\mathbf{N}(s)}{\Psi(s)} \\ \mathbf{N}(s) = \mathbf{V}^{-1}(s)\mathbf{S}(s)\mathbf{W}^{-1}(s) \end{cases} \Rightarrow \mathbf{F}(s) = \mathbf{V}^{-1}(s)\frac{\mathbf{S}(s)}{\Psi(s)}\mathbf{W}^{-1}(s) \tag{C.83}$$

Matrix $\frac{\mathbf{S}(s)}{\Psi(s)}$ is the Smith-McMillan form of transfer matrix $\mathbf{F}(s)$: this is a quasi-diagonal matrix of rational functions which may possibly simplify. Furthermore $\mathbf{S}(s)$ is a polynomial matrix of dimension $p \times m$ and $\Psi(s)$ is a polynomial. Let $\frac{\varepsilon_i(s)}{\psi_i(s)}$ be the elements obtained after possible simplifications of this quasi-diagonal matrix. Then matrix $\frac{\mathbf{S}(s)}{\Psi(s)}$ has the following expression:

$$\frac{\mathbf{S}(s)}{\Psi(s)} = \begin{bmatrix} \frac{\varepsilon_1(s)}{\psi_1(s)} & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & \frac{\varepsilon_r(s)}{\psi_r(s)} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{C.84}$$

Let $\mathbf{D}_{\Psi_p}(s)$ and $\mathbf{D}_{\Psi_m}(s)$ be the following square polynomial matrices, build from polynomials $\psi_i(s)$ and of dimension $p \times p$ and $m \times m$ respectively (those 2 matrices just differ by the number of 1 on their diagonal), and $\mathbf{N}_{\varepsilon_{pm}}(s)$ the following polynomial matrix build from polynomials $\varepsilon_i(s)$ and of dimension $p \times m$:

$$\begin{cases} \mathbf{D}_{\Psi_p}(s) = diag\left(\psi_1(s), \cdots, \psi_r(s), 1, \cdots, 1\right) : \ p \times p \text{ square matrix} \\ \mathbf{D}_{\Psi_m}(s) = diag\left(\psi_1(s), \cdots, \psi_r(s), 1, \cdots, 1\right) : \ m \times m \text{ square matrix} \\ \mathbf{N}_{\varepsilon_{pm}}(s) = diag\left(\varepsilon_1(s), \cdots, \varepsilon_r(s), 0, \cdots, 0\right) : \ p \times m \text{ matrix} \end{cases} \tag{C.85}$$

Matrix $\frac{\mathbf{S}(s)}{\Psi(s)}$ of dimension $p \times m$ can then be expressed in 2 different ways:

$$\frac{\mathbf{S}(s)}{\Psi(s)} = \mathbf{N}_{\varepsilon_{pm}}(s)\mathbf{D}_{\Psi_m}^{-1}(s) = \mathbf{D}_{\Psi_p}^{-1}(s)\mathbf{N}_{\varepsilon_{pm}}(s) \tag{C.86}$$

By carrying these expressions in (C.83), transfer function $\mathbf{F}(s)$ is then written as follows:

$$\begin{aligned} \mathbf{F}(s) &= \mathbf{V}^{-1}(s)\tfrac{\mathbf{S}(s)}{\Psi(s)}\mathbf{W}^{-1}(s) \\ &= \mathbf{V}^{-1}(s)\mathbf{N}_{\varepsilon_{pm}}(s)\mathbf{D}_{\Psi_m}^{-1}(s)\mathbf{W}^{-1}(s) \\ &= \mathbf{V}^{-1}(s)\mathbf{D}_{\Psi_p}^{-1}(s)\mathbf{N}_{\varepsilon_{pm}}(s)\mathbf{W}^{-1}(s) \end{aligned} \tag{C.87}$$

By identifying (C.87) to (C.81) we get the following expression of the *right* polynomial matrix fraction description of transfer matrix $\mathbf{F}(s)$:

$$\begin{aligned} \mathbf{F}(s) &= \mathbf{F}_{RN}(s)\mathbf{F}_{RD}^{-1}(s) \\ &= \mathbf{V}^{-1}(s)\mathbf{N}_{\varepsilon_{pm}}(s)\mathbf{D}_{\Psi_m}^{-1}(s)\mathbf{W}^{-1}(s) \\ &= \mathbf{V}^{-1}(s)\mathbf{N}_{\varepsilon_{pm}}(s)\left(\mathbf{W}(s)\mathbf{D}_{\Psi_m}(s)\right)^{-1} \\ \Rightarrow & \begin{cases} \mathbf{F}_{RN}(s) = \mathbf{V}^{-1}(s)\mathbf{N}_{\varepsilon_{pm}}(s) \\ \mathbf{F}_{RD}(s) = \mathbf{W}(s)\mathbf{D}_{\Psi_m}(s) \end{cases} \end{aligned} \tag{C.88}$$

Similarly, by identifying (C.87) to (C.82) we get the following expression of the *left* polynomial matrix fraction description of transfer matrix $\mathbf{F}(s)$:

$$\begin{aligned} \mathbf{F}(s) &= \mathbf{F}_{LD}^{-1}(s)\mathbf{F}_{LN}(s) \\ &= \mathbf{V}^{-1}(s)\mathbf{D}_{\Psi_p}^{-1}(s)\mathbf{N}_{\varepsilon_{pm}}(s)\mathbf{W}^{-1}(s) \\ &= \left(\mathbf{D}_{\Psi_p}(s)\mathbf{V}(s)\right)^{-1}\mathbf{N}_{\varepsilon_{pm}}(s)\mathbf{W}^{-1}(s) \\ \Rightarrow & \begin{cases} \mathbf{F}_{LD}(s) = \mathbf{D}_{\Psi_p}(s)\mathbf{V}(s) \\ \mathbf{F}_{LN}(s) = \mathbf{N}_{\varepsilon_{pm}}(s)\mathbf{W}^{-1}(s) \end{cases} \end{aligned} \tag{C.89}$$

Note that $\mathbf{V}(s)$ and $\mathbf{W}(s)$ are unimodular polynomial matrices. Therefore $\mathbf{V}^{-1}(s)$ and $\mathbf{W}^{-1}(s)$ are also unimodular polynomial matrices and matrices $\mathbf{F}_{RN}(s)$, $\mathbf{F}_{RD}(s)$, $\mathbf{F}_{LN}(s)$ and $\mathbf{F}_{LD}(s)$ are polynomial matrices.

Consider now the following linear and time invariant system with m inputs and p outputs:

$$\begin{cases} \dot{\underline{x}} = \mathbf{A}\underline{x} + \mathbf{B}\underline{u} \\ \underline{y} = \mathbf{C}\underline{x} \end{cases} \tag{C.90}$$

Once the Smith-McMillan form of this system is obtained, the right polynomial matrix fraction description of $\mathbf{F}(s)$ leads to the following realization of the system in the controllable canonical form:

$$\begin{aligned} \mathbf{F}(s) &= \mathbf{F}_{RN}(s)\mathbf{F}_{RD}^{-1}(s) \\ &= \left(\mathbf{N}_{r-1}s^{r-1} + \cdots + \mathbf{N}_1 s + \mathbf{N}_0\right)\left(\mathbb{I}_m s^r + \mathbf{D}_{r-1}s^{r-1} + \cdots + \mathbf{D}_1 s + \mathbf{D}_0\right)^{-1} \\ &= \mathbf{C}_c\left(s\mathbb{I} - \mathbf{A}_c\right)^{-1}\mathbf{B}_c := \left(\begin{array}{c|c} \mathbf{A}_c & \mathbf{B}_c \\ \hline \mathbf{C}_c & \mathbf{0}_{p\times m} \end{array}\right) \end{aligned} \tag{C.91}$$

where:

$$\begin{cases} \mathbf{A}_c = \begin{bmatrix} \mathbf{0}_m & \mathbb{I}_m & \mathbf{0}_m & \cdots & \mathbf{0}_m \\ \mathbf{0}_m & \mathbf{0}_m & \mathbb{I}_m & \cdots & \vdots \\ & & & \ddots & \\ & & & & \mathbb{I}_m \\ -\mathbf{D}_0 & -\mathbf{D}_1 & \cdots & & -\mathbf{D}_{r-1} \end{bmatrix} \\ \mathbf{B}_c = \begin{bmatrix} \mathbf{0}_m \\ \vdots \\ \mathbf{0}_m \\ \mathbb{I}_m \end{bmatrix} \\ \mathbf{C}_c = \begin{bmatrix} \mathbf{N}_0 & \mathbf{N}_1 & \cdots & \mathbf{N}_{r-1} \end{bmatrix} \end{cases} \tag{C.92}$$

Similarly, the left polynomial matrix fraction description of $\mathbf{F}(s)$ leads to the following realization of the system in the observable canonical form:

$$\begin{aligned} \mathbf{F}(s) &= \mathbf{F}_{LD}^{-1}(s)\mathbf{F}_{LN}(s) \\ &= \left(\mathbb{I}_m s^r + \mathbf{D}_{r-1}s^{r-1} + \cdots + \mathbf{D}_1 s + \mathbf{D}_0\right)^{-1}\left(\mathbf{N}_{r-1}s^{r-1} + \cdots + \mathbf{N}_1 s + \mathbf{N}_0\right) \\ &= \mathbf{C}_o\left(s\mathbb{I} - \mathbf{A}_o\right)^{-1}\mathbf{B}_o := \left(\begin{array}{c|c} \mathbf{A}_o & \mathbf{B}_o \\ \hline \mathbf{C}_o & \mathbf{0}_{p\times m} \end{array}\right) \end{aligned} \tag{C.93}$$

where:

$$\begin{cases} \mathbf{A}_o = \begin{bmatrix} \mathbf{0}_m & \mathbf{0}_m & \cdots & \mathbf{0}_m & -\mathbf{D}_0 \\ \mathbb{I}_m & \mathbf{0}_m & \cdots & \mathbf{0}_m & -\mathbf{D}_1 \\ & \ddots & & & \\ \mathbf{0}_m & \cdots & & \mathbb{I}_m & -\mathbf{D}_{r-1} \end{bmatrix} \\ \mathbf{B}_o = \begin{bmatrix} \mathbf{N}_0 \\ \mathbf{N}_1 \\ \vdots \\ \mathbf{N}_{r-1} \end{bmatrix} \\ \mathbf{C}_o = \begin{bmatrix} \mathbf{0}_m & \cdots & \mathbf{0}_m & \mathbb{I}_m \end{bmatrix} \end{cases} \tag{C.94}$$

Example C.9. *We consider (again) the following transfer matrix* $\mathbf{F}(s)$*:*

$$\mathbf{F}(s) = \begin{bmatrix} \frac{1}{(s+1)(s+2)} & \frac{-1}{(s+1)(s+2)} \\ \frac{s^2+s-4}{(s+1)(s+2)} & \frac{2s^2-s-8}{(s+1)(s+2)} \\ \frac{s-2}{s+1} & \frac{2s-4}{s+1} \end{bmatrix} \tag{C.95}$$

Here the number of outputs p *of the system is equal to* 3 *whereas the number of inputs* m *of the system is equal to* 2*.*

We have seen in a previous example that the Smith-McMillan form of the transfer matrix $\mathbf{F}(s)$ *is the following:*

$$\begin{aligned} \mathbf{F}(s) &= \mathbf{V}^{-1}(s)\tfrac{\mathbf{S}(s)}{\Psi(s)}\mathbf{W}^{-1}(s) \\ &= \mathbf{V}^{-1}(s)\tfrac{1}{(s+1)(s+2)}\begin{bmatrix} 1 & 0 \\ 0 & s^2-4 \\ 0 & 0 \end{bmatrix}\mathbf{W}^{-1}(s) \\ &= \mathbf{V}^{-1}(s)\begin{bmatrix} \frac{1}{(s+1)(s+2)} & 0 \\ 0 & \frac{s-2}{s+1} \\ 0 & 0 \end{bmatrix}\mathbf{W}^{-1}(s) \\ &= \mathbf{V}^{-1}(s)\mathbf{N}_{\varepsilon_{pm}}(s)\mathbf{D}_{\Psi_m}^{-1}(s)\mathbf{W}^{-1}(s) \\ &= \mathbf{V}^{-1}(s)\mathbf{D}_{\Psi_p}^{-1}(s)\mathbf{N}_{\varepsilon_{pm}}(s)\mathbf{W}^{-1}(s) \end{aligned} \tag{C.96}$$

where:

$$\begin{cases} \mathbf{N}_{\varepsilon_{pm}}(s) = \begin{bmatrix} 1 & 0 \\ 0 & s-2 \\ 0 & 0 \end{bmatrix} \\ \mathbf{D}_{\Psi_m}(s) = \begin{bmatrix} (s+1)(s+2) & 0 \\ 0 & s+1 \end{bmatrix} \\ \mathbf{D}_{\Psi_p}(s) = \begin{bmatrix} (s+1)(s+2) & 0 & 0 \\ 0 & s+1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{cases} \tag{C.97}$$

and:

$$\begin{cases} \mathbf{V}(s) = \begin{bmatrix} 1/3 & 0 & 0 \\ -s/3 & 1/3 & 0 \\ s & -1 & 1 \end{bmatrix} \\ \mathbf{W}(s) = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \end{cases} \Leftrightarrow \begin{cases} \mathbf{V}^{-1}(s) = \begin{bmatrix} 3 & 0 & 0 \\ 3s & 3 & 0 \\ 0 & 3 & 1 \end{bmatrix} \\ \mathbf{W}^{-1}(s) = \begin{bmatrix} 1/3 & -1/3 \\ 1/3 & 2/3 \end{bmatrix} \end{cases} \tag{C.98}$$

Then the right polynomial matrix fraction description of $\mathbf{F}(s)$ *reads as follows:*

$$\mathbf{F}(s) = \mathbf{F}_{RN}(s)\mathbf{F}_{RD}^{-1}(s)$$
$$\textit{where} \begin{cases} \mathbf{F}_{RN}(s) = \mathbf{V}^{-1}(s)\mathbf{N}_{\varepsilon_{pm}}(s) = \begin{bmatrix} 3 & 0 \\ 3s & 3s-6 \\ 0 & 3s-6 \end{bmatrix} \\ \mathbf{F}_{RD}(s) = \mathbf{W}(s)\mathbf{D}_{\Psi_m}(s) = \begin{bmatrix} 2s^2+6s+4 & s+1 \\ -s^2-3s-2 & s+1 \end{bmatrix} \end{cases} \tag{C.99}$$

And the left polynomial matrix fraction description of $\mathbf{F}(s)$ *reads as follows:*

$$\mathbf{F}(s) = \mathbf{F}_{LD}^{-1}(s)\mathbf{F}_{LN}(s)$$
$$\textit{where} \begin{cases} \mathbf{F}_{LD}(s) = \mathbf{D}_{\Psi_p}(s)\mathbf{V}(s) = \begin{bmatrix} s^2/3+s+2/3 & 0 & 0 \\ -s^2/3-s/3 & s/3+1/3 & 0 \\ s & -1 & 1 \end{bmatrix} \\ \mathbf{F}_{LN}(s) = \mathbf{N}_{\varepsilon_{pm}}(s)\mathbf{W}^{-1}(s) = \begin{bmatrix} 1/3 & -1/3 \\ -2/3+s/3 & 2/3s-4/3 \\ 0 & 0 \end{bmatrix} \end{cases} \tag{C.100}$$

■

C.1.7 Right and left coprimeness

Consider two polynomial matrices $\mathbf{D}(s)$ and $\mathbf{N}(s)$ with the same number of columns. A square polynomial matrix $\mathbf{R}(s)$ is called a *common right divisor* of $\mathbf{D}(s)$ and $\mathbf{N}(s)$ if there exists polynomial matrices $\overline{\mathbf{D}}(s)$ and $\overline{\mathbf{N}}(s)$ such that:

$$\begin{bmatrix} \mathbf{D}(s) \\ \mathbf{N}(s) \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{D}}(s) \\ \overline{\mathbf{N}}(s) \end{bmatrix} \mathbf{R}(s) \tag{C.101}$$

Assume $\mathbf{D}(s)$ and $\mathbf{N}(s)$ have both m columns and such that the sum of their rows is greater of equal to m. Then the following statements are equivalent[3]:

- $\mathbf{D}(s)$ and $\mathbf{N}(s)$ are *right* coprime, meaning that the *greatest common right divisor* of $\mathbf{D}(s)$ and $\mathbf{N}(s)$ is an unimodular matrix. We will see in section C.1.8 how to compute the *greatest common right divisor* of two polynomials matrices.

- There exists polynomial matrices $\mathbf{X}(s)$ and $\mathbf{Y}(s)$ which solve the following Bezout identity:
$$\mathbf{X}(s)\,\mathbf{D}(s) + \mathbf{Y}(s)\,\mathbf{N}(s) = \mathbb{I}_m \tag{C.102}$$

- For any complex number s:
$$\operatorname{rank}\left(\begin{bmatrix} \mathbf{D}(s) \\ \mathbf{N}(s) \end{bmatrix}\right) = m \tag{C.103}$$

[3]Panos J. Antsaklis, Anthony N. Michel, A Linear Systems Primer, Birkhäuser Boston 2007, https://doi.org/10.1007/978-0-8176-4661-5

– The Smith form of $\begin{bmatrix} \mathbf{D}(s) \\ \mathbf{N}(s) \end{bmatrix}$ is $\begin{bmatrix} \mathbb{I}_m \\ \mathbf{0} \end{bmatrix}$.

Assume now $\mathbf{D}(s)$ and $\mathbf{N}(s)$ have both p rows and such that the sum of their columns is greater of equal to p. Then the following statements are equivalent:

– $\mathbf{D}(s)$ and $\mathbf{N}(s)$ are *left* coprime, meaning that the *greatest common left divisor* of $\mathbf{D}(s)$ and $\mathbf{N}(s)$ is an unimodular matrix.

– There exists polynomial matrices $\mathbf{X}(s)$ and $\mathbf{Y}(s)$ which solve the following Bezout identity:

$$\mathbf{D}(s)\,\mathbf{X}(s) + \mathbf{N}(s)\,\mathbf{Y}(s) = \mathbb{I}_p \tag{C.104}$$

– For any complex number s:

$$\text{rank}\left(\begin{bmatrix} \mathbf{D}(s) & \mathbf{N}(s) \end{bmatrix}\right) = p \tag{C.105}$$

– The Smith form of $\begin{bmatrix} \mathbf{D}(s) & \mathbf{N}(s) \end{bmatrix}$ is $\begin{bmatrix} \mathbb{I}_p & \mathbf{0} \end{bmatrix}$.

C.1.8 Hermite form and applications

Greatest Common Right Divisor (GCD)

The row Hermite form of a matrix, initially discovered by Charles Hermite (1822-1901) in 1851 for the domain of integers, is an upper triangular matrix that has the added constraints that the diagonals have the largest degrees in each column.

Let $\mathbf{D}(s)$ and $\mathbf{N}(s)$ be two polynomial matrices with the same number of columns and $\mathbf{P}(s) := \begin{bmatrix} \mathbf{D}(s) \\ \mathbf{N}(s) \end{bmatrix}$. The row Hermite form of $\mathbf{P}(s)$ can be computed by the following algorithm[4]:

1. In the first column of $\mathbf{P}(s)$ use row interchange to bring to the first row a lowest-degree entry among nonzero first-column entries;

2. Multiply the first row by a real number so that the first column entry $m_{1,1}(s)$ is monic, meaning that the coefficient corresponding to the highest degree monomial is 1;

3. For each entry $m_{i,1}(s)$ below the first row in the first column, use polynomial division to write the following relation where each remainder $r_{i,1}(s)$ is such that $deg(r_{i,1}(s)) < deg(m_{1,1}(s))$:

$$m_{i,1}(s) = q_i(s)m_{1,1}(s) + r_{i,1}(s), \quad i \geq 2 \tag{C.106}$$

If $m_{i,1}(s) = 0$ we set $q_i(s) = r_{i,1}(s) = 0$. If $deg(m_{i,1}(s)) = 0$ then $deg(m_{1,1}(s)) = 0$ and $r_{i,1}(s) = 0$.

[4] Rugh W. J., Linear System Theory, Prentice Hall, 1992

4. For $i \geq 2$, add to the ith-row the product between $-q_i(s)$ and the first row, that is the product $-q_i(s)\, m_{1,1}(s)$. The resulting entries in the first column, below the first row, are polynomials $r_{2,1}(s), \cdots, r_{p+m,1}(s)$, all of which having degrees less than $deg(m_{1,1}(s))$.

5. Repeat the preceding steps until all entries of the first column are zero except the first entry. Since the degrees of the entries below the first entry are lowered by at least one in each iteration, a finite number of operations is required.

Next, proceed with the second column of $\mathbf{P}(s)$ and repeat the above steps while ignoring the first row. This results in a monic, nonzero entry $m_{2,2}(s)$, with all entries below it are zero.

If $m_{1,2}(s)$ does not have lower degree than $m_{2,2}(s)$, then polynomial division of $m_{1,2}(s)$ by $m_{2,2}(s)$ as in step 4 replaces $m_{1,2}(s)$ by a polynomial of degree less than $deg(m_{2,2}(s))$.

Next repeat the process for the third column of $\mathbf{P}(s)$, while ignoring the first two rows. Continuing until the last column of $\mathbf{P}(s)$ yields to the so-called row Hermite form of polynomial matrix $\mathbf{P}(s) = \begin{bmatrix} \mathbf{D}(s) \\ \mathbf{N}(s) \end{bmatrix}$, where $\mathbf{U}(s)$ is an unimodular polynomial matrix (meaning that $\det(\mathbf{U}(s)) = \text{constant}$):

$$\boxed{\begin{cases} \mathbf{U}(s) \begin{bmatrix} \mathbf{D}(s) \\ \mathbf{N}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{R}(s) \\ \mathbf{0} \end{bmatrix} \\ \det(\mathbf{U}(s)) = \text{constant} \end{cases}} \tag{C.107}$$

Polynomial matrix $\mathbf{R}(s)$ is the *Greatest Common Right Divisor* (GCD) of polynomial matrices $\mathbf{N}(s)$ and $\mathbf{D}(s)$.

Polynomial matrix fraction description

The results presented in section 5.12.4 lead to a matrix *fraction* description of $\mathbf{F}(s)$, where $\mathbf{N}_{FR}(s)$, $\mathbf{D}_{FR}(s)$, $\mathbf{X}_R(s)$ and $\mathbf{Y}_R(s)$ are *rational* matrices.

To achieve a *polynomial* matrix description of $\mathbf{F}(s)$, first write $\mathbf{F}(s)$ as follows:

$$\mathbf{F}(s) = \mathbf{N}(s)\mathbf{D}^{-1}(s) \text{ where } \mathbf{D}(s) = diag\,(\Psi(s), \cdots, \Psi(s)) \tag{C.108}$$

where $\mathbf{N}(s)$ is a $p \times m$ polynomial matrix, $\mathbf{D}(s)$ a $m \times m$ diagonal polynomial matrix and $\Psi(s)$ a polynomial.

From the row Hermite form (see Section C.1.8) of polynomial matrix $\mathbf{P}(s) = \begin{bmatrix} \mathbf{D}(s) \\ \mathbf{N}(s) \end{bmatrix}$, we get relation (C.107) where $\mathbf{U}(s)$ is a $(p+m) \times (p+m)$ unimodular polynomial matrix (meaning that $\det(\mathbf{U}(s)) = \text{constant}$):

$$\begin{cases} \mathbf{U}(s) \begin{bmatrix} \mathbf{D}(s) \\ \mathbf{N}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{R}(s) \\ \mathbf{0} \end{bmatrix} \\ \det(\mathbf{U}(s)) = \text{constant} \end{cases} \tag{C.109}$$

Then $\mathbf{U}(s)$ is partitioned as follows, where $\mathbf{U}_{11}(s)$ is a $m \times m$ polynomial matrix and $\mathbf{U}_{22}(s)$ a $p \times p$ polynomial matrix:

$$\mathbf{U}(s) = \begin{bmatrix} \mathbf{U}_{11}(s) & \mathbf{U}_{12}(s) \\ \mathbf{U}_{21}(s) & \mathbf{U}_{22}(s) \end{bmatrix} \tag{C.110}$$

Matrix $\mathbf{U}_{22}(s)$ is square and invertible. Then the *left* coprime polynomial matrix description of $\mathbf{F}(s)$ reads:

$$\boxed{\mathbf{F}(s) = -\mathbf{U}_{22}^{-1}(s)\mathbf{U}_{21}(s)} \tag{C.111}$$

Indeed, from (C.107) we have:

$$\begin{aligned} &\begin{bmatrix} \mathbf{U}_{11}(s) & \mathbf{U}_{12}(s) \\ \mathbf{U}_{21}(s) & \mathbf{U}_{22}(s) \end{bmatrix} \begin{bmatrix} \mathbf{D}(s) \\ \mathbf{N}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{R}(s) \\ \mathbf{0} \end{bmatrix} \\ &\Rightarrow \mathbf{U}_{21}(s)\mathbf{D}(s) + \mathbf{U}_{22}(s)\mathbf{N}(s) = \mathbf{0} \\ &\Rightarrow \mathbf{F}(s) = \mathbf{N}(s)\mathbf{D}^{-1}(s) = -\mathbf{U}_{22}^{-1}(s)\mathbf{U}_{21}(s) \end{aligned} \tag{C.112}$$

With $\mathbf{V}(s) = \mathbf{U}^{-1}(s)$ we get the following relation where $\mathbf{V}_{11}(s)$ is a $m \times m$ polynomial matrix and $\mathbf{V}_{22}(s)$ a $p \times p$ polynomial matrix:

$$\mathbf{U}^{-1}(s) = \mathbf{V}(s) = \begin{bmatrix} \mathbf{V}_{11}(s) & \mathbf{V}_{12}(s) \\ \mathbf{V}_{21}(s) & \mathbf{V}_{22}(s) \end{bmatrix} \tag{C.113}$$

Matrix $\mathbf{V}_{11}(s)$ is square and invertible. Then the *right* coprime polynomial matrix description of $\mathbf{F}(s)$ reads:

$$\boxed{\mathbf{F}(s) = \mathbf{V}_{21}(s)\mathbf{V}_{11}^{-1}(s)} \tag{C.114}$$

Indeed we have:

$$\begin{aligned} &\begin{bmatrix} \mathbf{D}(s) \\ \mathbf{N}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{V}_{11}(s) & \mathbf{V}_{12}(s) \\ \mathbf{V}_{21}(s) & \mathbf{V}_{22}(s) \end{bmatrix} \begin{bmatrix} \mathbf{R}(s) \\ \mathbf{0} \end{bmatrix} \Rightarrow \begin{cases} \mathbf{D}(s) = \mathbf{V}_{11}(s)\mathbf{R}(s) \\ \mathbf{N}(s) = \mathbf{V}_{21}(s)\mathbf{R}(s) \end{cases} \\ &\Rightarrow \mathbf{F}(s) = \mathbf{N}(s)\mathbf{D}^{-1}(s) = \mathbf{V}_{21}(s)\mathbf{R}(s)\mathbf{R}^{-1}(s)\mathbf{V}_{11}^{-1}(s) = \mathbf{V}_{21}(s)\mathbf{V}_{11}^{-1}(s) \end{aligned} \tag{C.115}$$

Furthermore, as far as $\mathbf{V}(s) = \mathbf{U}^{-1}(s)$, the following relation, known as Bézout identity, holds:

$$\boxed{\begin{bmatrix} \mathbf{U}_{11}(s) & \mathbf{U}_{12}(s) \\ \mathbf{U}_{21}(s) & \mathbf{U}_{22}(s) \end{bmatrix} \begin{bmatrix} \mathbf{V}_{11}(s) & \mathbf{V}_{12}(s) \\ \mathbf{V}_{21}(s) & \mathbf{V}_{22}(s) \end{bmatrix} = \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbb{I} \end{bmatrix}} \tag{C.116}$$

With the same kind of work on the columns of the achieved row Hermite form, matrix $\mathbf{R}(s)$ can be diagonalized. Then we get the following relation where matrices $\mathbf{U}(s)$ (rows operations) and $\mathbf{W}(s)$ (columns operations) are unimodular and $diag(p_i(s))$ is a diagonal polynomial matrix:

$$\mathbf{U}(s) \begin{bmatrix} \mathbf{D}(s) \\ \mathbf{N}(s) \end{bmatrix} \mathbf{W}(s) = \begin{bmatrix} diag(p_i(s)) \\ \mathbf{0} \end{bmatrix} \tag{C.117}$$

Example C.10. *We consider the following transfer function:*

$$\mathbf{F}(s) = \begin{bmatrix} 0 & \frac{s}{s^2-s-1} \\ \frac{1}{s^2} & \frac{s-1}{s(s^2-s-1)} \\ \frac{1}{s} & \frac{s-1}{s^2-s-1} \\ 0 & \frac{1}{s^2-s-1} \end{bmatrix} \tag{C.118}$$

It is clear that $\mathbf{F}(s)$ *can be written as follows:*

$$\begin{aligned} \mathbf{F}(s) &= \tfrac{1}{s^2(s^2-s-1)} \begin{bmatrix} 0 & s^3 \\ (s^2-s-1) & s(s-1) \\ s(s^2-s-1) & s^2(s-1) \\ 0 & s^2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & s^3 \\ (s^2-s-1) & s(s-1) \\ s(s^2-s-1) & s^2(s-1) \\ 0 & s^2 \end{bmatrix} \begin{bmatrix} s^2(s^2-s-1) & 0 \\ 0 & s^2(s^2-s-1) \end{bmatrix}^{-1} \end{aligned} \tag{C.119}$$

Then we seek the row Hermite form of $\mathbf{P}(s) := \begin{bmatrix} \mathbf{D}(s) \\ \mathbf{N}(s) \end{bmatrix}$ *where:*

$$\begin{cases} \mathbf{D}(s) = \begin{bmatrix} s^2(s^2-s-1) & 0 \\ 0 & s^2(s^2-s-1) \end{bmatrix} \\ \mathbf{N}(s) = \begin{bmatrix} 0 & s^3 \\ (s^2-s-1) & s(s-1) \\ s(s^2-s-1) & s^2(s-1) \\ 0 & s^2 \end{bmatrix} \end{cases} \tag{C.120}$$

Following the algorithm proposed by Rugh[4]*, the row Hermite form of* $\mathbf{P}(s)$ *can be computed step by step as shown in Table C.4 for the first column.*

Then we proceed with the second column while ignoring the first row, as shown in Table C.5.

Finally matrix $\mathbf{U}(s)$ *is partitioned as follows, where* $\mathbf{U}_{11}(s)$ *is a* $m \times m$ *polynomial matrix and* $\mathbf{U}_{22}(s)$ *is a* $p \times p$ *polynomial matrix:*

$$\mathbf{U}(s) = \left[\begin{array}{cc|ccc} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 1 & 0 & 0 & -s \\ 1 & 0 & 0 & -s^2 & 0 & (s^2-s) \\ 0 & 0 & 0 & -s & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -(s^2-s-1) \end{array}\right] = \left[\begin{array}{c|c} \mathbf{U}_{11}(s) & \mathbf{U}_{12}(s) \\ \hline \mathbf{U}_{21}(s) & \mathbf{U}_{22}(s) \end{array}\right] \tag{C.121}$$

Matrix $\mathbf{U}_{22}(s)$ *is a square invertible polynomial matrix. The* left *coprime polynomial matrix description of* $\mathbf{F}(s)$ *reads:*

$$\mathbf{F}(s) = -\mathbf{U}_{22}^{-1}(s)\mathbf{U}_{21}(s) = -\begin{bmatrix} 1 & 0 & 0 & -s \\ 0 & -s^2 & 0 & (s^2-s) \\ 0 & -s & 1 & 0 \\ 0 & 0 & 0 & -(s^2-s-1) \end{bmatrix}^{-1} \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{C.122}$$

Row operation	$\mathbf{U}$(s)	$\mathbf{P}$(s)
	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} s^2(s^2-s-1) & 0 \\ 0 & s^2(s^2-s-1) \\ 0 & s^3 \\ (s^2-s-1) & s(s-1) \\ s(s^2-s-1) & s^2(s-1) \\ 0 & s^2 \end{bmatrix}$
$r_1 \leftrightarrow r_4$	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} (s^2-s-1) & s(s-1) \\ 0 & s^2(s^2-s-1) \\ 0 & s^3 \\ s^2(s^2-s-1) & 0 \\ s(s^2-s-1) & s^2(s-1) \\ 0 & s^2 \end{bmatrix}$
$r_4 \leftarrow r_4 - s^2\, r_1$	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -s^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} (s^2-s-1) & s(s-1) \\ 0 & s^2(s^2-s-1) \\ 0 & s^3 \\ 0 & -s^3(s-1) \\ s(s^2-s-1) & s^2(s-1) \\ 0 & s^2 \end{bmatrix}$
$r_5 \leftarrow r_5 - s\, r_1$	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -s^2 & 0 & 0 \\ 0 & 0 & 0 & -s & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} (s^2-s-1) & s(s-1) \\ 0 & s^2(s^2-s-1) \\ 0 & s^3 \\ 0 & -s^3(s-1) \\ 0 & 0 \\ 0 & s^2 \end{bmatrix}$

Table C.4: Row Hermite form of $\mathbf{P}(s)$: process with the first column

Row operation	$\mathbf{U}$(s)	$\mathbf{P}$(s)
$r_2 \leftrightarrow r_6$	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -s^2 & 0 & 0 \\ 0 & 0 & 0 & -s & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} (s^2-s-1) & s(s-1) \\ 0 & s^2 \\ 0 & s^3 \\ 0 & -s^3(s-1) \\ 0 & 0 \\ 0 & s^2(s^2-s-1) \end{bmatrix}$
$r_3 \leftarrow r_3 - s\, r_2$	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -s \\ 1 & 0 & 0 & -s^2 & 0 & 0 \\ 0 & 0 & 0 & -s & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} (s^2-s-1) & s(s-1) \\ 0 & s^2 \\ 0 & 0 \\ 0 & -s^3(s-1) \\ 0 & 0 \\ 0 & s^2(s^2-s-1) \end{bmatrix}$
$r_4 \leftarrow r_4 + (s^2 - s)\, r_2$	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -s \\ 1 & 0 & 0 & -s^2 & 0 & (s^2-s) \\ 0 & 0 & 0 & -s & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} (s^2-s-1) & s(s-1) \\ 0 & s^2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & s^2(s^2-s-1) \end{bmatrix}$
$r_6 \leftarrow r_6 - (s^2 - s - 1)\, r_2$	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -s \\ 1 & 0 & 0 & -s^2 & 0 & (s^2-s) \\ 0 & 0 & 0 & -s & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -(s^2-s-1) \end{bmatrix}$	$\begin{bmatrix} (s^2-s-1) & s(s-1) \\ 0 & s^2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

Table C.5: Row Hermite form of $\mathbf{P}(s)$: process with the second column

Opération (colonne ou ligne)	$\mathbf{U}(s)$	$\mathbf{P}(s)$	$\mathbf{W}(s)$
	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -s \\ 1 & 0 & 0 & -s^2 & 0 & (s^2-s) \\ 0 & 0 & 0 & -s & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -(s^2-s-1) \end{bmatrix}$	$\begin{bmatrix} (s^2-s-1) & s(s-1) \\ 0 & s^2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$c_2 \leftarrow c_2 - c_1$	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -s \\ 1 & 0 & 0 & -s^2 & 0 & (s^2-s) \\ 0 & 0 & 0 & -s & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -(s^2-s-1) \end{bmatrix}$	$\begin{bmatrix} (s^2-s-1) & 1 \\ 0 & s^2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$
$c_2 \leftrightarrow c_1$	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -s \\ 1 & 0 & 0 & -s^2 & 0 & (s^2-s) \\ 0 & 0 & 0 & -s & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -(s^2-s-1) \end{bmatrix}$	$\begin{bmatrix} 1 & (s^2-s-1) \\ s^2 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$
$r_2 \leftarrow r_2 - s^2\, r_1$	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -s^2 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -s \\ 1 & 0 & 0 & -s^2 & 0 & (s^2-s) \\ 0 & 0 & 0 & -s & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -(s^2-s-1) \end{bmatrix}$	$\begin{bmatrix} 1 & (s^2-s-1) \\ 0 & -s^2(s^2-s-1) \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$
$c_2 \leftarrow c_2 - (s^2-s-1)\, c_1$	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -s^2 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -s \\ 1 & 0 & 0 & -s^2 & 0 & (s^2-s) \\ 0 & 0 & 0 & -s & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -(s^2-s-1) \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & -s^2(s^2-s-1) \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} -1 & s^2-s \\ 1 & -s^2+s+1) \end{bmatrix}$

Table C.6: Getting the Smith-McMillan form from the row Hermite form

With $\mathbf{V}(s) = \mathbf{U}^{-1}(s)$ *we get:*

$$\begin{aligned} \mathbf{U}^{-1}(s) = \mathbf{V}(s) &= \left[\begin{array}{c|c} \mathbf{V}_{11}(s) & \mathbf{V}_{12}(s) \\ \hline \mathbf{V}_{21}(s) & \mathbf{V}_{22}(s) \end{array}\right] \\ &= \left[\begin{array}{cc|cccc} s^2 & s-s^2 & 0 & 1 & 0 & 0 \\ 0 & -1-s+s^2 & 0 & 0 & 0 & 1 \\ \hline 0 & s & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ s & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{array}\right] \end{aligned} \tag{C.123}$$

Matrix $\mathbf{V}_{11}(s)$ *is a square invertible polynomial matrix. The* right *coprime polynomial matrix description of* $\mathbf{F}(s)$ *reads as follows:*

$$\mathbf{F}(s) = \mathbf{V}_{21}(s)\mathbf{V}_{11}^{-1}(s) = \begin{bmatrix} 0 & s \\ 1 & 0 \\ s & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s^2 & s-s^2 \\ 0 & -1-s+s^2 \end{bmatrix}^{-1} \tag{C.124}$$

With the same kind of work on the columns of the achieved row Hermite form, matrix $\mathbf{R}(s)$ *can be diagonalized. Starting with Table C.5, rows operations (matrix* $\mathbf{U}(s)$*) and columns operations (matrix* $\mathbf{W}(s)$*) on* $\mathbf{P}(s)$ *are presented in Table C.6.*

Then we get the following relation where matrices $\mathbf{U}(s)$ *(rows operations)*

and $\mathbf{W}(s)$ *(columns operations) are unimodular:*

$$\mathbf{U}(s)\begin{bmatrix}\mathbf{D}(s)\\ \mathbf{N}(s)\end{bmatrix}\mathbf{W}(s) = \begin{bmatrix} diag(p_i(s)) \\ \mathbf{0}\end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -s^2 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -s \\ 1 & 0 & 0 & -s^2 & 0 & (s^2-s) \\ 0 & 0 & 0 & -s & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -(s^2-s-1) \end{bmatrix} \left[\begin{array}{cc} s^2(s^2-s-1) & 0 \\ 0 & s^2(s^2-s-1) \\ \hdashline 0 & s^3 \\ (s^2-s-1) & s(s-1) \\ s(s^2-s-1) & s^2(s-1) \\ 0 & s^2 \end{array}\right] \begin{bmatrix} -1 & s^2-s \\ 1 & -s^2+s+1) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & -s^2(s^2-s-1) \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{C.125}$$

■

C.1.9 McMillan degree, poles and zeros of a transfer function

From the *left* and *right* coprime polynomial matrix descriptions of $\mathbf{F}(s)$ reads, the following properties hold[4]:

- The determinant of $\mathbf{V}_{11}(s)$ is equal to the determinant of $\mathbf{U}_{22}(s)$ times a constant c:

$$det\left(\mathbf{V}_{11}(s)\right) = c\, det\left(\mathbf{U}_{22}(s)\right), \quad c = \text{constant} \tag{C.126}$$

- The degree of $det\left(\mathbf{V}_{11}(s)\right)$ (or $det\left(\mathbf{U}_{22}(s)\right)$) is called the McMillan degree of $\mathbf{F}(s)$. This degree equals the dimension of any minimal realization of $\mathbf{F}(s)$.

- The poles of $\mathbf{F}(s)$ are either the roots of $det\left(\mathbf{V}_{11}(s)\right) = 0$ or the roots of $det\left(\mathbf{U}_{22}(s)\right) = 0$.

- The zeros of $\mathbf{F}(s)$ are the values of s for which the rank of either $\mathbf{V}_{21}(s)$ or $\mathbf{U}_{21}(s)$ drops below its normal rank.

C.1.10 Column reduced and row reduced polynomial matrix

Let $\mathbf{D}(s)$ a *square* $p \times p$ polynomial matrix and denote its column degree by $\{dc_i\}$. Then we can write the following relation, where $\mathbf{\Gamma}_c$ is a constant matrix and $\mathbf{D}_0(s)$ a polynomial matrix with column degrees strictly lower than those of $\mathbf{D}(s)$:

$$\mathbf{D}(s) = \mathbf{\Gamma}_c \operatorname{diag}\left(s^{dc_1}, \cdots, s^{dc_p}\right) + \mathbf{D}_0(s) \tag{C.127}$$

Matrix $\mathbf{\Gamma}_c$ is called the highest column degree coefficient matrix, or the leading column coefficient matrix, of $\mathbf{D}(s)$.

By definition[5], a square nonsingular polynomial matrix is *column reduced* if and only if its leading column coefficient matrix $\mathbf{\Gamma}_c$ is nonsingular, or equivalently if:

$$det\,(\mathbf{\Gamma}_c) \neq 0 \Leftrightarrow deg\ det\,(\mathbf{D}(s)) = \sum_{i=1}^{p} dc_i \tag{C.128}$$

The example provided by Kailath[6] will give an insight. Let $\mathbf{D}(s)$ be the following polynomial matrix:

$$\begin{aligned}\mathbf{D}(s) &= \begin{bmatrix} (s+1)^2(s+2)^2 & -(s+1)^2(s+2) \\ 0 & s+2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} s^4 & 0 \\ 0 & s^3 \end{bmatrix} \\ &\quad + \begin{bmatrix} 6s^3+13s^2+12s+4 & -4s^2-5s-2 \\ 0 & s+2 \end{bmatrix} \\ &:= \mathbf{\Gamma}_c\ \mathrm{diag}\left(s^4, s^3\right) + \mathbf{D}_0(s)\end{aligned} \tag{C.129}$$

Then the leading column coefficient matrix $\mathbf{\Gamma}_c$ of $\mathbf{D}(s)$ reads:

$$\mathbf{\Gamma}_c = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \tag{C.130}$$

The column coefficient matrix $\mathbf{\Gamma}_c$ is singular, thus $\mathbf{D}(s)$ is *not column reduced* and $deg\ det\,(\mathbf{D}(s)) = 5 \neq \sum_{i=1}^{p} dc_i = 4+3$.

Similarly, a square nonsingular polynomial matrix is *row reduced* if and only if the matrix formed by its *row* leading coefficients is nonsingular. Coming back to polynomial matrix $\mathbf{D}(s)$ in (C.129), its leading row coefficient matrix $\mathbf{\Gamma}_r$ reads:

$$\mathbf{\Gamma}_r = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{C.131}$$

We can see that $\mathbf{D}(s)$ is *row reduced*. Indeed denoting by dr_i the row degree of $\mathbf{D}(s)$ we have:

$$det\,(\mathbf{\Gamma}_r) \neq 0 \Leftrightarrow deg\ det\,(\mathbf{D}(s)) = 5 = \sum_{i=1}^{2} dr_i = 4+1 \tag{C.132}$$

Elementary row / column operations can be used to make a polynomial matrix *column* reduced or *row* reduced. For example, putting in the first column of $\mathbf{D}(s)$ in (C.129) the sum between its first column and polynomial $(s+2)$ times its second column leads to the following matrix:

$$\underline{c}_1 \leftarrow \underline{c}_1 + (s+2)\underline{c}_2 \Rightarrow \mathbf{D}_{cr}(s) = \begin{bmatrix} 0 & -(s+1)^2(s+2) \\ (s+2)^2 & s+2 \end{bmatrix} \tag{C.133}$$

[5] J. Ruiz-León, F. Kraffer, A. Castellanos and V. L. E. Ramos, Column reduced proper rational matrices, 1999 European Control Conference (ECC), 1999, pp. 936-941, doi: 10.23919/ECC.1999.7099427.

[6] Kailath T., Linear Systems, Prentice Hall, Englewood Cliffs, N.J., 1980

Matrix $\mathbf{D}_{cr}(s)$. Indeed the highest column degree coefficient matrix of $\mathbf{D}_{cr}(s)$ reads:

$$\mathbf{\Gamma}_c = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{C.134}$$

Because $\mathbf{\Gamma}_c$ is nonsingular, $\mathbf{D}_{cr}(s)$ is *column reduced* and the following relation holds:

$$det\,(\mathbf{\Gamma}_c) \neq 0 \Leftrightarrow deg\ det\,(\mathbf{D}_{cr}(s)) = 5 = \sum_{i=1}^{p} dc_i = 2+3 \tag{C.135}$$

Furthermore, it can be shown[6] that transfer function $\mathbf{N}(s)\mathbf{D}^{-1}(s)$ where $\mathbf{D}(s)$ is *column reduced* is strictly proper (proper) if and only if each column of $\mathbf{N}(s)$ has a degree less than (less than or equal to) the degree of the corresponding column of $\mathbf{D}(s)$.

C.1.11 Row index of transfer matrix

The *row index* ν of a transfer matrix $\mathbf{F}(s)$ is the largest row degree of $\mathbf{D}_L(s)$ in any left coprime factorization of $\mathbf{F}(s) = \mathbf{D}_L(s)^{-1}\mathbf{N}_L(s)$ with $\mathbf{D}_L(s)$ row reduced.

Alternatively *row index* ν is also equal to the *observability index* of any irreducible realization of $\mathbf{F}(s) = \mathbf{C}\,(s\mathbb{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} := \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right)$. To be more specific, let $\underline{c}_i^T$ be the ith row of $\mathbf{C}$. Then the observability matrix $\mathbf{Q}_o$ can be written explicitly as follows:

$$\mathbf{Q}_o = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} = \begin{bmatrix} \underline{c}_1^T \\ \vdots \\ \underline{c}_p^T \\ \hdashline \underline{c}_1^T\mathbf{A} \\ \vdots \\ \underline{c}_p^T\mathbf{A} \\ \hdashline \underline{c}_1^T\mathbf{A}^{n-1} \\ \vdots \\ \underline{c}_p^T\mathbf{A}^{n-1} \end{bmatrix} \tag{C.136}$$

Let ν_j be the number of the linearly independent rows associated with $\underline{c}_j^T$ in $\mathbf{Q}_o$. In other words rows $\left(\ \underline{c}_j^T,\ \underline{c}_j^T\mathbf{A},\ \cdots,\ \underline{c}_j^T\mathbf{A}^{\nu_j-1}\ \right)$ are linearly independent and $\underline{c}_j^T\mathbf{A}^{\nu_j+i}\,\forall\, i \geq 0$ are linearly dependent. Then if $\mathbf{Q}_o$ has rank n we have $\sum_{j=1}^{p}\nu_j = n$ and $\nu = max\,(\nu_1, \cdots, \nu_p)$ is called the *observability index.*

A similar definition exists for *controllability index* based on the columns of the controllability matrix $\mathbf{Q}_c$.

C.1.12 Flat outputs

Let $\mathbf{F}(s) = (s\mathbb{I} - \mathbf{A})^{-1}\mathbf{B}$ (here the output matrix $\mathbf{C}$ has been set to the identity matrix: $\mathbf{C} = \mathbb{I}$) and consider the *right* polynomial matrix fraction description (C.81) of transfer matrix $\mathbf{F}(s)$:

$$\underline{X}(s) = \mathbf{F}(s)\underline{U}(s) = \mathbf{F}_{RN}(s)\mathbf{F}_{RD}^{-1}(s)\underline{U}(s) \tag{C.137}$$

By introducing vector $\underline{\xi}(s)$, we can write:

$$\begin{cases} \underline{X}(s) = \mathbf{F}_{RN}(s)\underline{\xi}(s) \\ \underline{U}(s) = \mathbf{F}_{RD}(s)\underline{\xi}(s) \end{cases} \tag{C.138}$$

Matrices $\mathbf{F}_{RN}(s)$ and $\mathbf{F}_{RD}(s)$ are polynomial maytrices. Thus the preceding relation states that state vector $\underline{x}(t)$ and control vector $\underline{u}(t)$ can be expressed as a function of vector $\underline{\xi}(t)$ and its derivatives: vector $\underline{\xi}(t)$ is called vector of flat outputs[7].

Interestingly, when flat vector is used, the system dynamics can be obtained without integrating differential equations, just differentiations.

C.2 ICD controllers

C.2.1 Controller structure

ICD stands for *Individual Control Channel.* ICD controllers are also named *reversed frame normalizing controllers* after Hung & MacFarlane in 1982[8]. In the ICD framework, which is basically the Smith-McMillan form of the transfer matrix, SISO techniques can be used to control MIMO systems.

We consider the following transfer matrix $\mathbf{F}(s)$ where $\mathbf{N}(s)$ is a $p \times m$ polynomial matrix and $\Psi(s)$ is a polynomial which is actually the least common multiple of the denominators of all the elements in $\mathbf{F}(s)$):

$$\mathbf{F}(s) = \frac{\mathbf{N}(s)}{\Psi(s)} \tag{C.139}$$

We have seen in Section C.1.2 that matrix $\mathbf{N}(s)$ is equivalent to matrix $\mathbf{S}(s)$, where polynomial $\mathbf{S}(s)$ is called the Smith form $\mathbf{N}(s)$:

$$\mathbf{S}(s) = \mathbf{V}(s)\mathbf{N}(s)\mathbf{W}(s) \Leftrightarrow \mathbf{N}(s) = \mathbf{V}^{-1}(s)\mathbf{S}(s)\mathbf{W}^{-1}(s) \tag{C.140}$$

Moreover:

– $\mathbf{S}(s)$ is a $p \times m$ quasi-diagonal polynomial matrix:

$$\mathbf{S}(s) = \begin{bmatrix} s_1(s) & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & s_r(s) & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{C.141}$$

[7] J. Lévine, D.V. Nguyen, Flat output characterization for linear systems using polynomial matrices, Systems & Control Letters 48 (2003) 69 - 75

[8] Y.S. Hung & A.G.J. MacFarlane, Multivariable Feedback: A Quasi-Classical Approach, Lecture Notes in Control and Information Sciences, vol 40. Springer, Berlin, 1982

- Square matrices $\mathbf{V}(s)$ and $\mathbf{W}(s)$ are unimodular polynomial matrices;
- The dimension of matrix $\mathbf{V}(s)$ is $p \times p$ whereas the dimension of matrix $\mathbf{W}(s)$ is $m \times m$.

Dividing polynomial matrix $\mathbf{N}(s)$ by polynomial $\Psi(s)$ we get:

$$\mathbf{F}(s) = \frac{\mathbf{N}(s)}{\Psi(s)} = \mathbf{V}^{-1}(s)\frac{\mathbf{S}(s)}{\Psi(s)}\mathbf{W}^{-1}(s) := \mathbf{V}^{-1}(s)\mathbf{F}_S(s)\mathbf{W}^{-1}(s) \tag{C.142}$$

Matrix $\mathbf{F}_S(s) := \frac{\mathbf{S}(s)}{\Psi(s)}$ is called the Smith-McMillan form of transfer matrix $\mathbf{F}(s)$: this is a quasi-diagonal matrix of rational functions which may possibly simplify. Let $\frac{\varepsilon_i(s)}{\psi_i(s)}$ be the rational functions obtained after simplification of each element of quasi-diagonal matrix $\frac{\mathbf{S}(s)}{\Psi(s)}$. Then the Smith-McMillan form $\mathbf{F}_S(s)$ of transfer matrix $\mathbf{F}(s)$ reads:

$$\mathbf{F}_S(s) := \frac{\mathbf{S}(s)}{\Psi(s)} = \begin{bmatrix} \frac{s_1(s)}{\Psi(s)} & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & \frac{s_r(s)}{\Psi(s)} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \frac{\varepsilon_1(s)}{\psi_1(s)} & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & \frac{\varepsilon_r(s)}{\psi_r(s)} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{C.143}$$

The polynomials $\varepsilon_k(s)$ and $\psi_k(s)$ are coprime. Furthermore $\varepsilon_k(s)$ divides $\varepsilon_k(s)(s)$ and $\psi_{k+1}(s)$ divides $\psi_k(s) \forall k = 1, \ldots, r-1$:

$$\left.\begin{array}{l} s_k(s) \,|\, s_{k+1}(s) \\ \psi_{k+1}(s) \,|\, \psi_k(s) \end{array}\right\} \forall k = 1, \ldots, r-1 \tag{C.144}$$

We present hereafter some results presented by Mohsenizadeh & al.[9].

We consider Figure C.1 where $\mathbf{F}(s)$ is a MIMO plant with m inputs and p outputs, and $\mathbf{C}(s)$ a MIMO controller with p inputs and m outputs:

The main idea is to change the actual control loop in the top of Figure C.1 by an equivalent control loop using the Smith-McMillan form $\mathbf{F}_S(s)$ of plant $\mathbf{F}(s)$, as shown in the bottom of Figure C.1. The connection between $\mathbf{C}(s)$ and $\mathbf{C}_S(s)$ is obtained through the square unimodular polynomial matrices $\mathbf{V}(s)$ and $\mathbf{W}(s)$:

$$\mathbf{F}(s) = \mathbf{V}^{-1}(s)\mathbf{F}_S(s)\mathbf{W}^{-1}(s) \Rightarrow \mathbf{C}(s) = \mathbf{W}(s)\mathbf{C}_S(s)\mathbf{V}(s) \tag{C.145}$$

A controller defined by (C.145) is called SVD controller or reversed-frame normalizing controller.

[9]Daniel N. Mohsenizadeh, Lee H. Keel and Shankar P. Bhattacharyya, Multivariable Controller Synthesis using SISO Design Methods, 2015 IEEE 54th Annual Conference on Decision and Control (CDC), December 15-18, 2015. Osaka, Japan

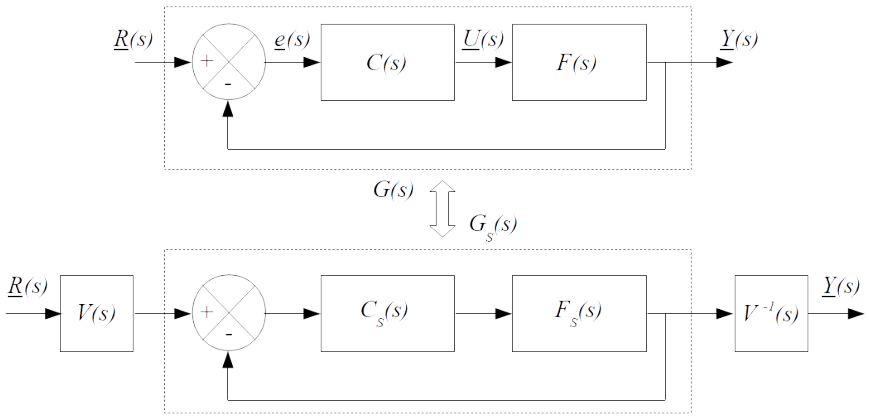

Figure C.1: Equivalent control loops with MIMO controller $C(s)$

C.2.2 Closed loop transfer function

We will assume that $\mathbf{C}_S(s)$ is a quasi-diagonal $m \times p$ controller with transfer functions $C_1(s), \cdots, C_r(s)$, where r is the number of non-zero terms in the Smith form $\mathbf{S}(s)$ of $\mathbf{N}(s)$. Then, denoting by $n_{C_k}(s)$ the numerator of transfer function $C_{S_k}(s)$ and by $d_{C_k}(s)$ its denominator, we get:

$$C_{S_k}(s) = \frac{n_{C_k}(s)}{d_{C_k}(s)} \Rightarrow \mathbf{C}_S(s) = \begin{bmatrix} \frac{n_{C_1}(s)}{d_{C_1}(s)} & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & \frac{n_{C_r}(s)}{d_{C_r}(s)} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{C.146}$$

Thanks to the definition (C.145) of controller $\mathbf{C}(s)$, the product $\mathbf{F}(s)\mathbf{C}(s)$ reads as follows:

$$\begin{cases} \mathbf{F}(s) = \mathbf{V}^{-1}(s)\mathbf{F}_S(s)\mathbf{W}^{-1}(s) \\ \mathbf{C}(s) = \mathbf{W}(s)\mathbf{C}_S(s)\mathbf{V}(s) \end{cases} \Rightarrow \mathbf{F}(s)\mathbf{C}(s) = \mathbf{V}^{-1}(s)\mathbf{F}_S(s)\mathbf{C}_S(s)\mathbf{V}(s) \tag{C.147}$$

Then we can use this result to compute the closed-loop transfer function:

$$\begin{aligned} &\underline{Y}(s) = \mathbf{F}(s)\mathbf{C}(s)\left(\underline{R}(s) - \underline{Y}(s)\right) \\ &\Rightarrow \left(\mathbb{I} + \mathbf{F}(s)\mathbf{C}(s)\right)\underline{Y}(s) = \mathbf{F}(s)\mathbf{C}(s)\underline{R}(s) \end{aligned} \tag{C.148}$$

We finally get:

$$\begin{aligned} \underline{Y}(s) &= \left(\mathbb{I} + \mathbf{F}(s)\mathbf{C}(s)\right)^{-1}\mathbf{F}(s)\mathbf{C}(s)\underline{R}(s) \\ &= \left(\mathbb{I} + \mathbf{V}^{-1}(s)\mathbf{F}_S(s)\mathbf{C}_S(s)\mathbf{V}(s)\right)^{-1}\mathbf{F}(s)\mathbf{C}(s)\underline{R}(s) \\ &= \left(\mathbf{V}^{-1}(s)\left(\mathbb{I} + \mathbf{F}_S(s)\mathbf{C}_S(s)\right)\mathbf{V}(s)\right)^{-1}\mathbf{F}(s)\mathbf{C}(s)\underline{R}(s) \\ &= \mathbf{V}^{-1}(s)\left(\mathbb{I} + \mathbf{F}_S(s)\mathbf{C}_S(s)\right)^{-1}\mathbf{F}_S(s)\mathbf{C}_S(s)\mathbf{V}(s)\underline{R}(s) \\ &:= \mathbf{G}(s)\underline{R}(s) \end{aligned} \tag{C.149}$$

Thus the closed-loop transfer matrix $\mathbf{G}(s)$ reads:

$$\begin{aligned} \mathbf{G}(s) &= \mathbf{V}^{-1}(s)\left(\mathbb{I} + \mathbf{F}_S(s)\mathbf{C}_S(s)\right)^{-1}\mathbf{F}_S(s)\mathbf{C}_S(s)\mathbf{V}(s) \\ &:= \mathbf{V}^{-1}(s)\mathbf{G}_S(s)\mathbf{V}(s) \\ \text{where } \mathbf{G}_S(s) &= \left(\mathbb{I} + \mathbf{F}_S(s)\mathbf{C}_S(s)\right)^{-1}\mathbf{F}_S(s)\mathbf{C}_S(s) \end{aligned} \tag{C.150}$$

This equivalence between closed-loop transfer matrix $\mathbf{G}(s)$ and $\mathbf{G}_S(s)$ is represented in Figure C.1.

C.2.3 Proper controller

Let $\underline{\widetilde{w}}_k(s)$ be the column polynomial vectors of unimodular of matrix $\mathbf{W}(s)$ and $\underline{\widetilde{v}}_k^T(s)$ be the row polynomial vectors of unimodular of matrices $\mathbf{V}(s)$. Then we get:

$$\begin{cases} \mathbf{W}(s) := \begin{bmatrix} \underline{\widetilde{w}}_1(s) & \cdots & \underline{\widetilde{w}}_r(s) \end{bmatrix} \\ \mathbf{V}(s) := \begin{bmatrix} \underline{\widetilde{v}}_1^T(s) \\ \vdots \\ \underline{\widetilde{v}}_r^T(s) \end{bmatrix} \end{cases} \tag{C.151}$$
$$\Rightarrow \mathbf{C}(s) = \mathbf{W}(s)\mathbf{C}_S(s)\mathbf{V}(s) = \textstyle\sum_{k=1}^{r} \underline{\widetilde{w}}_k(s)\underline{\widetilde{v}}_k^T(s)C_{S_k}(s)$$

In order to achieve a proper transfer matrix $\mathbf{C}(s)$ for the actual MIMO controller, the relative degree of each transfer function $C_{S_k}(s)$, $rdeg\,(C_{S_k}(s))$, shall be chosen such that $\underline{\widetilde{w}}_k(s)\underline{\widetilde{v}}_k^T(s)C_{S_k}(s)$ is proper:

$$rdeg\,(C_{S_k}(s)) \text{ such that } lim_{s\to\infty}\underline{\widetilde{w}}_k(s)\underline{\widetilde{v}}_k^T(s)C_{S_k}(s) < \infty \;\forall k = 1, \cdots, r \tag{C.152}$$

From relation (C.150), it is clear that the characteristic polynomial of $\mathbb{I} + \mathbf{F}(s)\mathbf{C}(s)$ is equal to the characteristic polynomial of $\mathbb{I} + \mathbf{F}_S(s)\mathbf{C}_S(s)$. Then we use the assumption (C.146) that $\mathbf{C}_S(s)$ is a quasi-diagonal controller with transfer functions $C_{S_k}(s) = \frac{n_{C_k}(s)}{d_{C_k}(s)}$:

$$\mathbf{F}_S(s)\mathbf{C}_S(s) = \begin{bmatrix} \frac{\varepsilon_1(s)}{\psi_1(s)}\frac{n_{C_1}(s)}{d_{C_1}(s)} & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & \frac{\varepsilon_r(s)}{\psi_r(s)}\frac{n_{C_r}(s)}{d_{C_r}(s)} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix}$$
$$\Rightarrow \mathbb{I} + \mathbf{F}_S(s)\mathbf{C}_S(s) = \begin{bmatrix} 1 + \frac{\varepsilon_1(s)\,n_{C_1}(s)}{\psi_1(s)\,d_{C_1}(s)} & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & 1 + \frac{\varepsilon_r(s)\,n_{C_r}(s)}{\psi_r(s)\,d_{C_r}(s)} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbb{I} \end{bmatrix}$$
$$\Rightarrow (\mathbb{I} + \mathbf{F}_S(s)\mathbf{C}_S(s))^{-1} = \begin{bmatrix} \frac{\psi_1(s)\,d_{C_1}(s)}{\psi_1(s)\,d_{C_1}(s)+\varepsilon_1(s)\,n_{C_1}(s)} & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & \frac{\psi_r(s)\,d_{C_r}(s)}{\psi_r(s)\,d_{C_r}(s)+\varepsilon_r(s)\,n_{C_r}(s)} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbb{I} \end{bmatrix} \tag{C.153}$$

We finally get:

$$\begin{aligned}\mathbf{G}_S(s) &= (\mathbb{I}+\mathbf{F}_S(s)\mathbf{C}_S(s))^{-1}\,\mathbf{F}_S(s)\mathbf{C}_S(s)\\ &= \begin{bmatrix} \frac{\varepsilon_1(s)\,n_{C_1}(s)}{\psi_1(s)\,d_{C_1}(s)+\varepsilon_1(s)\,n_{C_1}(s)} & & & \mathbf{0}\\ & \ddots & & \vdots\\ & & \frac{\varepsilon_r(s)\,n_{C_r}(s)}{\psi_r(s)\,d_{C_r}(s)+\varepsilon_r(s)\,n_{C_r}(s)} & \mathbf{0}\\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\end{bmatrix}\end{aligned} \tag{C.154}$$

Then the poles of $\mathbf{G}_S(s)$, which are also the poles of $\mathbf{G}(s)$ because matrix $\mathbf{V}(s)$ is unimodular, are the roots of the following polynomial $\chi_G(s)$:

$$\chi_G(s) = \prod_{k=1}^{r}\left(\psi_k(s)\,d_{C_k}(s)+\varepsilon_k(s)\,n_{C_k}(s)\right) \tag{C.155}$$

The interest in this expression is that each elementary transfer function $C_{S_k}(s) = \frac{n_{C_k}(s)}{d_{C_k}(s)}$ is used to set the roots of each polynomial $\psi_k(s)\,d_{C_k}(s) + \varepsilon_k(s)\,n_{C_k}(s)$, $k = 1,\cdots,r$ under the constraint (C.152). Assume that $\psi_k(s)$ is a polynomial of degree n_k and that polynomial $\varepsilon_k(s)$ has a degree lower or equal to n_k. Then denoting by m_k the degree of polynomial $d_{C_k}(s)$ and by $m_k - rdeg\,(C_{S_k}(s))$ the degree of polynomial $n_{C_k}(s)$:

$$\begin{cases} deg\,(\varepsilon_k(s)) \le deg\,(\psi_k(s)) := n_k\\ deg\,(d_{C_k}(s)) := m_k\\ deg\,(n_{C_k}(s)) = m_k - rdeg\,(C_{S_k}(s))\end{cases} \tag{C.156}$$

Polynomials $d_{C_k}(s)$ and $n_{C_k}(s)$ have a total of $m_k + 1 + (m_k - rdeg\,(C_{S_k}(s)) + 1) = 2m_k - rdeg\,(C_{S_k}(s)) + 2$ coefficients to be determined. On the other hand, and with the assumptions that we made, polynomial $\psi_k(s)\,d_{C_k}(s) + \varepsilon_k(s)\,n_{C_k}(s)$ is a polynomial of degree $n_k + m_k$ which has $n_k + m_k + 1$ coefficients. Thus the number of known coefficients, $n_k + m_k + 1$, will be equal to the number of coefficients to be determined, $2m_k - rdeg\,(C_{S_k}(s)) + 2$, as soon as the following equality holds:

$$\begin{aligned} n_k + m_k + 1 &= 2m_k - rdeg\,(C_{S_k}(s)) + 2\\ \Rightarrow m_k &= n_k + rdeg\,(C_{S_k}(s)) - 1\end{aligned} \tag{C.157}$$

Furthermore we have:

$$deg\,(\psi_k(s)\,d_{C_k}(s)+\varepsilon_k(s)\,n_{C_k}(s)) = n_k+m_k = 2n_k+rdeg\,(C_{S_k}(s))-1 \tag{C.158}$$

In other words, as soon as a polynomial of degree $2n_k + rdeg\,(C_{S_k}(s)) - 1$ is provided, the coefficients of this polynomial can be identified with the coefficients of polynomial $\psi_k(s)\,d_{C_k}(s) + \varepsilon_k(s)\,n_{C_k}(s)$, where $d_{C_k}(s)$ is a polynomial of degree m_k and $n_{C_k}(s)$ a polynomial of degree $m_k - rdeg\,(C_{S_k}(s))$.

Once each elementary transfer function $C_{S_k}(s) = \frac{n_{C_k}(s)}{d_{C_k}(s)}$, $k = 1,\cdots,r$, of $\mathbf{C}_S(s)$ have been set, the actual controller $\mathbf{C}(s)$ is obtained thanks to (C.145).

C.2.4 Decoupling

It is worth noticing that for square systems with p inputs and p outputs, input-output decoupling can be achieved through the following setting, provided that the relative degree condition (C.152) is satisfied and that each open loop transfer function $F_{S_i}(S)$ is stable (in order to ensure internal stability):

$$F_{S_i}(S)\, C_{S_i}(s) = F_{S_j}(S)\, C_{S_j}(s)\ \forall\, i, j \tag{C.159}$$

Indeed, in this situation $\mathbf{F}_S(s)\mathbf{C}_S(s)$ is a $p \times p$ square matrix:

$$\begin{cases} \mathbf{F}_S(s)\mathbf{C}_S(s) = \begin{bmatrix} G_{S_1}(s) & & & \\ & G_{S_2}(s) & & \\ & & \ddots & \\ & & & G_{S_p}(s) \end{bmatrix} \\ G_{S_i}(s) = \frac{F_{S_i}(S)\, C_{S_i}(s)}{1+F_{S_i}(S)\, C_{S_i}(s)}\ \forall\, i \end{cases} \tag{C.160}$$

Furthermore the closed-loop transfer matrix $\mathbf{G}(s)$ is also a $p \times p$ square matrix. From (C.150) we get:

$$\begin{aligned} \mathbf{G}(s) &= \mathbf{V}^{-1}(s)\mathbf{G}_S(s)\mathbf{V}(s) \\ &= \mathbf{V}^{-1}(s) \begin{bmatrix} G_{S_1}(s) & & & \\ & G_{S_2}(s) & & \\ & & \ddots & \\ & & & G_{S_p}(s) \end{bmatrix} \mathbf{V}(s) \end{aligned} \tag{C.161}$$

Thus it is clear that if $F_{S_i}(S)\, C_{S_i}(s) = F_{S_j}(S)\, C_{S_j}(s)\ \forall\, i, j$ then $G_{S_i}(s) = G_{S_j}(s)\ \forall\, i, j$. Then provided than the relative degree condition (C.152) is satisfied, we achieve decoupling control:

$$\begin{aligned} G_{S_i}(s) = G_{S_j}(s)\ \forall\, i, j \Rightarrow \mathbf{G}(s) &= \mathbf{V}^{-1}(s) \begin{bmatrix} G_{S_1}(s) & & & \\ & G_{S_1}(s) & & \\ & & \ddots & \\ & & & G_{S_1}(s) \end{bmatrix} \mathbf{V}(s) \\ &= G_{S_1}(s)\, \mathbf{V}^{-1}(s) \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \mathbf{V}(s) \\ &= G_{S_1}(s)\, \mathbf{V}^{-1}(s)\, \mathbb{I}_p\, \mathbf{V}(s) \\ &= G_{S_1}(s)\, \mathbb{I}_p \end{aligned} \tag{C.162}$$

Furthermore let $rdeg\,(C_{S_k}(s))$ be the relative degree of controller $C_{S_k}(s)$ such that (C.152) holds and let $rdeg\,(F_{S_k}(s))$ be the relative degree of transfer function $F_{S_k}(s)$. Denoting $r = max\,(rdeg\,(C_{S_k}(s)) + rdeg\,(F_{S_k}(s)))$, the relation $F_{S_i}(S)\, C_{S_i}(s) = F_{S_j}(S)\, C_{S_j}(s)\ \forall\, i, j$ implies that $rdeg\,(C_{S_k}(s))$ shall be chosen such that the following relation holds:

$$rdeg\,(C_{S_k}(s)) = r - rdeg\,(F_{S_k}(s)) \tag{C.163}$$

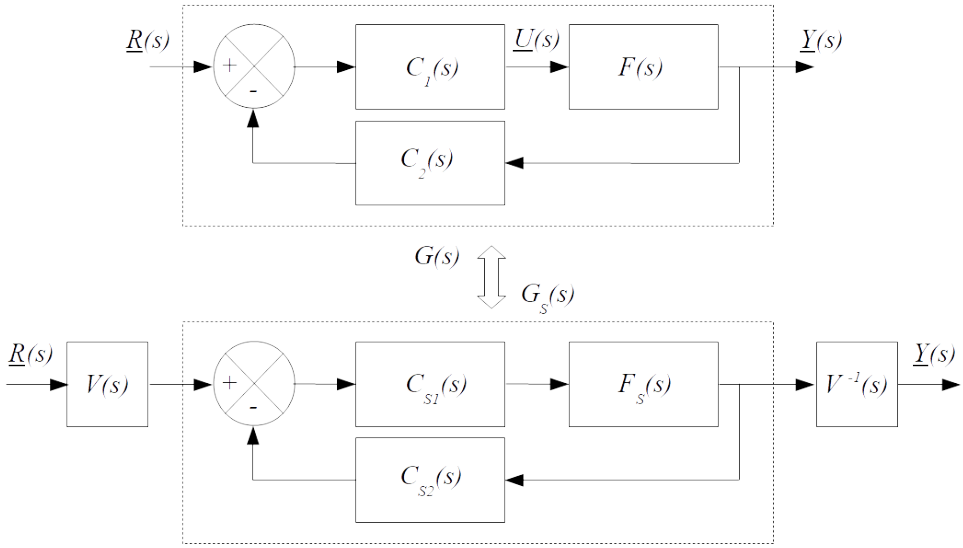

Figure C.2: Equivalent control loops with MIMO two-degree-of-freedom control loop

C.2.5 Two-degree-of-freedom control loop

Consider the configuration shown in Figure C.2 where $\mathbf{F}(s)$ is the plant transfer function and $\mathbf{C}_1(s)$ and $\mathbf{C}_2(s)$ two controllers. This feedback loop is called a two-degree-of-freedom control loop. From Figure C.2 control $\underline{u}(s)$ reads:

$$\underline{u}(s) = \mathbf{C}_1(s)\underline{r}(s) - \mathbf{C}_2(s)\underline{y}(s) \tag{C.164}$$

The relation between the output $\underline{y}(s)$ and the reference input $\underline{r}(s)$ reads:

$$\begin{aligned} &\underline{y}(s) = \mathbf{F}(s)\,\underline{u}(s) = \mathbf{F}(s)\,\left(\mathbf{C}_1(s)\,\underline{r}(s) - \mathbf{C}_2(s)\,\underline{y}(s)\right) \\ &\Rightarrow (\mathbb{I} + \mathbf{F}(s)\mathbf{C}_2(s))\;\underline{y}(s) = \mathbf{F}(s)\mathbf{C}_1(s)\,\underline{r}(s) \\ &\Leftrightarrow \underline{y}(s) = (\mathbb{I} + \mathbf{F}(s)\mathbf{C}_2(s))^{-1}\;\mathbf{F}(s)\mathbf{C}_1(s)\,\underline{r}(s) \end{aligned} \tag{C.165}$$

Thus the closed loop transfer function $\mathbf{G}(s)$ reads as follows:

$$\mathbf{G}(s) = (\mathbb{I} + \mathbf{F}(s)\mathbf{C}_2(s))^{-1}\;\mathbf{F}(s)\mathbf{C}_1(s) \tag{C.166}$$

Now as in section C.2.1, plant transfer function $\mathbf{F}(s)$ and controllers transfer function $\mathbf{C}_1(s)$ and $\mathbf{C}_2(s)$ are written as follows, where $\mathbf{C}_{S_1}(s)$ and $\mathbf{C}_{S_1}(s)$ are quasi-diagonal transfer functions:

$$\mathbf{F}(s) = \mathbf{V}^{-1}(s)\mathbf{F}_S(s)\mathbf{W}^{-1}(s) \Rightarrow \left\{ \begin{array}{l} \mathbf{C}_1(s) = \mathbf{W}(s)\mathbf{C}_{S_1}(s)\mathbf{V}(s) \\ \mathbf{C}_2(s) = \mathbf{W}(s)\mathbf{C}_{S_2}(s)\mathbf{V}(s) \end{array} \right. \tag{C.167}$$

Then the closed-loop transfer matrix $\mathbf{G}(s)$ becomes:

$$\begin{aligned} \mathbf{G}(s) &= \mathbf{V}^{-1}(s)\,(\mathbb{I} + \mathbf{F}_S(s)\mathbf{C}_{S_2}(s))^{-1}\,\mathbf{F}_S(s)\mathbf{C}_{S_1}(s)\mathbf{V}(s) \\ &:= \mathbf{V}^{-1}(s)\mathbf{G}_S(s)\mathbf{V}(s) \\ \text{where } \mathbf{G}_S(s) &= (\mathbb{I} + \mathbf{F}_S(s)\mathbf{C}_{S_2}(s))^{-1}\,\mathbf{F}_S(s)\mathbf{C}_{S_1}(s) \end{aligned} \tag{C.168}$$

The equivalent two-degree-of-freedom control loop using the Smith-McMillan form $\mathbf{F}_S(s)$ of plant $\mathbf{F}(s)$ is shown in the bottom of Figure C.2. The connection between $\mathbf{C}_i(s)$ and $\mathbf{C}_{S_i}(s)$ is obtained through the square unimodular polynomial matrices $\mathbf{V}(s)$ and $\mathbf{W}(s)$:

C.3 Matrix polynomials and solvents of λ-matrices

C.3.1 Latent roots and latent vectors

We present hereafter some results provided by Yaici & al.[10]. Let's consider the following r-degree, mth order monic matrix polynomials $\mathbf{D}(\lambda)$ where λ is a complex number and $\mathbf{D}_i \in \mathbb{R}^{m\times m}$:

$$\mathbf{D}(\lambda) = \mathbb{I}_m\lambda^r + \mathbf{D}_{r-1}\lambda^{r-1} + \cdots + \mathbf{D}_1\lambda + \mathbf{D}_0 \tag{C.169}$$

A latent root λ_i of $\mathbf{D}$ is a complex number satisfying:

$$\det\left(\mathbf{D}\left(\lambda_i\right)\right) = 0 \tag{C.170}$$

Matrix polynomials $\mathbf{D}(\lambda)$ has at most $m \times r$ latent roots. In the following we will assume that matrix polynomials $\mathbf{D}(\lambda)$ has exactly $m \times r$ latent roots $\{\lambda_1, \cdots, \lambda_{m\times r}\}$, including multiplicity.

A right latent vector corresponding to λ_i is a vector $\underline{v}_i \in \mathbb{R}^{m\times 1}$ belonging to the kernel of $\mathbf{D}(\lambda_i)$:

$$\mathbf{D}(\lambda_i)\,\underline{v}_i = \underline{0} \tag{C.171}$$

Similarly, a left latent vector is a row vector $\underline{w}_i \in \mathbb{R}^{m\times 1}$ belonging to the kernel of $\mathbf{D}(\lambda_i)^T$:

$$\mathbf{D}(\lambda_i)^T\,\underline{w}_i = \underline{0} \Leftrightarrow \underline{w}_i^T\,\mathbf{D}(\lambda_i) = \underline{0}^T \tag{C.172}$$

C.3.2 Right and left solvents

A right solvent of $\mathbf{D}(\lambda)$ is a square matrix $\mathbf{R} \in \mathbb{R}^{m\times m}$ satisfying:

$$\mathbf{R}^r + \mathbf{D}_{r-1}\mathbf{R}^{r-1} + \cdots + \mathbf{D}_1\mathbf{R} + \mathbf{D}_0 = \mathbf{0}_{m\times m} \tag{C.173}$$

Similarly, a left solvent of $\mathbf{D}(\lambda)$ is a square matrix $\mathbf{L} \in \mathbb{R}^{m\times m}$ satisfying:

$$\mathbf{L}^r + \mathbf{L}^{r-1}\mathbf{D}_{r-1} + \cdots + \mathbf{L}\mathbf{D}_1 + \mathbf{D}_0 = \mathbf{0}_{m\times m} \tag{C.174}$$

Assume that $\mathbf{D}(\lambda)$ have m linearly independent right latent vectors $\{\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_m\}$ and m linearly independent left latent vectors $\{\underline{w}_1, \underline{w}_2, \cdots, \underline{w}_m\}$ corresponding to latent roots $\{\lambda_1, \lambda_2, \cdots, \lambda_m\}$. Let $\mathbf{V}$ be the $m \times m$ matrix whose columns are the m linearly independent right latent

[10] Malika Yaici and Kamel Hariche, On Solvents of Matrix Polynomials, International Journal of Modeling and Optimization, Vol. 4, No. 4, August 2014

vectors $\underline{v}_i$ and $\mathbf{W}$ its inverse whose rows are the m linearly independent left latent vectors $\underline{w}_i$:

$$\begin{cases} \mathbf{V} = \begin{bmatrix} \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_m \end{bmatrix} \\ \mathbf{W} := \mathbf{V}^{-1} = \begin{bmatrix} \underline{w}_1^T \\ \underline{w}_2^T \\ \vdots \\ \underline{w}_m^T \end{bmatrix} \end{cases} \tag{C.175}$$

Then the following $m \times m$ matrices $\mathbf{R}$ and $\mathbf{L}$ are right and left solvents of $\mathbf{D}(\lambda)$, respectively:

$$\begin{cases} \mathbf{R} = \mathbf{V}\,\mathrm{diag}\,(\lambda_1, \cdots, \lambda_m)\,\mathbf{V}^{-1} \\ \mathbf{L} = \mathbf{W}^{-1}\,\mathrm{diag}\,(\lambda_1, \cdots, \lambda_m)\,\mathbf{W} \end{cases} \quad \text{where } \mathrm{diag}\,(\lambda_1, \cdots, \lambda_m) := \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix} \tag{C.176}$$

It can be shown that:

- Solvents of λ-matrices $\mathbf{D}(\lambda)$ do not always exist.
- If $\mathbf{R}$ is a right solvent of $\mathbf{D}(\lambda)$ then there exists a λ-matrix $\mathbf{Q}(\lambda)$ of degree $r-1$ such that $\mathbf{D}(\lambda) = \mathbf{Q}(\lambda)\,(\lambda\mathbb{I}_m - \mathbf{R})$.
- If $\mathbf{L}$ is a left solvent of $\mathbf{D}(\lambda)$ then there exists a λ-matrix $\mathbf{S}(\lambda)$ of degree $r-1$ such that $\mathbf{D}(\lambda) = (\lambda\mathbb{I}_m - \mathbf{L})\,\mathbf{S}(\lambda)$.
- Generalized right or left eigenvectors of a right or left solvent are the generalized latent vectors of $\mathbf{D}(\lambda)$.

C.3.3 Block Vandermonde matrices

Let split the set of $m \times r$ latent roots, including multiplicity, of $\mathbf{D}(\lambda)$ into r sets corresponding to m latent roots. For each set of m latent roots, let $\mathbf{R}_i$, $i = 1, \cdots, r$ and $\mathbf{L}_i$, $i = 1, \cdots, r$ be right and left solvents of $\mathbf{D}(\lambda)$, respectively. We denote $\lambda\,(\mathbf{R}_i)$ the set of eigenvalues of right solvent $\mathbf{R}_i$ and $\lambda\,(\mathbf{L}_i)$ the set of eigenvalues of left solvent $\mathbf{L}_i$. A complete set of right solvents and left solvents is obtained if we can find r right solvents and r left solvents such that:

$$\begin{cases} \bigcup_{i=1}^{r} \lambda\,(\mathbf{R}_i) = \{\lambda_1, \cdots, \lambda_{m\times r}\} \\ \bigcup_{i=1}^{r} \lambda\,(\mathbf{L}_i) = \{\lambda_1, \cdots, \lambda_{m\times r}\} \end{cases} \tag{C.177}$$

Furthermore the following block Vandermonde matrices $\mathbf{V}_R$ and $\mathbf{V}_L$ are

non-singulars:

$$\begin{cases} \mathbf{V}_R = \begin{bmatrix} \mathbb{I}_m & \mathbb{I}_m & \cdots & \mathbb{I}_m \\ \mathbf{R}_1 & \mathbf{R}_2 & \cdots & \mathbf{R}_r \\ \vdots & \vdots & & \vdots \\ \mathbf{R}_1^{r-1} & \mathbf{R}_2^{r-1} & \cdots & \mathbf{R}_r^{r-1} \end{bmatrix} \\ \mathbf{V}_L = \begin{bmatrix} \mathbb{I}_m & \mathbf{L}_1 & \cdots & \mathbf{L}_1^{r-1} \\ \mathbb{I}_m & \mathbf{L}_2 & \cdots & \mathbf{L}_2^{r-1} \\ \vdots & \vdots & & \vdots \\ \mathbb{I}_m & \mathbf{L}_r & \cdots & \mathbf{L}_r^{r-1} \end{bmatrix} \end{cases} \tag{C.178}$$

C.3.4 Modal form based on right solvents

Now consider the following linear time invariant realization with m inputs and p outputs:

$$\begin{cases} \dot{\underline{x}} = \mathbf{A}\underline{x} + \mathbf{B}\underline{u} \\ \underline{y} = \mathbf{C}\underline{x} \end{cases} \tag{C.179}$$

The corresponding state space equation in block controller form is the following:

$$\begin{aligned} \mathbf{F}(s) &= \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} \\ &= \left(\mathbf{N}_{r-1}s^{r-1} + \cdots + \mathbf{N}_1 s + \mathbf{N}_0\right)\left(\mathbb{I}_m s^r + \mathbf{D}_{r-1}s^{r-1} + \cdots + \mathbf{D}_1 s + \mathbf{D}_0\right)^{-1} \\ &= \mathbf{C}_c\left(s\mathbb{I} - \mathbf{A}_c\right)^{-1}\mathbf{B}_c := \left(\begin{array}{c|c} \mathbf{A}_c & \mathbf{B}_c \\ \hline \mathbf{C}_c & \mathbf{0}_{p\times m} \end{array}\right) \end{aligned} \tag{C.180}$$

where:

$$\begin{cases} \mathbf{A}_c = \begin{bmatrix} \mathbf{0}_m & \mathbb{I}_m & \mathbf{0}_m & \cdots & \mathbf{0}_m \\ \mathbf{0}_m & \mathbf{0}_m & \mathbb{I}_m & \cdots & \vdots \\ & & & \ddots & \\ & & & & \mathbb{I}_m \\ -\mathbf{D}_0 & -\mathbf{D}_1 & \cdots & & -\mathbf{D}_{r-1} \end{bmatrix} \\ \mathbf{B}_c = \begin{bmatrix} \mathbf{0}_m \\ \vdots \\ \mathbf{0}_m \\ \mathbb{I}_m \end{bmatrix} \\ \mathbf{C}_c = \begin{bmatrix} \mathbf{N}_0 & \mathbf{N}_1 & \cdots & \mathbf{N}_{r-1} \end{bmatrix} \end{cases} \tag{C.181}$$

Using the transformation $\underline{x}_m = \mathbf{V}_R\,\underline{x}_c$ where $\mathbf{V}_R$ is a non-singular right block Vandermonde matrix, we get the following modal form based on a complete set

of right solvents:

$$\begin{cases} \dot{\underline{x}}_m = \mathbf{A}_m \underline{x}_m + \mathbf{B}_m \underline{u} \\ \underline{y} = \mathbf{C}_m \underline{x} \end{cases} \quad \text{where} \quad \begin{cases} \mathbf{A}_m = \mathbf{V}_R^{-1} \mathbf{A}_c \mathbf{V}_R = \begin{bmatrix} \mathbf{R}_1 & \mathbf{0}_m & \cdots & \\ \mathbf{0}_m & \mathbf{R}_2 & \mathbf{0}_m & \cdots \\ & \ddots & \ddots & \ddots \\ \mathbf{0}_m & \mathbf{0}_m & \cdots & \mathbf{R}_r \end{bmatrix} \\ \mathbf{B}_m = \mathbf{V}_R^{-1} \mathbf{B}_c \\ \mathbf{C}_m = \mathbf{C}_c \mathbf{V}_R \end{cases} \tag{C.182}$$

It is worth noticing that:

- The state matrix $\mathbf{A}_m$ is block diagonal.
- Input matrix $\mathbf{B}_m$ is the last block column of $\mathbf{V}_R^{-1}$.
- The ith block column of $\mathbf{C}_m$ is $\mathbf{C}_{m_i} = \mathbf{N}_0 + \mathbf{N}_1 \mathbf{R}_i + \mathbf{N}_2 \mathbf{R}_i^2 + \cdots + \mathbf{N}_{r-1} \mathbf{R}_i^{r-1}$.

This leads to the following block partial fraction expansion of $\mathbf{F}(s)$ where $\mathbf{B}_{m_i}$ is i^{th} block row of $\mathbf{B}_m$:

$$\mathbf{F}(s) = \mathbf{C}_m \left(s\mathbb{I} - \mathbf{A}_m\right)^{-1} \mathbf{B}_m = \sum_{i=1}^{r} \mathbf{C}_{m_i} \left(s\mathbb{I}_m - \mathbf{R}_i\right)^{-1} \mathbf{B}_{m_i} \tag{C.183}$$

C.3.5 Modal form based on left solvents

Similarly, the state space equation in block observer form is the following:

$$\begin{aligned} \mathbf{F}(s) &= \mathbf{C} \left(s\mathbb{I} - \mathbf{A}\right)^{-1} \mathbf{B} \\ &= \left(\mathbb{I}_m s^r + \mathbf{D}_{r-1} s^{r-1} + \cdots + \mathbf{D}_1 s + \mathbf{D}_0\right)^{-1} \left(\mathbf{N}_{r-1} s^{r-1} + \cdots + \mathbf{N}_1 s + \mathbf{N}_0\right) \\ &= \mathbf{C}_o \left(s\mathbb{I} - \mathbf{A}_o\right)^{-1} \mathbf{B}_o := \left(\begin{array}{c|c} \mathbf{A}_o & \mathbf{B}_o \\ \hline \mathbf{C}_o & \mathbf{0}_{p\times m} \end{array}\right) \end{aligned} \tag{C.184}$$

where:

$$\begin{cases} \mathbf{A}_o = \begin{bmatrix} \mathbf{0}_m & \mathbf{0}_m & \cdots & \mathbf{0}_m & -\mathbf{D}_0 \\ \mathbb{I}_m & \mathbf{0}_m & \cdots & \mathbf{0}_m & -\mathbf{D}_1 \\ & \ddots & & & \\ \mathbf{0}_m & \cdots & & \mathbb{I}_m & -\mathbf{D}_{r-1} \end{bmatrix} \\ \mathbf{B}_o = \begin{bmatrix} \mathbf{N}_0 \\ \mathbf{N}_1 \\ \vdots \\ \mathbf{N}_{r-1} \end{bmatrix} \\ \mathbf{C}_o = \begin{bmatrix} \mathbf{0}_m & \cdots & \mathbf{0}_m & \mathbb{I}_m \end{bmatrix} \end{cases} \tag{C.185}$$

Using the transformation $\underline{x}_m = \mathbf{V}_L^{-1} \underline{x}_o$ where $\mathbf{V}_L$ is a non-singular left block Vandermonde matrix, we get the following modal form based on a complete set

of left solvents:

$$\begin{cases} \underline{\dot{x}}_m = \mathbf{A}_m \underline{x}_m + \mathbf{B}_m \underline{u} \\ \underline{y} = \mathbf{C}_m \underline{x} \end{cases} \quad \text{where} \quad \begin{cases} \mathbf{A}_m = \mathbf{V}_L \mathbf{A}_o \mathbf{V}_L^{-1} = \begin{bmatrix} \mathbf{L}_1 & \mathbf{0}_m & \cdots & \\ \mathbf{0}_m & \mathbf{L}_2 & \mathbf{0}_m & \cdots \\ & \ddots & \ddots & \ddots \\ \mathbf{0}_m & \mathbf{0}_m & \cdots & \mathbf{L}_r \end{bmatrix} \\ \mathbf{B}_m = \mathbf{V}_L \mathbf{B}_o \\ \mathbf{C}_m = \mathbf{C}_o \mathbf{V}_L^{-1} \end{cases} \tag{C.186}$$

It is worth noticing that:

- The state matrix $\mathbf{A}_m$ is block diagonal.
- Output matrix $\mathbf{C}_m$ is the last block row of $\mathbf{V}_L^{-1}$.
- The ith block row of $\mathbf{B}_m$ is $\mathbf{B}_{m_i} = \mathbf{N}_0 + \mathbf{L}_i \mathbf{N}_1 + \mathbf{L}_i^2 \mathbf{N}_2 + \cdots + \mathbf{L}_i^{r-1} \mathbf{N}_{r-1}$.

This leads to the following block partial fraction expansion of $\mathbf{F}(s)$ where $\mathbf{C}_{m_i}$ is i^{th} block column of $\mathbf{C}_m$:

$$\mathbf{F}(s) = \mathbf{C}_m \left(s\mathbb{I} - \mathbf{A}_m\right)^{-1} \mathbf{B}_m = \sum_{i=1}^{r} \mathbf{C}_{m_i} \left(s\mathbb{I}_m - \mathbf{L}_i\right)^{-1} \mathbf{B}_{m_i} \tag{C.187}$$

Appendix D

Singular perturbations and hierarchical control

D.1 Block triangular and block-diagonal forms

D.1.1 Block triangular form

Let's consider the following dynamical system arbitrarily partitioned as follows:

$$\boxed{\begin{cases} \begin{bmatrix} \dot{\underline{x}}_1 \\ \dot{\underline{x}}_2 \end{bmatrix} = \mathbf{A}\,\underline{x}(t) + \mathbf{B}\,\underline{u} := \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} \underline{u} \\ \underline{y} = \mathbf{C}\,\underline{x}(t) := \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} \end{cases}} \tag{D.1}$$

The preceding state space representation can be transformed into the following block triangular form[1]:

$$\begin{cases} \begin{bmatrix} \dot{\underline{x}}_1 \\ \dot{\underline{x}}_f \end{bmatrix} = \begin{bmatrix} \mathbf{A}_s & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_f \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_f \end{bmatrix} + \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_f \end{bmatrix} \underline{u} \\ \underline{y} = \begin{bmatrix} \mathbf{C}_s & \mathbf{C}_2 \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_f \end{bmatrix} \end{cases} \tag{D.2}$$

Where:

$$\begin{cases} \mathbf{A}_s = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{L} \\ \mathbf{A}_f = \mathbf{A}_{22} + \mathbf{L}\mathbf{A}_{12} \end{cases} \tag{D.3}$$

and:

$$\begin{cases} \mathbf{B}_f = \mathbf{L}\mathbf{B}_1 + \mathbf{B}_2 \\ \mathbf{C}_s = \mathbf{C}_1 - \mathbf{C}_2\mathbf{L} \end{cases} \tag{D.4}$$

Matrix $\mathbf{L}$ is a solution of the following non-symmetric algebraic Riccati equation:

$$\mathbf{L}\mathbf{A}_{11} - \mathbf{A}_{22}\mathbf{L} - \mathbf{L}\mathbf{A}_{12}\mathbf{L} + \mathbf{A}_{21} = \mathbf{0} \tag{D.5}$$

[1] Multi-Time Scale Systems, A.J.Fossard, IFAC Proceedings, Volume 17, Issue 2, July 1984, Pages 1139-1144

The corresponding similarity transformation is the following:

$$\begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{L} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_f \end{bmatrix} \Leftrightarrow \begin{bmatrix} \underline{x}_1 \\ \underline{x}_f \end{bmatrix} = \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ \mathbf{L} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} \tag{D.6}$$

It is worth noticing that the following relation holds as soon as square matrices $\mathbf{P}_{11}$ and $\mathbf{P}_{22}$ are invertible:

$$\begin{bmatrix} \mathbf{P}_{11} & \mathbf{0} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{P}_{11}^{-1} & \mathbf{0} \\ -\mathbf{P}_{22}^{-1}\mathbf{P}_{21}\mathbf{P}_{11}^{-1} & \mathbf{P}_{22}^{-1} \end{bmatrix}$$

D.1.2 Block-diagonal form

The block-diagonal form is obtained by introducing an additional similarity transformation:

$$\begin{bmatrix} \underline{x}_1 \\ \underline{x}_f \end{bmatrix} = \begin{bmatrix} \mathbb{I} & \mathbf{M} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \Leftrightarrow \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} = \begin{bmatrix} \mathbb{I} & -\mathbf{M} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_f \end{bmatrix} \tag{D.7}$$

We finally get the similarity transformation to the block-diagonal form:

$$\boxed{\begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{L} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \mathbb{I} & \mathbf{M} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} = \begin{bmatrix} \mathbb{I} & \mathbf{M} \\ -\mathbf{L} & \mathbb{I} - \mathbf{LM} \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix}} \tag{D.8}$$

Conversely:

$$\boxed{\begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} = \begin{bmatrix} \mathbb{I} & -\mathbf{M} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ \mathbf{L} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbb{I} - \mathbf{ML} & -\mathbf{M} \\ \mathbf{L} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix}} \tag{D.9}$$

The preceding similarity transformation leads to the following block-diagonal form:

$$\begin{cases} \begin{bmatrix} \dot{\underline{x}}_s \\ \dot{\underline{x}}_f \end{bmatrix} = \begin{bmatrix} \mathbf{A}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_f \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} + \begin{bmatrix} \mathbf{B}_s \\ \mathbf{B}_f \end{bmatrix} \underline{u} \\ \underline{y} = \begin{bmatrix} \mathbf{C}_s & \mathbf{C}_f \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \end{cases} \tag{D.10}$$

where:

$$\begin{cases} \begin{bmatrix} \mathbf{A}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_f \end{bmatrix} = \mathbf{P}^{-1} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \mathbf{P} \\ \begin{bmatrix} \mathbf{B}_s \\ \mathbf{B}_f \end{bmatrix} = \mathbf{P}^{-1} \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} \\ \begin{bmatrix} \mathbf{C}_s & \mathbf{C}_f \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix} \mathbf{P} \end{cases} \qquad \begin{cases} \mathbf{P} = \begin{bmatrix} \mathbb{I} & \mathbf{M} \\ -\mathbf{L} & \mathbb{I} - \mathbf{LM} \end{bmatrix} \\ \mathbf{P}^{-1} = \begin{bmatrix} \mathbb{I} - \mathbf{ML} & -\mathbf{M} \\ \mathbf{L} & \mathbb{I} \end{bmatrix} \end{cases}$$

$$\Leftrightarrow \begin{cases} \mathbf{A}_s = (\mathbb{I} - \mathbf{ML})\,\mathbf{A}_{11} - \mathbf{MA}_{21} + (\mathbf{MA}_{22} + (\mathbf{ML} - \mathbb{I})\,\mathbf{A}_{12})\,\mathbf{L} \\ \mathbf{A}_f = (\mathbf{A}_{21} + \mathbf{LA}_{11})\,\mathbf{M} + (\mathbf{A}_{22} + \mathbf{LA}_{12})\,(\mathbb{I} - \mathbf{LM}) \\ \mathbf{B}_s = \mathbf{B}_1 - \mathbf{MB}_f = \mathbf{B}_1 - \mathbf{M}\,(\mathbf{LB}_1 + \mathbf{B}_2) \\ \mathbf{C}_f = \mathbf{C}_s\mathbf{M} + \mathbf{C}_2 = (\mathbf{C}_1 - \mathbf{C}_2\mathbf{L})\,\mathbf{M} + \mathbf{C}_2 \end{cases} \tag{D.11}$$

Matrices $\mathbf{B}_f$ and $\mathbf{C}_s$ are still given by (D.4). Matrix $\mathbf{L}$ is still a solution of the non-symmetric algebraic Riccati equation (D.5). Thus matrices $\mathbf{A}_s$ and $\mathbf{A}_f$ simplify as follows:

$$\begin{aligned}
\mathbf{A}_s &= (\mathbb{I} - \mathbf{ML})\,\mathbf{A}_{11} - \mathbf{MA}_{21} + (\mathbf{MA}_{22} + (\mathbf{ML} - \mathbb{I})\,\mathbf{A}_{12})\,\mathbf{L} \\
&= \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{L} - \mathbf{M}\,(\mathbf{LA}_{11} + \mathbf{A}_{21} - \mathbf{A}_{22}\mathbf{L} - \mathbf{LA}_{12}\mathbf{L}) \\
&= \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{L} \\
\mathbf{A}_f &= (\mathbf{A}_{21} + \mathbf{LA}_{11})\,\mathbf{M} + (\mathbf{A}_{22} + \mathbf{LA}_{12})\,(\mathbb{I} - \mathbf{LM}) \\
&= \mathbf{A}_{22} + \mathbf{LA}_{12} + (\mathbf{A}_{21} + \mathbf{LA}_{11} - \mathbf{A}_{22}\mathbf{L} - \mathbf{LA}_{12}\mathbf{L})\,\mathbf{M} \\
&= \mathbf{A}_{22} + \mathbf{LA}_{12}
\end{aligned} \tag{D.12}$$

Finally matrix $\mathbf{M}$ is a solution of the following Sylvester equation:

$$\begin{aligned}
\mathbf{0} &= ((\mathbb{I} - \mathbf{ML})\,\mathbf{A}_{11} - \mathbf{MA}_{21})\,\mathbf{M} + (\mathbf{MA}_{22} + (\mathbf{ML} - \mathbb{I})\,\mathbf{A}_{12})\,(\mathbf{LM} - \mathbb{I}) \\
&= ((\mathbb{I} - \mathbf{ML})\,\mathbf{A}_{11} - \mathbf{MA}_{21} + (\mathbf{MA}_{22} + (\mathbf{ML} - \mathbb{I})\,\mathbf{A}_{12})\,\mathbf{L})\,\mathbf{M} \\
&\qquad -\mathbf{M}\,(\mathbf{A}_{22} + \mathbf{LA}_{12}) + \mathbf{A}_{12} \\
&= (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{L})\,\mathbf{M} - \mathbf{M}\,(\mathbf{A}_{22} + \mathbf{LA}_{12}) + \mathbf{A}_{12}
\end{aligned} \tag{D.13}$$

To summarize, we finally achieve the following block-diagonal form:

$$\boxed{\begin{cases}
\begin{bmatrix} \dot{\underline{x}}_s \\ \dot{\underline{x}}_f \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} + \mathbf{LA}_{12} \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \\
\qquad\qquad + \begin{bmatrix} \mathbf{B}_1 - \mathbf{M}\,(\mathbf{LB}_1 + \mathbf{B}_2) \\ \mathbf{LB}_1 + \mathbf{B}_2 \end{bmatrix} \underline{u} \\
\underline{y} = \begin{bmatrix} \mathbf{C}_1 - \mathbf{C}_2\mathbf{L} & (\mathbf{C}_1 - \mathbf{C}_2\mathbf{L})\,\mathbf{M} + \mathbf{C}_2 \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix}
\end{cases}} \tag{D.14}$$

where matrices $\mathbf{L}$ and $\mathbf{M}$ solve the following equations:

$$\boxed{\begin{cases}
\mathbf{LA}_{11} - \mathbf{A}_{22}\mathbf{L} - \mathbf{LA}_{12}\mathbf{L} + \mathbf{A}_{21} = \mathbf{0} \\
(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{L})\,\mathbf{M} - \mathbf{M}\,(\mathbf{A}_{22} + \mathbf{LA}_{12}) + \mathbf{A}_{12} = \mathbf{0}
\end{cases}} \tag{D.15}$$

It is useful to note that $\mathbf{L} = \mathbf{0}$ when $\mathbf{A}_{21} = \mathbf{0}$ and $\mathbf{M} = \mathbf{0}$ when $\mathbf{A}_{12} = \mathbf{0}$.

D.1.3 Similarity transformation

In order to get matrices $\mathbf{L}$ and $\mathbf{M}$, let $\chi_A(s)$ be the characteristic polynomial of matrix $\mathbf{A}$:

$$\chi_A(s) := \det(s\mathbb{I} - \mathbf{A}) = \det\left(s\mathbb{I} - \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}\right) \tag{D.16}$$

Let the n roots of $\chi_A(s)$ be split into two sets: the first set contains n_s roots, $\lambda_1, \cdots, \lambda_{n_s}$, which are dedicated to the roots of the characteristic polynomial of $\mathbf{A}_s$, whereas the second set contains $n - n_s$ roots, which are dedicated to the roots of the characteristic polynomial of $\mathbf{A}_f$. Then we can write:

$$\boxed{\chi_A(s) = \chi_{A_f}(s)\chi_{A_s}(s)} \tag{D.17}$$

Assuming that polynomials $\chi_{A_s}(s)$ and $\chi_{A_f}(s)$ are coprime (no common root), $(n-n_s)\times n_s$ matrix $\mathbf{L}$ and $n_s\times(n-n_s)$ matrix $\mathbf{M}$ can be obtained as follows[1]:

$$\begin{cases} \mathbf{L} = -\mathbf{T}\mathbf{S}^{-1} \\ \mathbf{M} = \mathbf{U}\left(\mathbf{V}+\mathbf{L}\mathbf{U}\right)^{-1} \end{cases} \tag{D.18}$$

Matrices $\mathbf{S}$, $\mathbf{T}$, $\mathbf{U}$ and $\mathbf{V}$ belongs to the nullspace (or Kernel) of $\chi_{A_s}(\mathbf{A})$ and $\chi_{A_f}(\mathbf{A})$ respectively (notice that in the characteristic polynomial the scalar variable s has been replaced by the $n\times n$ state matrix $\mathbf{A}$), each nullspace being partitioned appropriately:

$$\begin{cases} \begin{bmatrix} \mathbf{S} \\ \mathbf{T} \end{bmatrix} = \ker\left(\chi_{A_s}(\mathbf{A})\right) \\ \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix} = \ker\left(\chi_{A_f}(\mathbf{A})\right) \end{cases} \tag{D.19}$$

Furthermore the similarity transformation to the block-diagonal form reads as follows[1]:

$$\begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} := \mathbf{P}\begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \tag{D.20}$$

Where:

$$\boxed{\mathbf{P} := \begin{bmatrix} \mathbb{I} & \mathbf{M} \\ -\mathbf{L} & \mathbb{I}-\mathbf{L}\mathbf{M} \end{bmatrix} = \left[\begin{array}{c:c} \begin{bmatrix} \mathbf{S} \\ \mathbf{T} \end{bmatrix}\mathbf{S}^{-1} & \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix}\left(\mathbf{V}+\mathbf{L}\mathbf{U}\right)^{-1} \end{array}\right]} \tag{D.21}$$

Let n_s be the size of state vector $\underline{x}_s$ and n_f be the size of state vector $\underline{x}_f$. Then $\mathbf{L}$ is an $n_f\times n_s$ matrix whereas $\mathbf{M}$ is an $n_s\times n_f$ matrix.

D.2 Two-stage design for eigenvalue placement of block-diagonal systems

We consider hereafter the following block-diagonal state space representation (D.14) of a system:

$$\begin{bmatrix} \dot{\underline{x}}_s \\ \dot{\underline{x}}_f \end{bmatrix} = \begin{bmatrix} \mathbf{A}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_f \end{bmatrix}\begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} + \begin{bmatrix} \mathbf{B}_s \\ \mathbf{B}_f \end{bmatrix}\underline{u} \tag{D.22}$$

where:

$$\begin{cases} \mathbf{A}_s = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{L} \\ \mathbf{A}_f = \mathbf{A}_{22} + \mathbf{L}\mathbf{A}_{12} \\ \mathbf{B}_s = \mathbf{B}_1 - \mathbf{M}\left(\mathbf{L}\mathbf{B}_1 + \mathbf{B}_2\right) \\ \mathbf{B}_f = \mathbf{L}\mathbf{B}_1 + \mathbf{B}_2 \end{cases} \tag{D.23}$$

The block diagram representation corresponding to this block diagonal state space representation is shown in Figure D.1.

Following Phillips[2], the steps of the design dedicated to eigenvalue placement are the following:

[2] R. Phillips, A two-stage design of linear feedback controls, IEEE Transactions on Automatic Control, vol. 25, no. 6, pp. 1220-1223, December 1980, doi: 10.1109/TAC.1980.1102548.

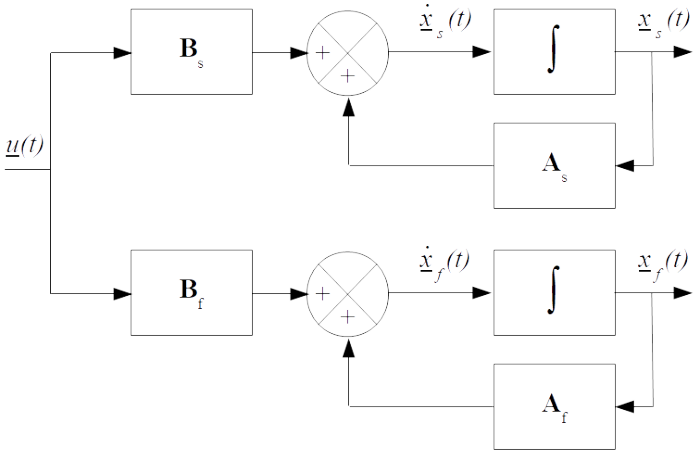

Figure D.1: Block diagram representation of a block diagonal system

– In the first stage of the design a state feedback gain $\mathbf{K}_f$ is set to place the eigenvalues of $\mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f$ at desired location:

$$\begin{aligned} \underline{u} &= \underline{u}_s - \begin{bmatrix} \mathbf{0} & \mathbf{K}_f \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \\ &\Rightarrow \begin{bmatrix} \dot{\underline{x}}_s \\ \dot{\underline{x}}_f \end{bmatrix} = \begin{bmatrix} \mathbf{A}_s & -\mathbf{B}_s\mathbf{K}_f \\ \mathbf{0} & \mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} + \begin{bmatrix} \mathbf{B}_s \\ \mathbf{B}_f \end{bmatrix} \underline{u}_s \end{aligned} \tag{D.24}$$

– Then transformation (D.9) is applied where it is worth noticing that $\widetilde{\mathbf{L}} = \mathbf{0}$ because the block at location $(2,1)$ in the closed-loop state matrix is zero:

$$\begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} = \begin{bmatrix} \mathbb{I} & \widetilde{\mathbf{M}} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \widetilde{\underline{x}}_s \\ \widetilde{\underline{x}}_f \end{bmatrix} \Leftrightarrow \begin{bmatrix} \widetilde{\underline{x}}_s \\ \widetilde{\underline{x}}_f \end{bmatrix} = \begin{bmatrix} \mathbb{I} & -\widetilde{\mathbf{M}} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \tag{D.25}$$

We get:

$$\begin{aligned} \begin{bmatrix} \dot{\widetilde{\underline{x}}}_s \\ \dot{\widetilde{\underline{x}}}_f \end{bmatrix} &= \begin{bmatrix} \mathbb{I} & -\widetilde{\mathbf{M}} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}_s & -\mathbf{B}_s\mathbf{K}_f \\ \mathbf{0} & \mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f \end{bmatrix} \begin{bmatrix} \mathbb{I} & \widetilde{\mathbf{M}} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \widetilde{\underline{x}}_s \\ \widetilde{\underline{x}}_f \end{bmatrix} \\ &\quad + \begin{bmatrix} \mathbb{I} & -\widetilde{\mathbf{M}} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \mathbf{B}_s \\ \mathbf{B}_f \end{bmatrix} \underline{u}_s \\ &= \begin{bmatrix} \mathbb{I} & -\widetilde{\mathbf{M}} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \mathbf{A}_s & \mathbf{A}_s\widetilde{\mathbf{M}} - \mathbf{B}_s\mathbf{K}_f \\ \mathbf{0} & \mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f \end{bmatrix} \begin{bmatrix} \widetilde{\underline{x}}_s \\ \widetilde{\underline{x}}_f \end{bmatrix} \\ &\quad + \begin{bmatrix} \mathbf{B}_s - \widetilde{\mathbf{M}}\mathbf{B}_f \\ \mathbf{B}_f \end{bmatrix} \underline{u}_s \\ &= \begin{bmatrix} \mathbf{A}_s & \mathbf{A}_s\widetilde{\mathbf{M}} - \widetilde{\mathbf{M}}(\mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f) - \mathbf{B}_s\mathbf{K}_f \\ \mathbf{0} & \mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f \end{bmatrix} \begin{bmatrix} \widetilde{\underline{x}}_s \\ \widetilde{\underline{x}}_f \end{bmatrix} \\ &\quad + \begin{bmatrix} \mathbf{B}_s - \widetilde{\mathbf{M}}\mathbf{B}_f \\ \mathbf{B}_f \end{bmatrix} \underline{u}_s \end{aligned} \tag{D.26}$$

Matrix $\widetilde{\mathbf{M}}$ is then defined as a solution of the following Lyapunov equation, as seen in (D.15):

$$\mathbf{A}_s\widetilde{\mathbf{M}} - \widetilde{\mathbf{M}}\left(\mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f\right) - \mathbf{B}_s\mathbf{K}_f = \mathbf{0} \tag{D.27}$$

We finally get, using (D.27):

$$\begin{bmatrix} \dot{\underline{\widetilde{x}}}_s \\ \dot{\underline{\widetilde{x}}}_f \end{bmatrix} = \begin{bmatrix} \mathbf{A}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f \end{bmatrix} \begin{bmatrix} \underline{\widetilde{x}}_s \\ \underline{\widetilde{x}}_f \end{bmatrix} + \begin{bmatrix} \mathbf{B}_s - \widetilde{\mathbf{M}}\mathbf{B}_f \\ \mathbf{B}_f \end{bmatrix} \underline{u}_s \tag{D.28}$$

– Then in the second stage of the design a state feedback gain $\mathbf{K}_s$ is set to place the eigenvalues of $\mathbf{A}_s - \left(\mathbf{B}_s - \widetilde{\mathbf{M}}\mathbf{B}_f\right)\mathbf{K}_s$ at desired location:

$$\begin{aligned} \underline{u}_s &= \underline{r} - \begin{bmatrix} \mathbf{K}_s & \mathbf{0} \end{bmatrix} \begin{bmatrix} \underline{\widetilde{x}}_s \\ \underline{\widetilde{x}}_f \end{bmatrix} \\ &\Rightarrow \begin{bmatrix} \dot{\underline{\widetilde{x}}}_s \\ \dot{\underline{\widetilde{x}}}_f \end{bmatrix} = \mathbf{A}_{cl} \begin{bmatrix} \underline{\widetilde{x}}_s \\ \underline{\widetilde{x}}_f \end{bmatrix} + \begin{bmatrix} \mathbf{B}_s - \widetilde{\mathbf{M}}\mathbf{B}_f \\ \mathbf{B}_f \end{bmatrix} \underline{r} \end{aligned} \tag{D.29}$$

where:

$$\begin{aligned} \widetilde{\mathbf{A}}_{cl} &= \begin{bmatrix} \mathbf{A}_s - \left(\mathbf{B}_s - \widetilde{\mathbf{M}}\mathbf{B}_f\right)\mathbf{K}_s & \mathbf{0} \\ -\mathbf{B}_f\mathbf{K}_s & \mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A}_s - \mathbf{B}_s\mathbf{K}_s + \widetilde{\mathbf{M}}\mathbf{B}_f\mathbf{K}_s & \mathbf{0} \\ -\mathbf{B}_f\mathbf{K}_s & \mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f \end{bmatrix} \end{aligned} \tag{D.30}$$

The preceding relation indicates that the closed-loop eigenvalues are obtained by the union of two sets: the set of the *fast* closed-loop eigenvalues $\lambda\left(\mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f\right)$ and the set of the *slow* closed-loop eigenvalues $\lambda\left(\mathbf{A}_s - \mathbf{B}_s\mathbf{K}_s + \widetilde{\mathbf{M}}\mathbf{B}_f\mathbf{K}_s\right)$:

$$\lambda\left(\widetilde{\mathbf{A}}_{cl}\right) = \lambda\left(\mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f\right) \cup \lambda\left(\mathbf{A}_s - \mathbf{B}_s\mathbf{K}_s + \widetilde{\mathbf{M}}\mathbf{B}_f\mathbf{K}_s\right) \tag{D.31}$$

– The control law is finally obtained by taking into account the control obtained during the first stage into the control obtained during the first stage. Using (D.9) we get the control $\mathbf{K}$ corresponding to state vector $\begin{bmatrix} \underline{x}_1^T & \underline{x}_2^T \end{bmatrix}^T$:

$$\begin{aligned} \underline{u} &= \underline{u}_s - \begin{bmatrix} \mathbf{0} & \mathbf{K}_f \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \\ &= \underline{r} - \begin{bmatrix} \mathbf{K}_s & \mathbf{0} \end{bmatrix} \begin{bmatrix} \underline{\widetilde{x}}_s \\ \underline{\widetilde{x}}_f \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{K}_f \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \\ &= \underline{r} - \left(\begin{bmatrix} \mathbf{K}_s & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbb{I} & -\widetilde{\mathbf{M}} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{K}_f \end{bmatrix} \right) \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \\ &= \underline{r} - \begin{bmatrix} \mathbf{K}_s & \mathbf{K}_f - \mathbf{K}_s\widetilde{\mathbf{M}} \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \\ &= \underline{r} - \begin{bmatrix} \mathbf{K}_s & \mathbf{K}_f - \mathbf{K}_s\widetilde{\mathbf{M}} \end{bmatrix} \mathbf{P}^{-1} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} \\ &:= \underline{r} - \mathbf{K} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} \end{aligned} \tag{D.32}$$

where:

$$\begin{aligned}\mathbf{K} &= \begin{bmatrix} \mathbf{K}_s & \mathbf{K}_f - \mathbf{K}_s\widetilde{\mathbf{M}} \end{bmatrix} \mathbf{P}^{-1} \\ &= \begin{bmatrix} \mathbf{K}_s & \mathbf{K}_f - \mathbf{K}_s\widetilde{\mathbf{M}} \end{bmatrix} \begin{bmatrix} \mathbb{I} - \mathbf{ML} & -\mathbf{M} \\ \mathbf{L} & \mathbb{I} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{K}_s\left(\mathbb{I} - \mathbf{ML}\right) + \left(\mathbf{K}_f - \mathbf{K}_s\widetilde{\mathbf{M}}\right)\mathbf{L} & -\mathbf{K}_s\mathbf{M} + \mathbf{K}_f - \mathbf{K}_s\widetilde{\mathbf{M}} \end{bmatrix}\end{aligned} \tag{D.33}$$

According to (D.30) and (D.9), the closed-loop state matrix $\mathbf{A}_{cl}$ corresponding to state vector $\begin{bmatrix} \underline{x}_1^T & \underline{x}_2^T \end{bmatrix}^T$ reads $\mathbf{A}_{cl} = \widetilde{\mathbf{P}}^{-1}\widetilde{\mathbf{A}}_{cl}\widetilde{\mathbf{P}}$ where matrix $\widetilde{\mathbf{P}}$ is defined as follows:

$$\begin{aligned}\begin{bmatrix} \underline{\widetilde{x}}_s \\ \underline{\widetilde{x}}_f \end{bmatrix} &= \begin{bmatrix} \mathbb{I} & -\widetilde{\mathbf{M}} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \\ &= \begin{bmatrix} \mathbb{I} & -\widetilde{\mathbf{M}} \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \mathbb{I} - \mathbf{ML} & -\mathbf{M} \\ \mathbf{L} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} \\ &= \begin{bmatrix} \mathbb{I} - \left(\mathbf{M} + \widetilde{\mathbf{M}}\right)\mathbf{L} & -\left(\mathbf{M} + \widetilde{\mathbf{M}}\right) \\ \mathbf{L} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} \\ &:= \widetilde{\mathbf{P}} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix}\end{aligned} \tag{D.34}$$

Example D.1. *We consider the following state space representation, which models the linearized longitudinal dynamics at a trimmed flight condition of a jet liner. In the following, V_p stands for the true airspeed, α the angle of attack, θ the pitch angle, q the pitch rate and δ_e the elevator deflection:*

$$\frac{d}{dt}\begin{bmatrix} V_p \\ \alpha \\ \theta \\ q \end{bmatrix} = \mathbf{A}\begin{bmatrix} V_p \\ \alpha \\ \theta \\ q \end{bmatrix} + \mathbf{B}\,\delta_e \tag{D.35}$$

Where:

$$\begin{cases} \mathbf{A} = \begin{bmatrix} -1.4900\times10^{-2} & 5.8649 & -9.8059 & -6.8000\times10^{-2} \\ -3.0000\times10^{-4} & -1.5863 & 0.0000 & 9.7250\times10^{-1} \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \\ 0.0000 & -4.9799 & 0.0000 & -2.2514 \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} -0.7137 \\ -0.2886 \\ 0. \\ -23.6403 \end{bmatrix} \end{cases} \tag{D.36}$$

The eigenvalues of $\mathbf{A}$ *are* $\{\lambda_f, \bar{\lambda}_f = -1.919 \pm 2.176j\}$, *which are the* fast *eigenvalues, and* $\{\lambda_s, \bar{\lambda}_s = -7.293\times10^{-3} \pm 4.108\times10^{-2}j\}$, *which are the* slow *eigenvalues.*

The block-diagonal form of (D.36) *is obtained with the similarity transformation* $\mathbf{P}$ *defined in* (D.21)*:*

$$\mathbf{P} = \begin{bmatrix} 1.000 & 0.000 & 2.809 \times 10^{-4} & -3.420 \times 10^{-1} \\ 0.000 & 1.000 & 8.122 \times 10^{-6} & 2.900 \times 10^{-5} \\ -9.857 & -1.228 \times 10^{5} & 1.549 \times 10^{-4} & -1.885 \times 10^{-1} \\ -7.671 \times 10^{-3} & -9.776 \times 10^{1} & -7.962 \times 10^{-4} & 9.998 \times 10^{-1} \end{bmatrix} \tag{D.37}$$

We get:

$$\begin{cases} \mathbf{A}_f = \begin{bmatrix} -96.657765 & 119379.91 \\ -0.0752234 & 92.819752 \end{bmatrix} \\ \mathbf{B}_f = \begin{bmatrix} -35434.23 \\ -51.859333 \end{bmatrix} \\ \mathbf{A}_s = \begin{bmatrix} 96.643387 & 1203739.2 \\ -0.0077603 & -96.657974 \end{bmatrix} \\ \mathbf{B}_s = \begin{bmatrix} -8.4969422 \\ 0.0007172 \end{bmatrix} \end{cases} \tag{D.38}$$

Assume that we wish to achieve the following closed-loop eigenvalues:

$$\begin{cases} \lambda_{clf}, \bar{\lambda}_{clf} = -1 \pm j \\ \lambda_{cls}, \bar{\lambda}_{cls} = -0.01 \pm 0.01j \end{cases} \tag{D.39}$$

It is worth noticing that the choice of the closed-loop eigenvalues maintain the distinction between the slow *and the* fast *modes.*

Then state feedback gains $\mathbf{K}_f$ *and* $\mathbf{K}_s$ *are set as follows:*

- *State feedback gain* $\mathbf{K}_f$ *is set such that the eigenvalues of* $\mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f$ *are equal to* $\{\lambda_{clf}, \bar{\lambda}_{clf}\}$*. This leads to the following value of* $\mathbf{K}_f$*:*

$$\mathbf{K}_f = \begin{bmatrix} -5.914 \times 10^{-5} & 7.585 \times 10^{-2} \end{bmatrix} \tag{D.40}$$

- *Matrix* $\widetilde{\mathbf{M}}$ *is then defined as the solution of Lyapunov equation* (D.27)*. We get:*

$$\begin{aligned} &\widetilde{\mathbf{M}}\left(\mathbf{A}_f - \mathbf{B}_f\mathbf{K}_f\right) - \mathbf{A}_s\widetilde{\mathbf{M}} + \mathbf{B}_s\mathbf{K}_f = \mathbf{0} \\ &\Rightarrow \widetilde{\mathbf{M}} = \begin{bmatrix} 4.255 \times 10^{-3} & -5.301 \\ -3.454 \times 10^{-7} & 4.304 \times 10^{-4} \end{bmatrix} \end{aligned} \tag{D.41}$$

- *In the second stage, state feedback gain* $\mathbf{K}_s$ *is set such that the eigenvalues of* $\mathbf{A}_s - \left(\mathbf{B}_s - \widetilde{\mathbf{M}}\mathbf{B}_f\right)\mathbf{K}_s$ *are equal to* $\{\lambda_{cls}, \bar{\lambda}_{cls}\}$*. This leads to the following value of* $\mathbf{K}_s$*:*

$$\mathbf{K}_s = \begin{bmatrix} 2.291 \times 10^{-3} & 2.865 \times 10^{1} \end{bmatrix} \tag{D.42}$$

State feedback gain $\mathbf{K}$ *for the actual system is finally obtained according to* (D.33)*:*

$$\begin{aligned} \mathbf{K} &= \begin{bmatrix} \mathbf{K}_s & \mathbf{K}_f - \mathbf{K}_s\widetilde{\mathbf{M}} \end{bmatrix} \mathbf{P}^{-1} \\ &= \begin{bmatrix} -1.011 \times 10^{-5} & 1.559 \times 10^{-1} & -2.923 \times 10^{-4} & 7.562 \times 10^{-2} \end{bmatrix} \end{aligned} \tag{D.43}$$

We check that the eigenvalues of $\mathbf{A} - \mathbf{BK}$ *of the whole system are exactly the expected eigenvalues* $\{\lambda_{clf}, \bar{\lambda}_{clf}, \lambda_{cls}, \bar{\lambda}_{cls}\}$.

■

D.3 Singularly perturbed system

D.3.1 Linear system

Dynamical system (D.1) is assumed to be *stable*. Let $\chi_A(s)$ be the characteristic polynomial of state matrix $\mathbf{A}$ and $\lambda(\chi_A(s))$ the roots of $\chi_A(s)$. Then dynamical system (D.1) is said to be singularly perturbed if $\chi_A(s)$ can be split as follows:

$$\begin{gathered}\chi_A(s) = \chi_{A_f}(s)\chi_{A_s}(s) \\ \text{where } \max\left(\mathrm{Re}\left(\lambda\left(\chi_{A_f}(s)\right)\right)\right) \ll \min\left(\mathrm{Re}\left(\lambda\left(\chi_{A_s}(s)\right)\right)\right) < 0\end{gathered} \tag{D.44}$$

Small number ϵ is related to the value of n_s which delimits the border between the *slow* and the *fast* modes.

Let $\epsilon \geq 0$ be a *small* number which may be defined as follows:

$$\epsilon = \frac{\min\left(\mathrm{Re}\left(\lambda\left(\chi_{A_s}(s)\right)\right)\right)}{\max\left(\mathrm{Re}\left(\lambda\left(\chi_{A_f}(s)\right)\right)\right)} \approx 0 \tag{D.45}$$

Alternatively, ϵ may be defined as the minimum of $\frac{|\lambda_{n_s}|}{|\lambda_{n_s+1}|}$, assuming that the real part of the eigenvalues λ_i of the open loop state matrix $\mathbf{A}$ are sorted in a descending manner.

Let $\mathbf{A}$ be a state matrix. The corresponding state vector can re-ordered using the following permutation matrix $\mathbf{T}$ where $\underline{e}_i^T$ is an elementary row vector whose i^{th} entry is 1:

$$\mathbf{T} = \begin{bmatrix} \underline{e}_{r_1}^T \\ \vdots \\ \underline{e}_{r_{ns}}^T \\ \underline{e}_{r_{ns+1}}^T \\ \vdots \\ \underline{e}_{r_n}^T \end{bmatrix} \quad \text{where} \quad \underline{e}_i^T = [0, \cdots, 0, \underbrace{1}_{i^{th}\text{ column}}, 0, \cdots, 0] \tag{D.46}$$

Since $\mathbf{T}$ is a permutation matrix we have $\mathbf{T}^{-1} = \mathbf{T}^T$. Assume that $\underline{\widetilde{x}} = \mathbf{T}\,\underline{x}$ corresponds to the following state space representation where $\epsilon \geq 0$ is a *small* number, $\widetilde{\mathbf{A}}_{11}$ is an $n_s \times n_s$ matrix and $\widetilde{\mathbf{A}}_{22}$ is an $(n - n_s) \times (n - n_s)$ matrix:

$$\boxed{\underline{\widetilde{x}} = \mathbf{T}\,\underline{x} \Rightarrow \begin{cases} \dot{\underline{\widetilde{x}}} = \mathbf{TAT}^T\,\underline{\widetilde{x}} + \mathbf{TB}\,\underline{u} := \begin{bmatrix} \widetilde{\mathbf{A}}_{11} & \widetilde{\mathbf{A}}_{12} \\ \frac{\widetilde{\mathbf{A}}_{21}}{\epsilon} & \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon} \end{bmatrix} \underline{\widetilde{x}} + \begin{bmatrix} \widetilde{\mathbf{B}}_1 \\ \frac{\widetilde{\mathbf{B}}_2}{\epsilon} \end{bmatrix} \underline{u} \\ \underline{y} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix} \mathbf{T}^T\,\underline{\widetilde{x}} := \begin{bmatrix} \widetilde{\mathbf{C}}_1 & \widetilde{\mathbf{C}}_2 \end{bmatrix} \underline{\widetilde{x}} \end{cases}} \tag{D.47}$$

Following Khalil[3], the system possesses a two-time scale property if the eigenvalues of $\widetilde{\mathbf{A}}$ can then be approximated as follows:

$$\boxed{\lambda\left(\widetilde{\mathbf{A}}\right) := \lambda\left(\begin{bmatrix} \widetilde{\mathbf{A}}_{11} & \widetilde{\mathbf{A}}_{12} \\ \frac{\widetilde{\mathbf{A}}_{21}}{\epsilon} & \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon} \end{bmatrix}\right) \approx \lambda\left(\frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right) \cup \lambda\left(\widetilde{\mathbf{A}}_0\right)} \tag{D.48}$$

where, assuming that matrix $\frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}$ is invertible:

$$\boxed{\widetilde{\mathbf{A}}_0 = \widetilde{\mathbf{A}}_{11} - \widetilde{\mathbf{A}}_{12}\left(\frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right)^{-1}\frac{\widetilde{\mathbf{A}}_{21}}{\epsilon} = \widetilde{\mathbf{A}}_{11} - \widetilde{\mathbf{A}}_{12}\widetilde{\mathbf{A}}_{22}^{-1}\widetilde{\mathbf{A}}_{21}} \tag{D.49}$$

The core of this result is the Schur's formula, which is stated hereafter:

$$\det\left(\begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{bmatrix}\right) = \det\left(\mathbf{X}_{22}\right)\det\left(\mathbf{X}_{11} - \mathbf{X}_{12}\mathbf{X}_{22}^{-1}\mathbf{X}_{21}\right) \tag{D.50}$$

The Schur's formula applied to the closed-loop state matrix $\mathbf{A}_{cl}$ reads:

$$\begin{aligned} &\det\left(\begin{bmatrix} s\mathbb{I} - \widetilde{\mathbf{A}}_{11} & -\widetilde{\mathbf{A}}_{12} \\ -\frac{\widetilde{\mathbf{A}}_{21}}{\epsilon} & s\mathbb{I} - \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon} \end{bmatrix}\right) \\ &\quad = \det\left(s\mathbb{I} - \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right)\det\left(s\mathbb{I} - \widetilde{\mathbf{A}}_{11} - \widetilde{\mathbf{A}}_{12}\left(s\mathbb{I} - \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right)^{-1}\frac{\widetilde{\mathbf{A}}_{21}}{\epsilon}\right) \end{aligned} \tag{D.51}$$

When $\epsilon \to 0$, we can write $s\mathbb{I} - \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon} \underbrace{\approx}_{\epsilon\to 0} -\frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}$. Then result (D.48) is obtained as follows:

$$\begin{aligned} &\det\left(\begin{bmatrix} s\mathbb{I} - \widetilde{\mathbf{A}}_{11} & \widetilde{\mathbf{A}}_{12} \\ \frac{\widetilde{\mathbf{A}}_{21}}{\epsilon} & s\mathbb{I} - \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon} \end{bmatrix}\right) \\ &\quad \approx \det\left(s\mathbb{I} - \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right)\det\left(s\mathbb{I} - \widetilde{\mathbf{A}}_{11} - \widetilde{\mathbf{A}}_{12}\left(-\frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right)^{-1}\frac{\widetilde{\mathbf{A}}_{21}}{\epsilon}\right) \\ &\quad = \det\left(s\mathbb{I} - \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right)\det\left(s\mathbb{I} - \widetilde{\mathbf{A}}_{11} + \widetilde{\mathbf{A}}_{12}\widetilde{\mathbf{A}}_{22}^{-1}\widetilde{\mathbf{A}}_{21}\right) \\ &\quad = \det\left(s\mathbb{I} - \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right)\det\left(s\mathbb{I} - \widetilde{\mathbf{A}}_{11} + \widetilde{\mathbf{A}}_{12}\left(\frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right)^{-1}\frac{\widetilde{\mathbf{A}}_{21}}{\epsilon}\right) \end{aligned} \tag{D.52}$$

This completes the proof. ■

Assuming that system possesses a two-time scale property and that $\frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}$ is not singular, a recursive determination of $\mathbf{L}$ and $\mathbf{M}$ can be obtained from (D.15) as follows:

$$\begin{cases} \mathbf{L}_{k+1} = \left(\frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right)^{-1}\left(\mathbf{L}_k\widetilde{\mathbf{A}}_{11} - \mathbf{L}_k\widetilde{\mathbf{A}}_{12}\mathbf{L}_k + \left(\frac{\widetilde{\mathbf{A}}_{21}}{\epsilon}\right)\right) \\ \mathbf{M}_{k+1} = \left(\left(\widetilde{\mathbf{A}}_{11} - \widetilde{\mathbf{A}}_{12}\mathbf{L}\right)\mathbf{M}_k - \mathbf{M}_k\mathbf{L}\widetilde{\mathbf{A}}_{12} + \widetilde{\mathbf{A}}_{12}\right)\left(\frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right)^{-1} \end{cases} \tag{D.53}$$

[3]Hassan K. Khalil, On the robustness of output feedback control methods to modeling errors, IEEE Transactions on Automatic Control, Vol. AC-26, April 1981, pp 524-526

where:

$$\boxed{\begin{cases} \mathbf{L}_0 = \left(\frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right)^{-1} \frac{\widetilde{\mathbf{A}}_{21}}{\epsilon} = \widetilde{\mathbf{A}}_{22}^{-1}\,\widetilde{\mathbf{A}}_{21} \\ \mathbf{M}_0 = \widetilde{\mathbf{A}}_{12}\left(\frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right)^{-1} = \epsilon\,\widetilde{\mathbf{A}}_{12}\,\widetilde{\mathbf{A}}_{22}^{-1} \end{cases}} \tag{D.54}$$

Usually, matrices $\mathbf{L}_0$ and $\mathbf{M}_0$ are sufficient to approximate matrices $\mathbf{L}$ and $\mathbf{M}$, respectively[4] . Then the similarity transformation $\mathbf{P}$ defined in (D.21) can be approximated as follows:

$$\begin{cases} \mathbf{P} &= \begin{bmatrix} \mathbb{I} & \mathbf{M} \\ -\mathbf{L} & \mathbb{I}-\mathbf{LM} \end{bmatrix} \approx \begin{bmatrix} \mathbb{I} & \mathbf{0} \\ -\mathbf{L}_0 & \mathbb{I} \end{bmatrix} \\ \mathbf{P}^{-1} &= \begin{bmatrix} \mathbb{I}-\mathbf{ML} & -\mathbf{M} \\ \mathbf{L} & \mathbb{I} \end{bmatrix} \approx \begin{bmatrix} \mathbb{I} & -\mathbf{M}_0 \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \end{cases} \tag{D.55}$$

Consequently the block diagonal state space representation (D.10) can be approximated as follows where $\epsilon \geq 0$ is a *small* number:

$$\boxed{\begin{cases} \begin{bmatrix} \dot{\underline{x}}_s \\ \dot{\underline{x}}_f \end{bmatrix} \approx \begin{bmatrix} \widetilde{\mathbf{A}}_{11} - \widetilde{\mathbf{A}}_{12}\mathbf{L}_0 & \mathbf{0} \\ \mathbf{0} & \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon} \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \\ \qquad\qquad + \begin{bmatrix} \widetilde{\mathbf{B}}_1 - \mathbf{M}_0\frac{\widetilde{\mathbf{B}}_2}{\epsilon} \\ \frac{\widetilde{\mathbf{B}}_2}{\epsilon} \end{bmatrix} \underline{u} \\ \underline{y} \approx \begin{bmatrix} \widetilde{\mathbf{C}}_1 - \widetilde{\mathbf{C}}_2\mathbf{L}_0 & \widetilde{\mathbf{C}}_2 \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \end{cases}} \tag{D.56}$$

The state equation equivalently reads as follows:

$$\begin{bmatrix} \dot{\underline{x}}_s \\ \dot{\underline{x}}_f \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{A}}_s & \mathbf{0} \\ \mathbf{0} & \widetilde{\mathbf{A}}_f \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} + \begin{bmatrix} \widetilde{\mathbf{B}}_s \\ \widetilde{\mathbf{B}}_f \end{bmatrix} \underline{u} \tag{D.57}$$

where:

$$\begin{cases} \widetilde{\mathbf{A}}_s = \widetilde{\mathbf{A}}_{11} - \widetilde{\mathbf{A}}_{12}\mathbf{L}_0 \\ \widetilde{\mathbf{A}}_f = \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon} \end{cases} \quad \text{and} \quad \begin{cases} \widetilde{\mathbf{B}}_s = \widetilde{\mathbf{B}}_1 - \mathbf{M}_0\frac{\widetilde{\mathbf{B}}_2}{\epsilon} \\ \widetilde{\mathbf{B}}_f = \frac{\widetilde{\mathbf{B}}_2}{\epsilon} \end{cases} \tag{D.58}$$

State vector $\underline{x}_s$ is related to the *slow* variables of the system whereas state vector $\underline{x}_f$ is related to the *fast* variables of the system and the two-time scale property (D.48) can be rewritten as follows:

$$\begin{cases} \lambda\left(s\mathbb{I} - \widetilde{\mathbf{A}}_s\right) \approx \lambda\left(\chi_{A_s}(s)\right) \\ \lambda\left(s\mathbb{I} - \widetilde{\mathbf{A}}_f\right) \approx \lambda\left(\chi_{A_f}(s)\right) \end{cases} \tag{D.59}$$

The two-stage design method for eigenvalue placement presented in Section D.2 can also be applied. Matrix $\widetilde{\mathbf{M}}$ defined in (D.27) is now approximated as follows:

$$\widetilde{\mathbf{M}} \approx -\widetilde{\mathbf{B}}_s\widetilde{\mathbf{K}}_f\left(\widetilde{\mathbf{A}}_f - \widetilde{\mathbf{B}}_f\widetilde{\mathbf{K}}_f\right)^{-1} \tag{D.60}$$

[4] Fu-Chuang Chen & Hassan K. Khalil, Two-Time-Scale Longitudinal Control Of Airplanes Using Singular Perturbation, Journal of Guidance Control and Dynamics, 1990, vol. 13, pp. 952-960

Finally, the state feedback gain which achieves approximate eigenvalue placement of of $\widetilde{\mathbf{A}} - \widetilde{\mathbf{B}}\widetilde{\mathbf{K}}$ is obtained from (D.33) as follows:

$$\begin{aligned}\widetilde{\mathbf{K}} &= \begin{bmatrix} \widetilde{\mathbf{K}}_s & \widetilde{\mathbf{K}}_f - \widetilde{\mathbf{K}}_s\widetilde{\mathbf{M}} \end{bmatrix} \mathbf{P}^{-1} \\ &= \begin{bmatrix} \widetilde{\mathbf{K}}_s & \widetilde{\mathbf{K}}_f - \widetilde{\mathbf{K}}_s\widetilde{\mathbf{M}} \end{bmatrix} \begin{bmatrix} \mathbb{I} & -\mathbf{M}_0 \\ \mathbf{0} & \mathbb{I} \end{bmatrix} \\ &= \begin{bmatrix} \widetilde{\mathbf{K}}_s & \widetilde{\mathbf{K}}_f - \widetilde{\mathbf{K}}_s\left(\widetilde{\mathbf{M}} + \mathbf{M}_0\right) \end{bmatrix}\end{aligned} \tag{D.61}$$

The state feedback gain $\mathbf{K}$ corresponding to state vector $\underline{x}$ is then obtained as follows:

$$\mathbf{K} = \widetilde{\mathbf{K}}\,\mathbf{T} \tag{D.62}$$

Example D.2. *We consider example D.1 where we can check that the two-time scale property* (D.48) *is satisfied with the following permutation matrix* $\mathbf{T}$*:*

$$\mathbf{T} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{D.63}$$

Matrix $\mathbf{T}$ *indicates that the third and the first components of the state vector are the* slow *variables, that are pitch angle* θ *and true airspeed* V_p*, whereas the fourth and the second components of the state vector are the* fast *variables, that are pitch rate* q *and angle of attack* α*.*

We get:

$$\widetilde{\mathbf{A}} = \mathbf{T}\mathbf{A}\mathbf{T}^T = \left[\begin{array}{cc|cc} 0.000 & 0.000 & 1.000 & 0.000 \\ -9.806 & -1.490\times10^{-2} & -6.800\times10^{-2} & 5.865 \\ \hline 0.000 & 0.000 & -2.251 & -4.980 \\ 0.000 & -3.000\times10^{-4} & 9.725\times10^{-1} & -1.586 \end{array}\right] \tag{D.64}$$

And thus:

$$\begin{cases} \widetilde{\mathbf{A}}_s = \widetilde{\mathbf{A}}_{11} - \widetilde{\mathbf{A}}_{12}\mathbf{L}_0 = \begin{bmatrix} 0.000 & 1.776\times10^{-4} \\ -9.806 & -1.538\times10^{-2} \end{bmatrix} \\ \widetilde{\mathbf{A}}_f = \frac{\widetilde{\mathbf{A}}_{22}}{\epsilon} = \begin{bmatrix} -2.251 & -4.980 \\ 9.725\times10^{-1} & -1.586 \end{bmatrix} \\ \widetilde{\mathbf{B}}_s = \widetilde{\mathbf{B}}_1 - \mathbf{M}_0\frac{\widetilde{\mathbf{B}}_2}{\epsilon} = \begin{bmatrix} -4.286 \\ -1.690\times10^{1} \end{bmatrix} \\ \widetilde{\mathbf{B}}_f = \frac{\widetilde{\mathbf{B}}_2}{\epsilon} = \begin{bmatrix} -2.364\times10^{1} \\ -2.886\times10^{-1} \end{bmatrix} \end{cases} \tag{D.65}$$

We check that the two-time scale property (D.48) *is satisfied. Indeed:*

$$\lambda\left(\widetilde{\mathbf{A}}\right) = \left\{\ -1.919 \pm 2.176j\,,\quad -7.293\times10^{-3} \pm 4.108\times10^{-2}j\ \right\} \tag{D.66}$$

and:

$$\begin{aligned}\lambda\left(\frac{\widetilde{\mathbf{A}}_{22}}{\epsilon}\right) \cup \lambda\left(\widetilde{\mathbf{A}}_0\right) &= \left\{\ -1.919 \pm 2.175j\ \right\} \\ &\qquad \cup \left\{\ -7.691\times10^{-3} \pm 4.101\times10^{-2}j\ \right\} \\ &\approx \lambda\left(\widetilde{\mathbf{A}}\right)\end{aligned} \tag{D.67}$$

Then state feedback gains $\widetilde{\mathbf{K}}_f$ *and* $\widetilde{\mathbf{K}}_s$ *are then set as follows:*

– *State feedback gain* $\widetilde{\mathbf{K}}_f$ *is set such that the eigenvalues of* $\widetilde{\mathbf{A}}_f - \widetilde{\mathbf{B}}_f\widetilde{\mathbf{K}}_f$ *are equal to* $\{\lambda_{clf}, \bar{\lambda}_{clf}\}$. *This leads to the following value of* $\widetilde{\mathbf{K}}_f$*:*

$$\widetilde{\mathbf{K}}_f = \left[\begin{array}{cc} 7.584 \times 10^{-2} & 1.556 \times 10^{-1} \end{array}\right] \tag{D.68}$$

– *Matrix* $\widetilde{\mathbf{M}}$ *is then approximated by* (D.60). *We get:*

$$\widetilde{\mathbf{M}} \approx -\widetilde{\mathbf{B}}_s\widetilde{\mathbf{K}}_f\left(\widetilde{\mathbf{A}}_f - \widetilde{\mathbf{B}}_f\widetilde{\mathbf{K}}_f\right)^{-1} = \left[\begin{array}{cc} -5.822 \times 10^{-1} & 5.836 \times 10^{-2} \\ -2.295 & 2.301 \times 10^{-1} \end{array}\right] \tag{D.69}$$

– *In the second stage, state feedback gain* $\widetilde{\mathbf{K}}_s$ *is set such that the eigenvalues of* $\widetilde{\mathbf{A}}_s - \left(\widetilde{\mathbf{B}}_s - \widetilde{\mathbf{M}}\widetilde{\mathbf{B}}_f\right)\widetilde{\mathbf{K}}_s$ *are equal to* $\{\lambda_{cls}, \bar{\lambda}_{cls}\}$. *This leads to the following value of* $\widetilde{\mathbf{K}}_s$*:*

$$\widetilde{\mathbf{K}}_s = \left[\begin{array}{cc} -2.203 \times 10^{-4} & -9.077 \times 10^{-6} \end{array}\right] \tag{D.70}$$

State feedback gain $\widetilde{\mathbf{K}}$ *for the actual system is finally obtained according to* (D.61)*:*

$$\begin{aligned} \widetilde{\mathbf{K}} &= \left[\begin{array}{cc} \widetilde{\mathbf{K}}_s & \widetilde{\mathbf{K}}_f - \widetilde{\mathbf{K}}_s\left(\widetilde{\mathbf{M}} + \mathbf{M}_0\right) \end{array}\right] \\ &= \left[\begin{array}{cccc} -2.203 \times 10^{-4} & -9.077 \times 10^{-6} & 7.564 \times 10^{-2} & 1.558 \times 10^{-1} \end{array}\right] \end{aligned} \tag{D.71}$$

We check that the eigenvalues of $\widetilde{\mathbf{A}} - \widetilde{\mathbf{B}}\widetilde{\mathbf{K}}$ *of the whole system are close to the expected eigenvalues* $\{\lambda_{clf}, \bar{\lambda}_{clf}, \lambda_{cls}, \bar{\lambda}_{cls}\}$*:*

$$\lambda\left(\widetilde{\mathbf{A}} - \widetilde{\mathbf{B}}\widetilde{\mathbf{K}}\right) = \left\{\ -1.000 \pm 1.001j\,,\quad -9.277 \times 10^{-3} \pm 1.675 \times 10^{-2}j\ \right\} \tag{D.72}$$

The state feedback gain $\mathbf{K}$ *corresponding to state vector* $\underline{x}$ *is then obtained as follows:*

$$\begin{aligned} \mathbf{K} &= \widetilde{\mathbf{K}}\,\mathbf{T} \\ &= \left[\begin{array}{cccc} -9.077 \times 10^{-6} & 1.558 \times 10^{-1} & -2.203 \times 10^{-4} & 7.564 \times 10^{-2} \end{array}\right] \end{aligned} \tag{D.73}$$

This state feedback gain has to be compared with state feedback gain obtained in (D.43) *to achieve exact eigenvalue placement: the exact eigenvalue placement is achieved by manipulating a state vector of dimension* 4 *whereas the approximate eigenvalue placement is achieved by manipulating* 2 *state vectors of dimension* 2*, which leads to simpler computations.*

■

D.3.2 Nonlinear system

Nonlinear singularly perturbed systems, also known as two-timescale systems, are described by ordinary differential equation of the form:

$$\left\{\begin{array}{l} \dot{\underline{x}}_s = f(\underline{x}_s, \underline{x}_f, \epsilon) \\ \epsilon\,\dot{\underline{x}}_f = g(\underline{x}_s, \underline{x}_f, \epsilon) \end{array}\right. \tag{D.74}$$

where $\underline{x}_s \in \mathbb{R}^{n_s}$ are the slow states, $\underline{x}_f \in \mathbb{R}^{n_f}$ the fast states, $0 < \epsilon \ll 1$ is a small parameter responsible for the timescale separation between slow and fast dynamics, and f and g are smooth vector fields. For $\epsilon > 0$ we can define a new time parameter $\tau = t/\epsilon$ and obtain an equivalent system to (D.74) of the form:

$$\tau = \frac{t}{\epsilon} \Rightarrow \left\{ \begin{array}{l} \frac{d}{d\tau}\underline{x}_s = \epsilon\, f(\underline{x}_s, \underline{x}_f, \epsilon) \\ \frac{d}{d\tau}\underline{x}_f = g(\underline{x}_s, \underline{x}_f, \epsilon) \end{array} \right. \tag{D.75}$$

In the study of slow-fast systems, two reduced subsystems are defined by taking the limit $\epsilon \to 0$ of (D.74) and (D.75). By doing so we obtain the DAE (Differential Algebraic Equation) and the Layer Equation, which read[5]:

$$\text{DAE: } \left\{ \begin{array}{l} \dot{\underline{x}}_s = f(\underline{x}_s, \underline{x}_f, 0) \\ \underline{0} = g(\underline{x}_s, \underline{x}_f, 0) \end{array} \right. \qquad \text{Layer Equation: } \left\{ \begin{array}{l} \frac{d}{d\tau}\underline{x}_s = \underline{0} \\ \frac{d}{d\tau}\underline{x}_f = g(\underline{x}_s, \underline{x}_f, 0) \end{array} \right. \tag{D.76}$$

The critical manifold $\mathcal{M}$ associated to the two-timescale systems (D.74) is defined as:

$$\mathcal{M} = \left\{ \left(\underline{x}_{se}, \underline{x}_{fe}\right) \in \mathbb{R}^{n_s+n_f} \mid g(\underline{x}_{se}, \underline{x}_{fe}, 0) = \underline{0} \right\} \tag{D.77}$$

Note that $\mathcal{M}$ is the phase-space of the DAE and the set of equilibrium points of the Layer Equation[5]. We say that $\mathcal{M}$ is Normally Hyperbolic if every point in $\mathcal{M}$ is a hyperbolic equilibrium point of the reduced dynamics $\frac{d}{d\tau}\underline{x}_f = g(\underline{x}_s, \underline{x}_f, 0)$.

D.4 Two-frequency-scale transfer function

We consider hereafter the following block-diagonal state space representation (D.14) of a system:

$$\left\{ \begin{array}{l} \begin{bmatrix} \dot{\underline{x}}_s \\ \dot{\underline{x}}_f \end{bmatrix} = \begin{bmatrix} \mathbf{A}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_f \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} + \begin{bmatrix} \mathbf{B}_s \\ \mathbf{B}_f \end{bmatrix} \underline{u} \\ \underline{y} = \begin{bmatrix} \mathbf{C}_s & \mathbf{C}_f \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \end{array} \right. \tag{D.78}$$

From this block-diagonal form we get the following transfer function of the system as follows:

$$\begin{aligned} \mathbf{F}(s) &= \begin{bmatrix} \mathbf{C}_s & \mathbf{C}_f \end{bmatrix} \left(s\mathbb{I} - \begin{bmatrix} \mathbf{A}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_f \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{B}_s \\ \mathbf{B}_f \end{bmatrix} \\ &= \mathbf{C}_s \left(s\mathbb{I} - \mathbf{A}_s\right)^{-1} \mathbf{B}_s + \mathbf{C}_f \left(s\mathbb{I} - \mathbf{A}_f\right)^{-1} \mathbf{B}_f \end{aligned} \tag{D.79}$$

In the preceding relations we assume that the system is stable; in others words all the eigenvalues of matrices $\mathbf{A}_s$ and $\mathbf{A}_f$ have negative real part.

Furthermore, we will assume that matrix $\mathbf{A}_s$ contains the *slow* modes of the system (that are the eigenvalues which are the closest to zero) whereas matrix $\mathbf{A}_f$ contains the *fast* modes of the system (that are the eigenvalues which are

[5] Hildeberto Jardon-Kojakhmetov & Jacquelien M.A. Scherpen, Stabilization of slow-fast control systems: the non-hyperbolic case, https://doi.org/10.48550/arXiv.1710.01629

the most distant to zero). From a practical point of view, and assuming that the real part of the eigenvalues λ_i of $\mathbf{A}$ are sorted in a descending manner, the value of n_s which delimits the border between the *slow* and the *fast* modes can be obtained by finding the minimum of $\frac{|\lambda_{n_s}|}{|\lambda_{n_s+1}|}$.

The *slow* subsystem is obtained by setting $\dot{\underline{x}}_f = \underline{0}$ in (D.14). Physically, it means that the *fast* components of the state vector have achieved the equilibrium point well before the *slow* components of the state vector. We get from (D.14):

$$\begin{cases} \dot{\underline{x}}_s = \mathbf{A}_s\,\underline{x}_s + \mathbf{B}_s\,\underline{u} \\ \dot{\underline{x}}_f := \underline{0} = \mathbf{A}_f\,\underline{x}_f + \mathbf{B}_f\,\underline{u} \\ \underline{y} = \mathbf{C}_s\,\underline{x}_s + \mathbf{C}_f\,\underline{x}_f \end{cases} \tag{D.80}$$

Assuming that $\mathbf{A}_f^{-1}$ exists, we get $\underline{x}_f = -\mathbf{A}_f^{-1}\mathbf{B}_f\,\underline{u}$. Thus the preceding relations reduce as follows:

$$\underline{x}_f = -\mathbf{A}_f^{-1}\mathbf{B}_f\,\underline{u} \Rightarrow \begin{cases} \dot{\underline{x}}_s = \mathbf{A}_s\,\underline{x}_s + \mathbf{B}_s\,\underline{u} \\ \underline{y} = \mathbf{C}_s\,\underline{x}_s - \mathbf{C}_f\mathbf{A}_f^{-1}\mathbf{B}_f\,\underline{u} \end{cases} \tag{D.81}$$

The transfer function of the *slow* dynamics $\mathbf{F}_s(s)$ is then obtained by taking the Laplace transform of the preceding relations, assuming no initial condition. We get:

$$\boxed{\mathbf{F}_s(s) = \mathbf{C}_s\,(s\mathbb{I} - \mathbf{A}_s)^{-1}\,\mathbf{B}_s - \mathbf{C}_f\mathbf{A}_f^{-1}\mathbf{B}_f} \tag{D.82}$$

The *fast* subsystem is obtained by setting $\underline{x}_s = \underline{0}$ in (D.14). Physically, it means that the *slow* components of the state vector stay at the equilibrium point while the *fast* components of the state vector are changing. We get from (D.14):

$$\begin{cases} \dot{\underline{x}}_f = \mathbf{A}_f\,\underline{x}_f + \mathbf{B}_f\,\underline{u} \\ \dot{\underline{x}}_s := \underline{0} \Rightarrow \underline{y} = \mathbf{C}_s\,\underline{x}_s + \mathbf{C}_f\,\underline{x}_f := \mathbf{C}_f\,\underline{x}_f \end{cases} \tag{D.83}$$

The transfer function of the *fast* dynamics $\mathbf{F}_f(s)$ is then obtained by taking the Laplace transform of the preceding relations, assuming no initial condition. We get:

$$\boxed{\mathbf{F}_f(s) = \mathbf{C}_f\,(s\mathbb{I} - \mathbf{A}_f)^{-1}\,\mathbf{B}_f} \tag{D.84}$$

The so-called *fast* outputs are the outputs for which the Bode magnitude plot of $\mathbf{F}_f(s)$ and $\mathbf{F}(s)$ match for *high* frequencies. In the time domain, the impulse response of $\mathbf{F}_f(s)$ and $\mathbf{F}(s)$ match on the *fast* scale time.

Furthermore it can be noticed that the Bode magnitude and phase plots of $\mathbf{F}_s(s)$ and $\mathbf{F}(s)$ match pretty well for *low* frequencies, both for *fast* and for *slow* outputs.

Finally the following property holds:

$$\lim_{s\to\infty} (s\mathbb{I} - \mathbf{A}_s)^{-1} = \mathbf{0} \Rightarrow \mathbf{F}_s(\infty) = \mathbf{F}_f(0) = -\mathbf{C}_f\mathbf{A}_f^{-1}\mathbf{B}_f \tag{D.85}$$

Example D.3. *We use again example D.1. Figure D.2 shows the Bode magnitude plot and the impulse response of the* fast *outputs, that are α and q: it can be seen that the Bode magnitude plot of $\mathbf{F}_f(s)$ and $\mathbf{F}(s)$ match for* high

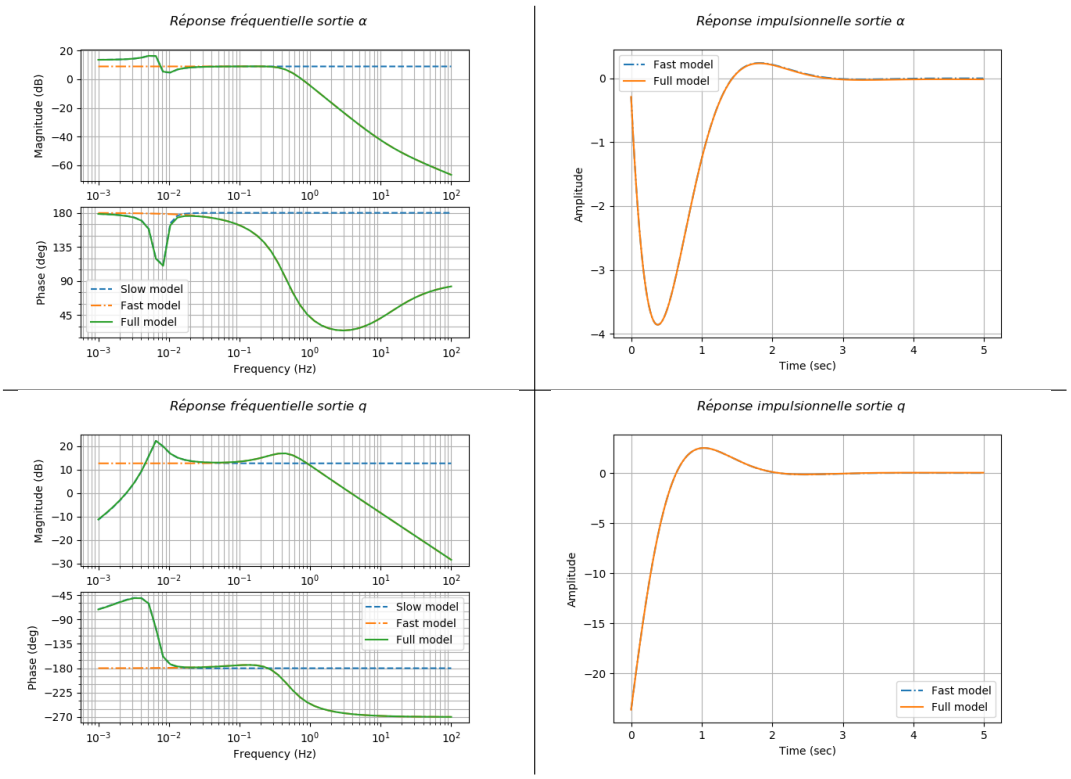

Figure D.2: Bode magnitude plot and impulse response of *fast* outputs

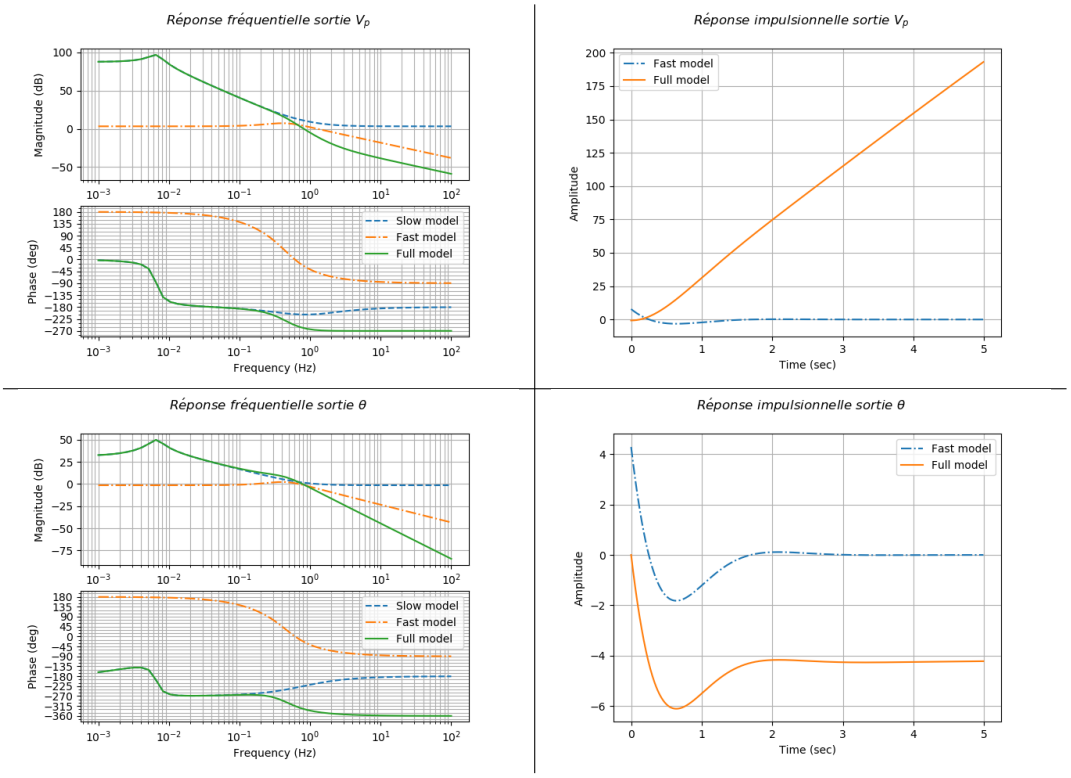

Figure D.3: Bode magnitude plot and impulse response of *slow* outputs

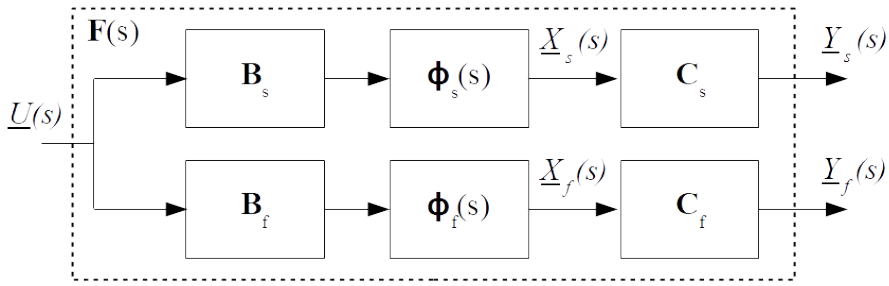

Figure D.4: Block diagonal system in the Laplace domain

frequencies. In the time domain, the impulse response of $\mathbf{F}_f(s)$ *and* $\mathbf{F}(s)$ *match on the* fast *scale time (here* 5 *seconds) .*

On the other hand, Figure D.3 shows the Bode magnitude plot and the impulse response of the slow *outputs, that are* V_p *and* θ*: contrary to the* fast *outputs, the* high *frequencies Bode magnitude plot of* $\mathbf{F}_f(s)$ *and* $\mathbf{F}(s)$ *do not match for* high *frequencies. The mismatch is also clear between the impulse response of* $\mathbf{F}_f(s)$ *and* $\mathbf{F}(s)$ *on the* fast *scale time (here* 5 *seconds) .*

Furthermore it can be noticed that the Bode magnitude and phase plots of $\mathbf{F}_s(s)$ *and* $\mathbf{F}(s)$ *match pretty well for* low *frequencies, both for* fast *and for* slow *outputs.*

■

D.5 Hierarchical output feedback of singularly perturbed system

We consider hereafter the following block-diagonal state space representation (D.14) of a system:

$$\begin{cases} \begin{bmatrix} \dot{\underline{x}}_s \\ \dot{\underline{x}}_f \end{bmatrix} = \begin{bmatrix} \mathbf{A}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_f \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} + \begin{bmatrix} \mathbf{B}_s \\ \mathbf{B}_f \end{bmatrix} \underline{u} \\ \begin{bmatrix} \underline{y}_s \\ \underline{y}_f \end{bmatrix} \begin{bmatrix} \mathbf{C}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_f \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \end{cases} \tag{D.86}$$

where:

- $\underline{x}_s(t)$ represents the *slow* components of the state vector;
- $\underline{x}_f(t)$ represents *fast* components of the state vector;
- $\underline{y}_s(t)$ is a *slow* output vector;
- $\underline{y}_f(t)$ is a *fast* output vector.

Let $\mathbf{F}(s)$ be the transfer function of the plant:

$$\begin{bmatrix} \underline{Y}_s(s) \\ \underline{Y}_f(s) \end{bmatrix} = \mathbf{F}(s)\underline{U}(s) \tag{D.87}$$

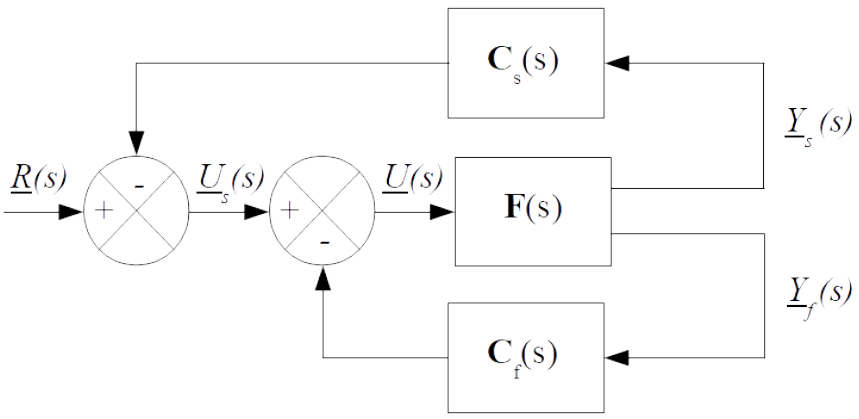

Figure D.5: Hierarchical output feedback of a singularly perturbed system

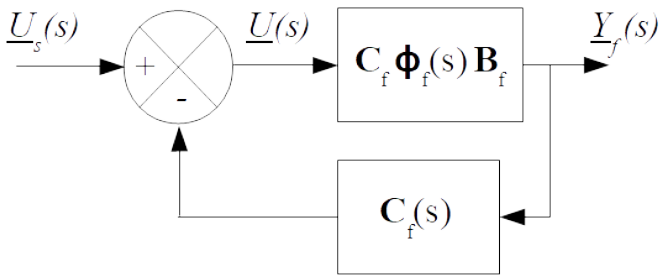

Figure D.6: Control of the *fast* subsystem

The transfer function $\mathbf{F}(s)$ expands as follows:

$$\begin{aligned}\mathbf{F}(s) &= \begin{bmatrix} \mathbf{C}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_f \end{bmatrix} \left(s\mathbb{I} - \begin{bmatrix} \mathbf{A}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_f \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{B}_s \\ \mathbf{B}_f \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{C}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_f \end{bmatrix} \begin{bmatrix} (s\mathbb{I} - \mathbf{A}_s)^{-1} & \mathbf{0} \\ \mathbf{0} & (s\mathbb{I} - \mathbf{A}_f)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{B}_s \\ \mathbf{B}_f \end{bmatrix}\end{aligned} \tag{D.88}$$

We finally get:

$$\begin{bmatrix} \underline{Y}_s(s) \\ \underline{Y}_f(s) \end{bmatrix} = \begin{bmatrix} \mathbf{C}_s\, \mathbf{\Phi}_s(s)\, \mathbf{B}_s \\ \mathbf{C}_f\, \mathbf{\Phi}_f(s)\, \mathbf{B}_f \end{bmatrix} \underline{U}(s) \tag{D.89}$$

where:

$$\begin{cases} \mathbf{\Phi}_s(s) = (s\mathbb{I} - \mathbf{A}_s)^{-1} \\ \mathbf{\Phi}_f(s) = (s\mathbb{I} - \mathbf{A}_f)^{-1} \end{cases} \tag{D.90}$$

The corresponding block diagram representation is shown in Figure D.4 where transfer function $\mathbf{C}_f\, \mathbf{\Phi}_f(s)\, \mathbf{B}_f$ corresponds to the *fast* subsystem whereas transfer function $\mathbf{C}_s\, \mathbf{\Phi}_s(s)\, \mathbf{B}_s$ corresponds to the *slow* subsystem.

The hierarchical output feedback of a singularly perturbed system is shown in Figure D.5 and uses results presented in Section D.4. It consists in splitting the control into two control loops:

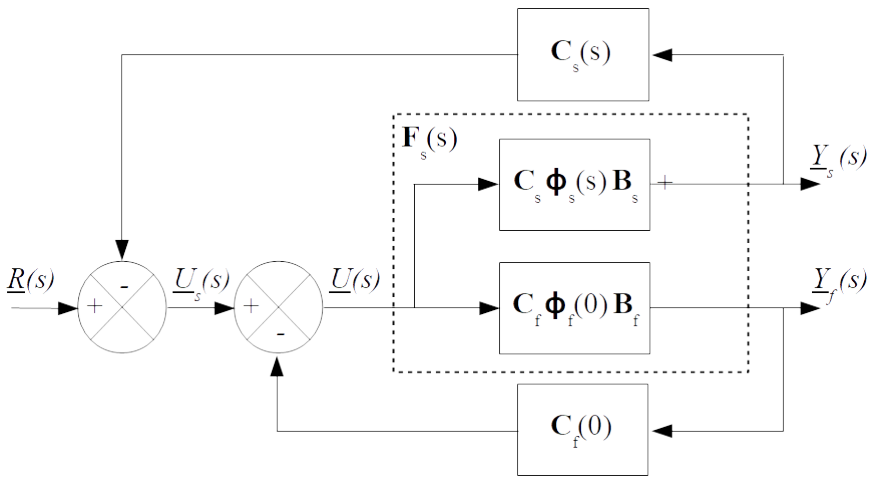

Figure D.7: Control of the *slow* subsystem

- In the first stage, the *fast* subsystem is first controlled through the *fast* controller $C_f(s)$ as shown in Figure D.6.

- In the second stage, the *slow* subsystem is then controlled through the *slow* controller $C_s(s)$ as shown in Figure D.7. The *fast* controller $C_f(s)$ is replaced by its low frequency approximation $C_f(0)$ as well as transfer function $\mathbf{C}_f\,\mathbf{\Phi}_f(s)\,\mathbf{B}_f$, which is replaced by its low frequency approximation $\mathbf{C}_f\,\mathbf{\Phi}_f(0)\,\mathbf{B}_f$. Following Figure D.7, the *open loop* transfer function between $Y_s(s)$ and $U_s(s)$ reads:

$$\begin{aligned} U(s) &= U_s(s) - C_f(s)\mathbf{C}_f\,\mathbf{\Phi}_f(s)\,\mathbf{B}_f\,U(s) \\ \Rightarrow U(s) &= (\mathbb{I} + C_f(s)\mathbf{C}_f\,\mathbf{\Phi}_f(s)\,\mathbf{B}_f)^{-1}\,U_s(s) \\ &\approx (\mathbb{I} + C_f(0)\mathbf{C}_f\,\mathbf{\Phi}_f(0)\,\mathbf{B}_f)^{-1}\,U_s(s) \\ \Rightarrow Y_s(s) &= \mathbf{C}_s\,\mathbf{\Phi}_s(s)\,\mathbf{B}_s\,U(s) \\ &\approx \mathbf{C}_s\,\mathbf{\Phi}_s(s)\,\mathbf{B}_s\,(\mathbb{I} + C_f(0)\mathbf{C}_f\,\mathbf{\Phi}_f(0)\,\mathbf{B}_f)^{-1}\,U_s(s) \end{aligned} \tag{D.91}$$

Thus, in the SISO case all is working as if transfer function $\mathbf{C}_s\,\mathbf{\Phi}_s(s)\,\mathbf{B}_s$ is feedback with static gain $C_f(0)\mathbf{C}_f\,\mathbf{\Phi}_f(0)\,\mathbf{B}_f$, as shown in Figure D.8

A similar approach can be applied when the *fast* controller $C_f(s)$ and the *slow* controller $C_s(s)$ are placed in the direct path rather than the feedback path.

D.6 Weierstrass decomposition and descriptor system

If $\det(s\mathbf{E} - \mathbf{A}) = c\prod_{i=1}^{r_E}(s - \lambda_i)$, where $c \neq 0$ is a constant and where $\mathbf{E}$ and $\mathbf{A}$ are $n \times n$ square matrices, and if $\lambda \in \mathbb{C}$ is such that $\lambda\mathbf{E} - \mathbf{A}$ is non-singular,

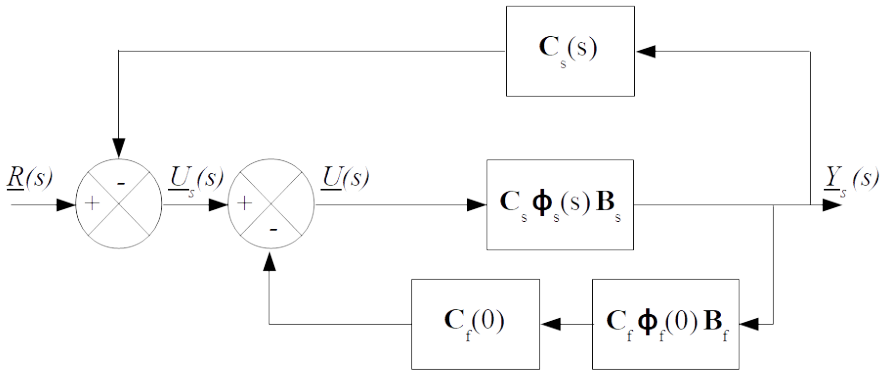

Figure D.8: Equivalent control of the *slow* subsystem

than it can be shown that[6]:

$$\begin{aligned} \det\left(s\mathbf{E}-\mathbf{A}\right) &= c\prod_{i=1}^{r_E}\left(s-\lambda_i\right) \\ \Rightarrow \det\left(s\mathbb{I}-\left(\lambda\mathbf{E}-\mathbf{A}\right)^{-1}\mathbf{E}\right) &= s^{n-r_E}\prod_{i=1}^{r_E}\left(s-\tfrac{1}{\lambda-\lambda_i}\right) \end{aligned} \tag{D.92}$$

The following relation also holds:

$$\left(\lambda\mathbf{E}-\mathbf{A}\right)^{-1}\mathbf{A} = \lambda\left(\lambda\mathbf{E}-\mathbf{A}\right)^{-1}\mathbf{E}-\mathbb{I} \tag{D.93}$$

Furthermore let $n_s \leq r_E$. Then it can be shown that range $(\mathbf{S})$+range $(\mathbf{F}) = \mathbb{R}^n$ and range $(\mathbf{ES})$ + range $(\mathbf{AF}) = \mathbb{R}^n$, where:

$$\begin{cases} \mathbf{S} &= \ker\left(\prod_{i=1}^{n_s}\left(\left(\lambda\mathbf{E}-\mathbf{A}\right)^{-1}\mathbf{E}-\frac{1}{\lambda-\lambda_i}\mathbb{I}\right)\right) \\ \mathbf{F} &= \ker\left(\left(\left(\lambda\mathbf{E}-\mathbf{A}\right)^{-1}\mathbf{E}\right)^{n-r_E}\prod_{i=n_s+1}^{r_E}\left(\left(\lambda\mathbf{E}-\mathbf{A}\right)^{-1}\mathbf{E}-\frac{1}{\lambda-\lambda_i}\mathbb{I}\right)\right) \end{cases} \tag{D.94}$$

Note that if matrices $\mathbf{E}$ and $\mathbf{A}$ are real-valued, the preceding relations reduce as follows where $Re\left(\mathbf{X}\right)$ designates the real part of matrix $\mathbf{X}$:

$$\begin{cases} \mathbf{S} &= \ker\left(Re\left(\prod_{i=1}^{n_s}\left(\left(\lambda\mathbf{E}-\mathbf{A}\right)^{-1}\mathbf{E}-\frac{1}{\lambda-\lambda_i}\mathbb{I}\right)\right)\right) \\ \mathbf{F} &= \ker\left(Re\left(\left(\left(\lambda\mathbf{E}-\mathbf{A}\right)^{-1}\mathbf{E}\right)^{n-r_E}\prod_{i=n_s+1}^{r_E}\left(\left(\lambda\mathbf{E}-\mathbf{A}\right)^{-1}\mathbf{E}-\frac{1}{\lambda-\lambda_i}\mathbb{I}\right)\right)\right) \end{cases} \tag{D.95}$$

Thus $\forall\ \underline{x} \in \mathbb{R}^n$, we can write:

$$\boxed{\underline{x} := \mathbf{S}\underline{x}_s + \mathbf{F}\underline{x}_f = \begin{bmatrix} \mathbf{S} & \mathbf{F} \end{bmatrix}\begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \ \forall\ \underline{x} \in \mathbb{R}^n} \tag{D.96}$$

Let $\mathbf{J}_s$ and $\mathbf{J}_f$ be the restriction of $\left(\lambda\mathbf{E}-\mathbf{A}\right)^{-1}\mathbf{E}$ to $\mathbf{S}$ and $\mathbf{F}$, respectively:

$$\begin{cases} \mathbf{J}_s := \mathbf{S}^T\left(\lambda\mathbf{E}-\mathbf{A}\right)^{-1}\mathbf{E}\,\mathbf{S} \\ \mathbf{J}_f := \mathbf{F}^T\left(\lambda\mathbf{E}-\mathbf{A}\right)^{-1}\mathbf{E}\,\mathbf{F} \end{cases} \tag{D.97}$$

[6]Cobb, J.D. Global analyticity of a geometric decomposition for linear singularly perturbed systems. Circuits Systems and Signal Process 5, 139–152 (1986). https://doi.org/10.1007/BF01600192

Then the following relations hold:

$$\begin{cases} \det\left(s\mathbb{I} - \mathbf{J}_s\right) = \prod_{i=1}^{n_s}\left(s - \frac{1}{\lambda-\lambda_i}\right) \\ \det\left(s\mathbb{I} - \mathbf{J}_f\right) = \prod_{i=n_s+1}^{r_E}\left(s - \frac{1}{\lambda-\lambda_i}\right) \end{cases} \tag{D.98}$$

Now consider the following descriptor system:

$$\begin{cases} \mathbf{E}\dot{\underline{x}} = \mathbf{A}\underline{x} + \mathbf{B}\underline{u} \\ \underline{y} = \mathbf{C}\,\underline{x} \end{cases} \tag{D.99}$$

From the Weierstrass decomposition theorem, it can be shown that there exists an invertible matrix $\mathbf{M}$ such that this system has the following canonical form[7]:

$$\begin{cases} \underline{x} := \begin{bmatrix} \mathbf{S} & \mathbf{F} \end{bmatrix} \begin{bmatrix} \underline{x}_s \\ \underline{x}_f \end{bmatrix} \\ \mathbf{ME}\begin{bmatrix} \mathbf{S} & \mathbf{F} \end{bmatrix} = \begin{bmatrix} \mathbb{I}_{n_s} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_f \end{bmatrix} \\ \mathbf{MA}\begin{bmatrix} \mathbf{S} & \mathbf{F} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_s & \mathbf{0} \\ \mathbf{0} & \mathbb{I}_{n-n_s} \end{bmatrix} \\ \mathbf{MB} = \begin{bmatrix} \mathbf{B}_s \\ \mathbf{B}_f \end{bmatrix} \\ \mathbf{C}\begin{bmatrix} \mathbf{S} & \mathbf{F} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_s & \mathbf{C}_f \end{bmatrix} \end{cases} \Leftrightarrow \begin{cases} \dot{\underline{x}}_s &= \mathbf{A}_s\underline{x}_s + \mathbf{B}_s\underline{u} \\ \mathbf{A}_f\dot{\underline{x}}_f &= \underline{x}_f + \mathbf{B}_f\underline{u} \\ \underline{y} &= \mathbf{C}_s\underline{x}_s + \mathbf{C}_f\underline{x}_f \end{cases} \tag{D.100}$$

where :

$$\boxed{\mathbf{M} = \begin{bmatrix} \mathbf{ES} & \mathbf{AF} \end{bmatrix}^{-1}} \tag{D.101}$$

Furthermore the following relations hold:

$$\begin{cases} \det\left(s\mathbb{I}_{n_s} - \mathbf{A}_s\right) = \prod_{i=1}^{n_s}\left(s - \lambda_i\right) \\ \det\left(s\mathbf{A}_f - \mathbb{I}_{n-n_s}\right) = \det\left(\mathbf{M}\right)\det\left(\begin{bmatrix} \mathbf{S} & \mathbf{F} \end{bmatrix}\right)\prod_{i=n_s+1}^{r_E}\left(s - \lambda_i\right) \\ \det\left(\mathbf{M}\right)\det\left(\begin{bmatrix} \mathbf{S} & \mathbf{F} \end{bmatrix}\right)\det\left(s\mathbf{E} - \mathbf{A}\right) = \det\left(s\mathbb{I}_{n_s} - \mathbf{A}_s\right)\det\left(s\mathbf{A}_f - \mathbb{I}_{n-n_s}\right) \end{cases} \tag{D.102}$$

Matrix $\mathbf{A}_f$ can alternatively be obtained as:

$$\mathbf{A}_f = \left(\lambda\mathbf{J}_f - \mathbb{I}\right)^{-1}\mathbf{J}_f \tag{D.103}$$

Moreover matrix $\mathbf{A}_f$ is nilpotent if $\det\left(s\mathbf{E} - \mathbf{A}\right)$ has degree n_s, that is when $r_E = n_s$. Assuming that $\mathbf{A}_f$ is nilpotent, matrix $\left(s\mathbf{E} - \mathbf{A}\right)^{-1}$ is a rational matrix with Laurent series expansion at infinity given by[8]:

$$\begin{aligned} \left(s\mathbf{E} - \mathbf{A}\right)^{-1} &= s^{-1}\sum_{i=-h}^{\infty}\mathbf{L}_i\, s^{-i} \\ &= \mathbf{L}_{-h}s^{h-1} + \cdots + \mathbf{L}_{-1}s^{0} + \mathbf{L}_0 s^{-1} + \mathbf{L}_1 s^{-2} + \cdots \end{aligned} \tag{D.104}$$

[7] Thomas Berger, On Differential-Algebraic Control Systems, 2013, PhD thesis

[8] L. Moysis, I. Kafetzis and N. P. Karampetakis, Reachability and controllability of discrete time descriptor systems using the Weierstrass decomposition, 2018 5th International Conference on Control, Decision and Information Technologies (CoDIT), 2018, pp. 379-384, doi: 10.1109/CoDIT.2018.8394863

where $h \leq \text{rank}(\mathbf{E}) - deg\left(\det\left(s\mathbf{E} - \mathbf{A}\right)\right) + 1$ is the index of nilpotency of the pair $(\mathbf{E}, \mathbf{A})$ and:

$$\mathbf{L}_i = \begin{cases} \begin{bmatrix} \mathbf{A}_s^i & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{n-n_s} \end{bmatrix} \mathbf{M} & \forall\, i \geq 0 \\ \begin{bmatrix} \mathbf{0}_{n_s} & \mathbf{0} \\ \mathbf{0} & -\mathbf{A}_f^{-i-1} \end{bmatrix} \mathbf{M} & i = -h, \cdots, -1 \end{cases} \tag{D.105}$$

If matrix $\mathbf{J}_s$ is built from the slowest eigenvalues of state matrix $\mathbf{A}$ and matrix $\mathbf{J}_f$ is built from the fastest eigenvalues of state matrix $\mathbf{A}$, then matrices $\mathbf{A}_s$ corresponds to the state matrix of the slowest modes of the system whereas matrix $\mathbf{A}_f$ encompasses its fastest modes. Matrix $\mathbf{A}_f$ can then be seen as a *small* valued matrix and can be written as $\mathbf{A}_f := \epsilon\, \mathbf{A}_{f0}$. When *small* parameter ϵ is set to zero, the following singularly perturbed system is obtained:

$$\begin{aligned} \mathbf{A}_f := \epsilon\, \mathbf{A}_{f0} \quad \Rightarrow & \begin{cases} \dot{\underline{x}}_s &= \mathbf{A}_s \underline{x}_s + \mathbf{B}_s \underline{u} \\ \epsilon\, \mathbf{A}_{f0}\, \dot{\underline{x}}_f &= \underline{x}_f + \mathbf{B}_f \underline{u} \\ \underline{y} &= \mathbf{C}_s \underline{x}_s + \mathbf{C}_f \underline{x}_f \end{cases} \\ \epsilon = 0 \quad \Rightarrow & \begin{cases} \dot{\underline{x}}_s &= \mathbf{A}_s \underline{x}_s + \mathbf{B}_s \underline{u} \\ \underline{0} &= \underline{x}_f + \mathbf{B}_f \underline{u} \\ \underline{y} &= \mathbf{C}_s \underline{x}_s + \mathbf{C}_f \underline{x}_f \end{cases} \end{aligned} \tag{D.106}$$

Example D.4. *We consider the linearized longitudinal dynamics at a trimmed flight condition of a jet liner presented in Example D.1. To get the descriptor form of* (D.99)*, we simply set* $\mathbf{E} := \mathbb{I}_4$*. Then applying* (D.95) *with* $\lambda = 1$*, and choosing* $n_s = 2$*, we get the following expression of matrices* $\mathbf{S}$*,* $\mathbf{F}$ *and* $\mathbf{M}$*:*

$$\begin{cases} \mathbf{S} = \begin{bmatrix} 0.9882059 & -0.1531309 \\ -0.0000781 & 0.0000203 \\ -0.1531309 & -0.9882056 \\ 0.0000547 & -0.0008144 \end{bmatrix} \\ \mathbf{F} = \begin{bmatrix} 0.1210859 & -0.6715563 \\ 0.4109715 & -0.5249282 \\ 0.0734018 & -0.378629 \\ 0.9005847 & 0.3606974 \end{bmatrix} \\ \mathbf{M} = \begin{bmatrix} 0.9880751 & -0.9410508 & -0.1533847 & 0.3090902 \\ -0.1530286 & 0.7444567 & -0.9880101 & -0.2386219 \\ 0.0000682 & 0.3446454 & 0.0001831 & -0.2264441 \\ 0.0000549 & 0.7795729 & -0.0000278 & 0.04285 \end{bmatrix} \end{cases} \tag{D.107}$$

Then:

$$\begin{cases} \mathbf{A}_s = \begin{bmatrix} 1.4688616 & 9.5785097 \\ -0.2276681 & -1.4834483 \end{bmatrix} \\ \mathbf{A}_f = \begin{bmatrix} -0.0622709 & -0.262707 \\ 0.3589769 & -0.3937903 \end{bmatrix} \end{cases} \tag{D.108}$$

We can check that:

$$\begin{cases} \det\left(s\mathbf{A}_f - \mathbb{I}\right) &= \det\left(\mathbf{M}\right) \det\left(\begin{bmatrix} \mathbf{S} & \mathbf{F} \end{bmatrix}\right) \left(s - \lambda_f\right)\left(s - \bar{\lambda}_f\right) \\ &= 1 + 0.4560613s + 0.1188274s^2 \\ \det\left(s\mathbb{I} - \mathbf{A}_s\right) &= \left(s - \lambda_s\right)\left(s - \bar{\lambda}_s\right) \\ &= 0.0017408 + 0.0145867s + s^2 \end{cases} \tag{D.109}$$

■

Appendix E

Carleman linearization

E.1 Reduced Kronecker product

Let $\underline{x} := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^T \in \mathbb{R}^n$. We will denote $\underline{x}^{[k]}$ the kth *reduced* Kronecker product of $\underline{x}$. Denoting by $\widetilde{\otimes}$ the *reduced* Kronecker product, vectors $\underline{x}^{[1]} \in \mathbb{R}^n$ and $\underline{x}^{[2]} \in \mathbb{R}^{n(n+1)/2}$ are defined as follows[1]:

$$\begin{aligned} \underline{x}^{[0]} &:= 1 \\ \underline{x}^{[1]} &:= \underline{x} \\ \underline{x}^{[2]} &:= \underline{x} \widetilde{\otimes} \underline{x} \\ &:= [\underbrace{x_1^2, x_1\, x_2, \cdots, x_1\, x_n}_{n\ terms}, \underbrace{x_2^2,\ x_2\, x_3, \cdots, x_2\, x_n}_{n-1\ terms}, \cdots, \underbrace{x_n^2}_{1\ term}]^T \end{aligned} \tag{E.1}$$

In this way we define:

$$\underline{x}^{[k]} := \underbrace{\underline{x} \widetilde{\otimes} \underline{x} \widetilde{\otimes} \cdots \widetilde{\otimes} \underline{x}}_{k\ times} \tag{E.2}$$

The components of $\underline{x}^{[k]}$ are the *unitary* monomials which appears in the development of polynomial $(x_1 + \cdots + x_n)^k$. The size of vector $\underline{x}^{[k]}$ is $\begin{pmatrix} n+k-1 \\ k \end{pmatrix} = \frac{(n+k-1)!}{k!\,(n-1)!}$.

Example E.1. *Assume that* $\underline{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$. *Then* $(x_1 + x_2)^3 = x_1^3 + 3x_1^2 x_2 + 3x_1 x_2^2 + x_2^3$. *Thus the* unitary *monomials which appears in the development of polynomial* $(x_1 + x_2)^3$ *are* $\left(x_1^3,\, x_1^2 x_2,\, x_1 x_2^2,\, x_2^3\right)$ *and consequently:*

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow \underline{x}^{[3]} = \begin{bmatrix} x_1^3 \\ x_1^2 x_2 \\ x_1 x_2^2 \\ x_2^3 \end{bmatrix} \tag{E.3}$$

[1] S. Irving and C. Joaquín, On stabilization of non linear systems by using Carleman linearization and periodic systems theory, 8th International Conference on Electrical Engineering, Computing Science and Automatic Control, 2011, pp. 1-6, doi: 10.1109/ICEEE.2011.6106597.

The size of vector $\underline{x}^{[3]}$ *is* $4 = \frac{(2+3-1)!}{3!\,(2-1)!}$.

■

Note that $(\mathbf{A}\,\underline{x})\,\widetilde{\otimes}\,(\mathbf{B}\,\underline{x}) = \left(\mathbf{A}\widetilde{\otimes}\mathbf{B}\right)\underline{x}^{[2]}$. In order to express the product $\mathbf{A}\widetilde{\otimes}\mathbf{B}$ in terms of $\mathbf{A}$ and $\mathbf{B}$, we will use the Kronecker product $\otimes$ (not the *reduced* one !). Let $\otimes$ be the Kronecker product between matrices $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{p\times q}$ and $\mathbf{B} \in \mathbb{R}^{r\times s}$. Then[2]:

$$\mathbf{A}\otimes\mathbf{B} := \begin{bmatrix} a_{11}\,\mathbf{B} & \cdots & a_{1q}\,\mathbf{B} \\ \vdots & & \vdots \\ a_{p1}\,\mathbf{B} & \cdots & a_{pq}\,\mathbf{B} \end{bmatrix} \tag{E.4}$$

Furthermore let $\mathbf{Q}$ and $\widetilde{\mathbf{Q}}$ the matrices such that:

$$\begin{cases} \mathbf{Q}\,(\underline{x}\otimes\underline{x}) = \underline{x}\widetilde{\otimes}\underline{x} \\ \widetilde{\mathbf{Q}}\,\left(\underline{x}\widetilde{\otimes}\underline{x}\right) = \underline{x}\otimes\underline{x} \end{cases} \tag{E.5}$$

The product $\mathbf{A}\widetilde{\otimes}\mathbf{B}$ can be written using $\mathbf{A}$ and $\mathbf{B}$ as follows:

$$\mathbf{A}\widetilde{\otimes}\mathbf{B} = \mathbf{Q}\,(\mathbf{A}\otimes\mathbf{B})\,\widetilde{\mathbf{Q}} \tag{E.6}$$

Example E.2. *Assume that* $\underline{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$. *Then:*

$$\begin{aligned} (\mathbf{A}\,\underline{x})\,\widetilde{\otimes}\,(\mathbf{B}\,\underline{x}) &= \left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)\widetilde{\otimes}\left(\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) \\ &= \begin{bmatrix} a_{11}\,x_1 + a_{12}\,x_2 \\ a_{21}\,x_1 + a_{22}\,x_2 \end{bmatrix}\widetilde{\otimes}\begin{bmatrix} b_{11}\,x_1 + b_{12}\,x_2 \\ b_{21}\,x_1 + b_{22}\,x_2 \end{bmatrix} \\ &= \begin{bmatrix} (a_{11}\,x_1 + a_{12}\,x_2)\,(b_{11}\,x_1 + b_{12}\,x_2) \\ (a_{11}\,x_1 + a_{12}\,x_2)\,(b_{21}\,x_1 + b_{22}\,x_2) \\ (a_{21}\,x_1 + a_{22}\,x_2)\,(b_{21}\,x_1 + b_{22}\,x_2) \end{bmatrix} \\ &= \begin{bmatrix} a_{11}\,b_{11} & a_{11}\,b_{12} + a_{12}\,b_{11} & a_{12}\,b_{12} \\ a_{11}\,b_{21} & a_{11}\,b_{22} + a_{12}\,b_{21} & a_{12}\,b_{22} \\ a_{21}\,b_{21} & a_{21}\,b_{22} + a_{22}\,b_{21} & a_{22}\,b_{22} \end{bmatrix}\begin{bmatrix} x_1^2 \\ x_1\,x_2 \\ x_2^2 \end{bmatrix} \\ &:= \left(\mathbf{A}\widetilde{\otimes}\mathbf{B}\right)\underline{x}^{[2]} \end{aligned} \tag{E.7}$$

Consequently:

$$\mathbf{A}\widetilde{\otimes}\mathbf{B} = \begin{bmatrix} a_{11}\,b_{11} & a_{11}\,b_{12} + a_{12}\,b_{11} & a_{12}\,b_{12} \\ a_{11}\,b_{21} & a_{11}\,b_{22} + a_{12}\,b_{21} & a_{12}\,b_{22} \\ a_{21}\,b_{21} & a_{21}\,b_{22} + a_{22}\,b_{21} & a_{22}\,b_{22} \end{bmatrix} \tag{E.8}$$

On the other hand we have:

$$\begin{cases} \mathbf{Q}\,(\underline{x}\otimes\underline{x}) = \underline{x}\widetilde{\otimes}\underline{x} \\ \widetilde{\mathbf{Q}}\,\left(\underline{x}\widetilde{\otimes}\underline{x}\right) = \underline{x}\otimes\underline{x} \end{cases} \Leftrightarrow \begin{cases} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1^2 \\ x_1x_2 \\ x_2x_1 \\ x_2^2 \end{bmatrix} = \begin{bmatrix} x_1^2 \\ x_1x_2 \\ x_2^2 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1^2 \\ x_1x_2 \\ x_2^2 \end{bmatrix} = \begin{bmatrix} x_1^2 \\ x_1x_2 \\ x_2x_1 \\ x_2^2 \end{bmatrix} \end{cases} \tag{E.9}$$

[2] Kathrin Schacke, On the Kronecker Product, August 1, 2013

It can be checked that:

$$\mathbf{A}\widetilde{\otimes}\mathbf{B} = \mathbf{Q}\left(\mathbf{A}\otimes\mathbf{B}\right)\widetilde{\mathbf{Q}} \tag{E.10}$$

■

E.2 Taylor series expansion of any continuous vector field

Let $f\left(\underline{x}\right)$ be a continuous vector field. The Nth order approximation of this vector field around any point $\underline{x}_0$ can be achieved through the following Taylor series expansion[3]:

$$f\left(\underline{x}\right) \approx f\left(\underline{x}_0\right) + \sum_{k=1}^{N} \frac{1}{k!}\, \partial f^{[k]}\Big|_{\underline{x}=\underline{x}_0} \left(\underline{x}-\underline{x}_0\right)^{[k]} \tag{E.11}$$

Assuming that where $\underline{x}\in\mathbb{R}^n$, each column of $\partial f^{[k]}$ formally corresponds to each term of the development of the following expression:

$$\left(\frac{\partial f}{\partial x_1}+\cdots+\frac{\partial f}{\partial x_n}\right)^k \tag{E.12}$$

We recall that the second partial derivative $\frac{\partial^2 f}{\partial x_i \partial x_j}$ is obtained as follows:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) \tag{E.13}$$

Schwarz's theorem states that if the second derivatives are continuous, the expression for the second partial derivative is unaffected by which variable the partial derivative is taken with respect to first and which is taken second[4]:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} \tag{E.14}$$

The higher order partial derivatives are obtained by successive differentiation.

[3] N. Hashemian and A. Armaou, Fast Moving Horizon Estimation of Nonlinear Processes via Carleman Linearization, 2015 American Control Conference, Palmer House Hilton, July 1-3, 2015. Chicago, IL, USA

[4] https://en.wikipedia.org/wiki/Partial_derivative

Example E.3. *Assume that* $\underline{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$*, then we get:*

$$\begin{aligned}
k=1 \quad & \Rightarrow \underline{x}^{[1]} = \underline{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T \\
& \Rightarrow \left(\frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2}\right)^1 = \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2} \\
& \Rightarrow \partial f^{[1]} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix} \\
k=2 \quad & \Rightarrow \underline{x}^{[2]} = \underline{x} \widetilde{\otimes} \underline{x} = \begin{bmatrix} x_1^2 & x_1 x_2 & x_2^2 \end{bmatrix}^T \\
& \Rightarrow \left(\frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2}\right)^2 = \frac{\partial^2 f}{\partial x_1^2} + 2\,\frac{\partial^2 f}{\partial x_1\,\partial x_2} + \frac{\partial^2 f}{\partial x_2^2} \\
& \Rightarrow \partial f^{[2]} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & 2\,\frac{\partial^2 f}{\partial x_1\,\partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} \\
k=3 \quad & \Rightarrow \underline{x}^{[3]} = \underline{x}^{[2]} \widetilde{\otimes} \underline{x} = \begin{bmatrix} x_1^3 & x_1^2 x_2 & x_1 x_2^2 & x_2^3 \end{bmatrix}^T \\
& \Rightarrow \left(\frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2}\right)^3 = \frac{\partial^3 f}{\partial x_1^3} + 3\,\frac{\partial^3 f}{\partial x_1^2\,\partial x_2} + 3\,\frac{\partial^3 f}{\partial x_1\,\partial x_2^2} + \frac{\partial^2 f}{\partial x_2^3} \\
& \Rightarrow \partial f^{[3]} = \begin{bmatrix} \frac{\partial^3 f}{\partial x_1^3} & 3\,\frac{\partial^3 f}{\partial x_1^2\,\partial x_2} & 3\,\frac{\partial^3 f}{\partial x_1\,\partial x_2^2} & \frac{\partial^2 f}{\partial x_2^3} \end{bmatrix}
\end{aligned} \tag{E.15}$$

Consequently, the third *order approximation of* $f(\underline{x})$ *around any point* $\underline{x}_0$ *reads as follows:*

$$\begin{aligned}
f(\underline{x}) \quad & \approx f(\underline{x}_0) + \textstyle\sum_{k=1}^{3} \frac{1}{k!}\, \partial f^{[k]}\big|_{\underline{x}=\underline{x}_0} (\underline{x} - \underline{x}_0)^{[k]} \\
& \approx f(\underline{x}_0) + \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix}_{\underline{x}=\underline{x}_0} \begin{bmatrix} x_1 - x_{10} \\ x_2 - x_{20} \end{bmatrix} \\
& \quad + \tfrac{1}{2!} \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & 2\,\frac{\partial^2 f}{\partial x_1\,\partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}_{\underline{x}=\underline{x}_0} \begin{bmatrix} (x_1 - x_{10})^2 \\ (x_1 - x_{10})\,(x_2 - x_{20}) \\ (x_2 - x_{20})^2 \end{bmatrix} \\
& \quad + \tfrac{1}{3!} \begin{bmatrix} \frac{\partial^3 f}{\partial x_1^3} & 3\,\frac{\partial^3 f}{\partial x_1^2\,\partial x_2} & 3\,\frac{\partial^3 f}{\partial x_1\,\partial x_2^2} & \frac{\partial^2 f}{\partial x_2^3} \end{bmatrix}_{\underline{x}=\underline{x}_0} \begin{bmatrix} (x_1 - x_{10})^3 \\ (x_1 - x_{10})^2\,(x_2 - x_{20}) \\ (x_1 - x_{10})\,(x_2 - x_{20})^2 \\ (x_2 - x_{20})^3 \end{bmatrix}
\end{aligned} \tag{E.16}$$

■

E.3 Time derivative

To get time derivative of *reduced* Kronecker product, the product rule holds:

$$\frac{d}{dt}\,\underline{x}^{[2]} = \frac{d}{dt}\,\underline{x} \widetilde{\otimes} \underline{x} = \dot{\underline{x}} \widetilde{\otimes} \underline{x} + \underline{x} \widetilde{\otimes} \dot{\underline{x}} \tag{E.17}$$

Example E.4. *Assume that* $\underline{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$*. Then expanding definition* (E.1) *we get:*

$$\begin{cases}
\dot{\underline{x}} \widetilde{\otimes} \underline{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} \widetilde{\otimes} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \dot{x}_1 x_1 \\ \dot{x}_1 x_2 \\ \dot{x}_2 x_2 \end{bmatrix} \\[2ex]
\underline{x} \widetilde{\otimes} \dot{\underline{x}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \widetilde{\otimes} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_1 \dot{x}_1 \\ x_1 \dot{x}_2 \\ x_2 \dot{x}_2 \end{bmatrix}
\end{cases} \tag{E.18}$$

It can be checked that:

$$\frac{d}{dt}\,\underline{x}^{[2]} = \frac{d}{dt}\begin{bmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix} = \frac{d}{dt}\,\underline{x}\widetilde{\otimes}\underline{x} = \dot{\underline{x}}\widetilde{\otimes}\underline{x} + \underline{x}\widetilde{\otimes}\dot{\underline{x}} = \begin{bmatrix} 2\,x_1\dot{x}_1 \\ x_1\dot{x}_2 + x_2\dot{x}_1 \\ 2\,x_2\dot{x}_2 \end{bmatrix} \tag{E.19}$$

■

Now assume that:

$$\boxed{\dot{\underline{x}} = \sum_{k=1}^{N_1} \mathbf{A}_k\,\underline{x}^{[k]} + u\sum_{k=0}^{N_2} \mathbf{B}_k\,\underline{x}^{[k]}} \tag{E.20}$$

Then it can be shown than[1]:

$$\boxed{\frac{d}{dt}\,\underline{x}^{[j]} = \sum_{k=1}^{N_1} \mathbf{A}_{jk}\,\underline{x}^{[k+j-1]} + u\sum_{k=0}^{N_2} \mathbf{B}_{jk}\,\underline{x}^{[k+j-1]}} \tag{E.21}$$

where $\mathbf{A}_{1k} = \mathbf{A}_k$ and, for $j > 1$, $\mathbf{A}_{jk}$ is obtained by summing j terms, each term being formed through $j-1$ *reduced* Kronecker product:

$$\mathbf{A}_{jk} = \mathbf{A}_k\,\underbrace{\widetilde{\otimes}\mathbb{I}_n\widetilde{\otimes}\cdots\widetilde{\otimes}\mathbb{I}_n}_{j-1\ times} + \mathbb{I}_n\widetilde{\otimes}\mathbf{A}_k\widetilde{\otimes}\mathbb{I}_n\widetilde{\otimes}\cdots\widetilde{\otimes}\mathbb{I}_n + \cdots + \mathbb{I}_n\widetilde{\otimes}\mathbb{I}_n\widetilde{\otimes}\cdots\widetilde{\otimes}\mathbf{A}_k \tag{E.22}$$

Notation for $\mathbf{B}_{jk}$ is likewise.

Similar to (E.6), $\mathbf{A}_{jk}$ is related to the Kronecker product $\otimes$ through the following relation:

$$\boxed{\begin{aligned} \mathbf{A}_{jk} \;&= \mathbf{Q}_j\,\big(\,(\mathbf{A}_k\mathbf{Q}_k)\,\underbrace{\otimes\mathbb{I}_n\otimes\cdots\otimes\mathbb{I}_n}_{j-1\ times} \\ &\quad + \mathbb{I}_n\otimes(\mathbf{A}_k\mathbf{Q}_k)\otimes\mathbb{I}_n\otimes\cdots\otimes\mathbb{I}_n + \cdots \\ &\quad \cdots + (\mathbb{I}_n\otimes\cdots\otimes\mathbb{I}_n\otimes(\mathbf{A}_k\mathbf{Q}_k))\,\big)\,\widetilde{\mathbf{Q}}_{k+j-1} \end{aligned}} \tag{E.23}$$

where n is the size of vector $\underline{x}$ and where matrices $\mathbf{Q}_j$, $\mathbf{Q}_k$ and $\widetilde{\mathbf{Q}}_{k+j-1}$ are such that:

$$\begin{cases} \mathbf{Q}_j\,(\underbrace{\underline{x}\otimes\underline{x}\cdots\otimes\underline{x}}_{j\ times}) = \underline{x}^{[j]} \\ \mathbf{Q}_k\,(\underbrace{\underline{x}\otimes\underline{x}\cdots\otimes\underline{x}}_{k\ times}) = \underline{x}^{[k]} \\ \widetilde{\mathbf{Q}}_{k+j-1}\,\underline{x}^{[k+j-1]} = (\underbrace{\underline{x}\otimes\underline{x}\cdots\otimes\underline{x}}_{k+j-1\ times}) \end{cases} \tag{E.24}$$

Example E.5. *Assume that* $\underline{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ *and that:*

$$\dot{\underline{x}} = \mathbf{A}_1\,\underline{x} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11}\,x_1 + a_{12}\,x_2 \\ a_{21}\,x_1 + a_{22}\,x_2 \end{bmatrix} \tag{E.25}$$

Then we have:

$$\underline{x}^{[3]} = \begin{bmatrix} x_1^3 \\ x_1^2 x_2 \\ x_1 x_2^2 \\ x_2^3 \end{bmatrix} \Rightarrow \frac{d}{dt}\underline{x}^{[3]} = \begin{bmatrix} 3x_1^2 \dot{x}_1 \\ 2x_1 x_2 \dot{x}_1 + x_1^2 \dot{x}_2 \\ x_2^2 \dot{x}_1 + 2x_1 x_2 \dot{x}_2 \\ 3x_2^2 \dot{x}_2 \end{bmatrix} \tag{E.26}$$

Using the relation $\dot{\underline{x}} = \mathbf{A}_1\, \underline{x}$ *we get:*

$$\begin{aligned} \tfrac{d}{dt}\underline{x}^{[3]} &= \begin{bmatrix} 3x_1^2\,(a_{11}\,x_1 + a_{12}\,x_2) \\ 2x_1 x_2\,(a_{11}\,x_1 + a_{12}\,x_2) + x_1^2\,(a_{21}\,x_1 + a_{22}\,x_2) \\ x_2^2\,(a_{11}\,x_1 + a_{12}\,x_2) + 2x_1 x_2\,(a_{21}\,x_1 + a_{22}\,x_2) \\ 3x_2^2\,(a_{21}\,x_1 + a_{22}\,x_2) \end{bmatrix} \\ &= \begin{bmatrix} 3a_{11}x_1^3 + 3a_{12}x_2x_1^2 \\ a_{21}x_1^3 + 2a_{11}x_1^2x_2 + 2a_{12}x_1x_2^2 + a_{22}x_1^2x_2 \\ a_{12}x_2^3 + a_{11}x_1x_2^2 + 2a_{21}x_1^2x_2 + 2a_{22}x_1x_2^2 \\ 3a_{22}x_2^3 + 3a_{21}x_1x_2^2 \end{bmatrix} \\ &= \begin{bmatrix} 3a_{11} & 3a_{12} & 0 & 0 \\ a_{21} & 2a_{11} + a_{22} & 2a_{12} & 0 \\ 0 & 2a_{21} & a_{11} + 2a_{22} & a_{12} \\ 0 & 0 & 3a_{21} & 3a_{22} \end{bmatrix} \underline{x}^{[3]} \\ &:= \mathbf{A}_{31}\, \underline{x}^{[3]} \end{aligned} \tag{E.27}$$

It can be checked that matrix $\mathbf{A}_{31}$ *can be computed as follows:*

$$\mathbf{A}_{31} = \mathbf{Q}_3 \left((\mathbf{A}_1\mathbf{Q}_1) \underbrace{\otimes \mathbb{I}_2 \otimes \mathbb{I}_2}_{3-1\ times} + \mathbb{I}_2 \otimes (\mathbf{A}_1\mathbf{Q}_1) \otimes \mathbb{I}_2 + \mathbb{I}_2 \otimes \mathbb{I}_2 \otimes (\mathbf{A}_1\mathbf{Q}_1) \right) \widetilde{\mathbf{Q}}_3 \tag{E.28}$$

where matrices $\mathbf{Q}_1$, $\mathbf{Q}_3$ *and* $\widetilde{\mathbf{Q}}_3$ *are such that:*

$$\mathbf{Q}_1\, \underline{x} = \underline{x} \Rightarrow \mathbf{Q}_1 = \mathbb{I}_2 \tag{E.29}$$

and:

$$\begin{cases} \mathbf{Q}_3\, (\underbrace{\underline{x} \otimes \underline{x} \otimes \underline{x}}_{1+3-1\ times}) = \underline{x}^{[3]} := \begin{bmatrix} x_1^3 \\ x_1^2 x_2 \\ x_1 x_2^2 \\ x_2^3 \end{bmatrix} \\ \\ \widetilde{\mathbf{Q}}_3\, \underline{x}^{[3]} = (\underbrace{\underline{x} \otimes \underline{x} \otimes \underline{x}}_{1+3-1\ times}) := \begin{bmatrix} x_1^3 \\ x_1^2 x2 \\ x_1^2 x2 \\ x_1 x2^2 \\ x_1^2 x2 \\ x_1 x2^2 \\ x_1 x2^2 \\ x2^3 \end{bmatrix} \end{cases} \tag{E.30}$$

We get:

$$\begin{cases} \mathbf{Q}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\ \widetilde{\mathbf{Q}}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{cases} \tag{E.31}$$

■

E.4 Carleman linearization of nonlinear system affine in control

We consider the following nonlinear system affine in the control vector $\underline{u} \in \mathbb{R}^m$:

$$\boxed{\dot{\underline{x}} = f(\underline{x}) + \sum_{j=1}^{m} g_j(\underline{x})\, u_j} \tag{E.32}$$

Let $\underline{v} = \begin{bmatrix} \underline{x}^T & \underline{u}^T \end{bmatrix}^T$. Then (E.32) reads:

$$\underline{v} = \begin{bmatrix} \underline{x} \\ \underline{u} \end{bmatrix} \Rightarrow \dot{\underline{x}} = h(\underline{v}) \tag{E.33}$$

Assuming that $h(\underline{v})$ is a continuous vector field, the Nth order approximation of the vector field $h(\underline{v})$ around any point $\underline{v}_e$ is achieved through the following Taylor series expansion according to (E.11):

$$\boxed{\dot{\underline{x}} = h(\underline{v}) \approx h(\underline{v}_e) + \sum_{k=1}^{N} \frac{1}{k!}\, \partial h^{[k]}\Big|_{\underline{v}=\underline{v}_e} \delta \underline{v}^{[k]}} \tag{E.34}$$

where:

$$\delta \underline{v} := \underline{v} - \underline{v}_e = \begin{bmatrix} \underline{x} - \underline{x}_e \\ \underline{u} - \underline{u}_e \end{bmatrix} := \begin{bmatrix} \delta \underline{x} \\ \delta \underline{u} \end{bmatrix} \tag{E.35}$$

Because the vector field $h(\underline{v}) := h(\underline{x}, \underline{u})$ is affine in the control $\underline{u}$, terms in factors of $\delta \underline{u}^{[k]}$ in (E.34) are null $\forall k > 1$. Thus (E.34) can be rewritten as follows:

$$\dot{\underline{x}} \approx h(\underline{v}_e) + \sum_{i=1}^{N} \mathbf{A}_{1i}\, \delta \underline{x}^{[i]} + \sum_{j=1}^{m} \delta u_j \sum_{i=0}^{N} \mathbf{B}_{1ij}\, \delta \underline{x}^{[i]} \tag{E.36}$$

Furthermore, as far as $\underline{x}_e$ is a constant vector, we have:

$$\dot{\underline{x}} = \delta\dot{\underline{x}} + \dot{\underline{x}}_e = \delta\dot{\underline{x}} \tag{E.37}$$

Thus, the (E.36) reads:

$$\delta\dot{\underline{x}} \approx h\left(\underline{v}_e\right) + \sum_{i=1}^{N} \mathbf{A}_{1i}\, \delta\underline{x}^{[i]} + \sum_{j=1}^{m} \delta u_j \sum_{i=0}^{N} \mathbf{B}_{1ij}\, \delta\underline{x}^{[i]} \tag{E.38}$$

Carleman linearization has been introduced by Torsten Carleman in 1932[5]. The common use of Carleman linearization consists in approximating a nonlinear state equation by a finite dimensional polynomial model. Following Minisini & al.[6], the original state vector $\underline{x} \in \mathbb{R}^n$ is extended by repeated application of the Kronecker product to itself to give the following extended state vector $\underline{z}$:

$$\boxed{\underline{z} := \begin{bmatrix} \delta\underline{x}^{[1]} \\ \delta\underline{x}^{[2]} \\ \vdots \\ \delta\underline{x}^{[N]} \end{bmatrix} = \begin{bmatrix} \delta\underline{x} \\ \delta\underline{x}\widetilde{\otimes}\delta\underline{x} \\ \vdots \\ \underbrace{\delta\underline{x}\widetilde{\otimes}\cdots\widetilde{\otimes}\delta\underline{x}}_{N \quad times} \end{bmatrix} \text{ where } \delta\underline{x} := \underline{x} - \underline{x}_e} \tag{E.39}$$

Furthermore we can use (E.21) to write:

$$\frac{d}{dt}\,\delta\underline{x}^{[k]} = \sum_{i=1}^{N_1} \mathbf{A}_{ki}\, \underline{x}^{[i+k-1]} + \sum_{j=1}^{m} \delta u_j \sum_{i=0}^{N_2} \mathbf{B}_{kij}\, \delta\underline{x}^{[i+k-1]} \tag{E.40}$$

When the series is truncated to its first N terms we get:

$$\frac{d}{dt}\,\delta\underline{x}^{[k]} = \sum_{i=1}^{N-k+1} \mathbf{A}_{ki}\, \underline{x}^{[i+k-1]} + \sum_{j=1}^{m} \delta u_j \sum_{i=0}^{N-k+1} \mathbf{B}_{kij}\, \delta\underline{x}^{[i+k-1]} \tag{E.41}$$

Finally, assuming that $\underline{v}_e$ is an equilibrium point, that is $h\left(\underline{v}_e\right) = \underline{0}$, Carleman linearization of order $N \geq 1$ leads to the following bilinear state

[5] Application de la théorie des équations intégrales linéaires aux systèmes d'équations différentielles non linéaires. Acta Math. 59 63 - 87, 1932. https://doi.org/10.1007/BF02546499

[6] J. Minisini, A. Rauh, and E. P. Hofer, Carleman Linearization for Approximate Solutions of Nonlinear Control Problems: Part 1 - Theory. In F. L. Chernousko, G. V. Kostin, and V. V. Saurin, editors, Advances in Mechanics: Dynamics and Control: Proc. of the 14th International Workshop on Dynamics and Control, 2007, Moscow-Zvenigorod, Russia, 2008. Nauka

space representation around the equilibrium point $\underline{v}_e := (\underline{x}_e, \underline{u}_e)$:

$$\boxed{\begin{aligned} \dot{\underline{z}} &= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1N} \\ \mathbf{0} & \mathbf{A}_{21} & \cdots & \mathbf{A}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_{NN} \end{bmatrix} \begin{bmatrix} \delta\underline{x}^{[1]} \\ \delta\underline{x}^{[2]} \\ \vdots \\ \delta\underline{x}^{[N]} \end{bmatrix} + \begin{bmatrix} \mathbf{B}_0 \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \delta\underline{u} \\ &\quad + \textstyle\sum_{j=1}^{m} \delta u_j \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \cdots & \mathbf{B}_{1N} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \cdots & \mathbf{B}_{2N} \\ \mathbf{0} & \mathbf{B}_{31} & \cdots & \mathbf{B}_{3N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_{NN} \end{bmatrix} \begin{bmatrix} \delta\underline{x}^{[1]} \\ \delta\underline{x}^{[2]} \\ \vdots \\ \delta\underline{x}^{[N]} \end{bmatrix} \\ &:= \mathbf{A}\,\underline{z} + \mathbf{B}\,\delta\underline{u} + \textstyle\sum_{j=1}^{m} \delta u_j\, \mathbf{B}_j\, \underline{z} \end{aligned}} \tag{E.42}$$

In practice components of the extended state vector $\underline{z} := \begin{bmatrix} \delta\underline{x}^{[1]} \\ \vdots \\ \delta\underline{x}^{[N]} \end{bmatrix}$ are eliminated if they do not appear in the right hand side of the expression of $\dot{\underline{z}}$.

Example E.6. *We consider the following single input nonlinear system model:*

$$\begin{cases} \dot{x}_1 = -x_1^3 + 1.5\,x_1^2 - 0.5\,x_1 + x_2 \\ \dot{x}_2 = -x_1 - x_2\,u \end{cases} \tag{E.43}$$

With $\underline{v} = \begin{bmatrix} \underline{x} \\ u \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ u \end{bmatrix}$ *we get the following expression:*

$$\dot{\underline{x}} = h\,(\underline{v}) \;\; \textit{where}\;\; h\,(\underline{v}) = \begin{bmatrix} -x_1^3 + 1.5\,x_1^2 - 0.5\,x_1 + x_2 \\ -x_1 - x_2\,u \end{bmatrix} \tag{E.44}$$

We can see that $h\,(\underline{0}) = \underline{0}$. *Thus* $\underline{v}_e = (\underline{x}_e, u_e) = (\underline{0}, 0)$ *is an equilibrium point. The second order linearization of the vector field* $h\,(\underline{v})$ *around the equilibrium point* $\underline{v}_e = \underline{0}$ *reads:*

$$\dot{\underline{x}} = h\,(\underline{v}) \approx h\,(\underline{v}_e) + \sum_{k=1}^{2} \frac{1}{k!}\, \partial h^{[k]}\Big|_{\underline{v}=\underline{v}_e} \delta\underline{v}^{[k]} = \sum_{k=1}^{2} \frac{1}{k!}\, \partial h^{[k]}\Big|_{\underline{v}=\underline{v}_e} \delta\underline{v}^{[k]} \tag{E.45}$$

where:

$$\begin{cases} \underline{v}^{[1]} &= \begin{bmatrix} x_1 & x_2 & u \end{bmatrix}^T \\ \partial h^{[1]}\big|_{\underline{v}=\underline{0}} &= \begin{bmatrix} -3x_1^2 + 3x_1 - 0.5 & 1 & 0 \\ -1 & -u & -x_2 \end{bmatrix}\Bigg|_{\underline{v}=\underline{0}} \\ &= \begin{bmatrix} -0.5 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \end{cases} \tag{E.46}$$

and:

$$\left\{\begin{array}{rl} \underline{v}^{[2]} & = \left[\begin{array}{cccccc} x_1^2 & x_1\,x_2 & x_1\,u & x_2^2 & x_2\,u & u^2 \end{array}\right]^T \\ \partial h^{[2]}\big|_{\underline{v}=\underline{0}} & = \left[\begin{array}{cccccc} -6x_1+3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{array}\right]\Bigg|_{\underline{v}=\underline{0}} \\ & = \left[\begin{array}{cccccc} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{array}\right] \end{array}\right. \tag{E.47}$$

Then we get:

$$\begin{array}{rl} h\left(\underline{v}\right) & \approx \sum_{k=1}^{2} \frac{1}{k!}\, \partial h^{[k]}\big|_{\underline{v}=\underline{v}_e}\, \delta\underline{v}^{[k]} \\ & \approx \left[\begin{array}{ccc} -0.5 & 1 & 0 \\ -1 & 0 & 0 \end{array}\right] \left[\begin{array}{c} \delta x_1 \\ \delta x_2 \\ \delta u \end{array}\right] \\ & \qquad + \frac{1}{2} \left[\begin{array}{cccccc} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{array}\right] \left[\begin{array}{c} \delta x_1^2 \\ \delta x_1\,x_2 \\ \delta x_1\,\delta u \\ \delta x_2^2 \\ \delta x_2\,\delta u \\ \delta u^2 \end{array}\right] \\ & \approx \left[\begin{array}{cc} -0.5 & 1 \\ -1 & 0 \end{array}\right] \left[\begin{array}{c} \delta x_1 \\ \delta x_2 \end{array}\right] + \frac{1}{2} \left[\begin{array}{ccc} 3 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] \left[\begin{array}{c} \delta x_1^2 \\ \delta x_1\,x_2 \\ \delta x_2^2 \end{array}\right] \\ & \qquad + \delta u \left[\begin{array}{cc} 0 & 0 \\ 0 & -0.5 \end{array}\right] \left[\begin{array}{c} \delta x_1 \\ \delta x_2 \end{array}\right] \end{array} \tag{E.48}$$

Thus around the considered equilibrium point the following approximation holds:

$$\begin{array}{rl} \left[\begin{array}{c} \delta\dot{x}_1 \\ \delta\dot{x}_2 \end{array}\right] = h\left(\underline{v}\right) \approx & \left[\begin{array}{cc} -0.5 & 1 \\ -1 & 0 \end{array}\right] \delta\underline{x}^{[1]} \\ & + \dfrac{1}{2} \left[\begin{array}{ccc} 3 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] \delta\underline{x}^{[2]} + \delta u \left[\begin{array}{cc} 0 & 0 \\ 0 & -0.5 \end{array}\right] \delta\underline{x}^{[1]} \end{array} \tag{E.49}$$

Furthermore, the time derivative of $\delta\underline{x}^{[2]}$ reads as follows;

$$\begin{array}{rl} \delta\dot{\underline{x}}^{[2]} & = \frac{d}{dt} \left[\begin{array}{c} \delta x_1^2 \\ \delta x_1\,\delta x_2 \\ \delta x_2^2 \end{array}\right] \\ & = \left[\begin{array}{c} 2\delta x_1 \delta\dot{x}_1 \\ \delta\dot{x}_1\,\delta x_2 + \delta x_1\,\delta\dot{x}_2 \\ 2\delta x_2 \delta\dot{x}_2 \end{array}\right] \end{array} \tag{E.50}$$

Then using the approximation (E.49) *we get:*

$$\begin{aligned}
\delta\dot{\underline{x}}^{[2]} &\approx \begin{bmatrix} 2\delta x_1\left(-0.5\delta x_1+\delta x_2+1.5\delta x_1^2\right) \\ \left(-0.5\delta x_1+\delta x_2+1.5\delta x_1^2\right)\,\delta x_2+\delta x_1\,\left(-\delta x_1-0.5\delta u\delta x_2\right) \\ 2\delta x_2\left(-\delta x_1-0.5\delta u\delta x_2\right) \end{bmatrix} \\
&\approx \begin{bmatrix} -\delta x_1^2+2\delta x_1\delta x_2+3\delta x_1^3 \\ -\delta x_1^2-0.5\delta x_1\delta x_2+\delta x_2^2+1.5\delta x_1^2\delta x_2-0.5\delta u\delta x_1\delta x_2 \\ -2\delta x_1\delta x_2-\delta u\delta x_2^2 \end{bmatrix} \\
&\approx \begin{bmatrix} -1 & 2 & 0 \\ -1 & -0.5 & 1 \\ 0 & -2 & 0 \end{bmatrix}\delta\underline{x}^{[2]}+\delta u\begin{bmatrix} 0 & 0 & 0 \\ 0 & -0.5 & 0 \\ 0 & 0 & -1 \end{bmatrix}\delta\underline{x}^{[2]} \\
&\qquad+\begin{bmatrix} 3\delta x_1^3 \\ 1.5\delta x_1^2\delta x_2 \\ -2\delta x_1\delta x_2 \end{bmatrix}
\end{aligned} \tag{E.51}$$

Limiting the approximation to the terms involving $\delta\underline{x}^{[2]}$ *we get:*

$$\delta\dot{\underline{x}}^{[2]} \approx \begin{bmatrix} -1 & 2 & 0 \\ -1 & -0.5 & 1 \\ 0 & -2 & 0 \end{bmatrix}\delta\underline{x}^{[2]}+\delta u\begin{bmatrix} 0 & 0 & 0 \\ 0 & -0.5 & 0 \\ 0 & 0 & -1 \end{bmatrix}\delta\underline{x}^{[2]} \tag{E.52}$$

Finally order 2 *Carlerman linearization is obtained by merging relations* (E.49) *and* (E.52)*:*

$$\begin{aligned}
\begin{bmatrix} \delta\dot{\underline{x}}^{[1]} \\ \delta\dot{\underline{x}}^{[2]} \end{bmatrix} &= \left[\begin{array}{cc|ccc} -0.5 & 1 & 1.5 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & -0.5 & 1 \\ 0 & 0 & 0 & -2 & 0 \end{array}\right]\begin{bmatrix} \delta\underline{x}^{[1]} \\ \delta\underline{x}^{[2]} \end{bmatrix} \\
&\qquad+\delta u\left[\begin{array}{cc|ccc} 0 & 0 & 0 & 0 & 0 \\ 0 & -0.5 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.5 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right]\begin{bmatrix} \delta\underline{x}^{[1]} \\ \delta\underline{x}^{[2]} \end{bmatrix}
\end{aligned} \tag{E.53}$$

■

E.5 Application to mechanical systems

We consider systems described by the following differential equation where $\underline{N}$ is the generalized coordinates vector, $\mathbf{M}(\underline{q})$ is a positive definite (and thus symmetric) matrix known as the inertia matrix, $\mathbf{C}(\underline{q},\dot{\underline{q}})$ is the matrix of terms generated by centrifugal and Coriolis forces, $\underline{G}(\underline{q})$ is the vector of conservative forces, $\mathbf{H}$ is a constant matrix and $\underline{u}$ the control vector:

$$\mathbf{M}(\underline{q})\,\ddot{\underline{q}}+\mathbf{C}(\underline{q},\dot{\underline{q}})\,\dot{\underline{q}}+\underline{G}(\underline{q})=\mathbf{H}\underline{u} \tag{E.54}$$

Let $\underline{x}$ be the following state vector:

$$\underline{x}=\begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix}:=\begin{bmatrix} \underline{q} \\ \dot{\underline{q}} \end{bmatrix} \tag{E.55}$$

Thus (E.54) can be written as follows:

$$\dot{\underline{x}} = \begin{bmatrix} \dot{\underline{x}}_1 \\ \dot{\underline{x}}_2 \end{bmatrix} := \begin{bmatrix} \underline{x}_2 \\ -\mathbf{M}^{-1}(\underline{x}_2)\,\left(\mathbf{C}(\underline{x}_1,\underline{x}_2)\,\underline{x}_2 + \underline{G}(\underline{x}_1)\right) \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{M}^{-1}(\underline{x}_2)\,\mathbf{H} \end{bmatrix} \underline{u} \tag{E.56}$$

Or, more generally:

$$\boxed{\dot{\underline{x}} = \begin{bmatrix} \dot{\underline{x}}_1 \\ \dot{\underline{x}}_2 \end{bmatrix} = \begin{bmatrix} \underline{x}_2 \\ f(\underline{x}_1,\underline{x}_2) \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ g(\underline{x}_2) \end{bmatrix} \underline{u}} \tag{E.57}$$

where:

$$\begin{cases} f(\underline{x}_1,\underline{x}_2) := -\mathbf{M}^{-1}(\underline{x}_2)\,\left(\mathbf{C}(\underline{x}_1,\underline{x}_2)\,\underline{x}_2 + \underline{G}(\underline{x}_1)\right) \\ g(\underline{x}_2) := \mathbf{M}^{-1}(\underline{x}_2)\,\mathbf{H} \end{cases} \tag{E.58}$$

Let $(\underline{x}_{1e}, \underline{u}_e)$ be an equilibrium point. Then $\underline{x}_{2e} = \dot{\underline{x}}_{1e} := \underline{0}$ and the following equality holds:

$$\boxed{\underline{0} = f(\underline{x}_{1e}, \underline{0}) + g(\underline{0})\,\underline{u}_e} \tag{E.59}$$

Let $\underline{v} = \begin{bmatrix} \underline{x}^T & \underline{u}^T \end{bmatrix}^T$. Then the second equation of (E.57) reads:

$$\underline{v} = \begin{bmatrix} \underline{x} \\ \underline{u} \end{bmatrix} \Rightarrow \dot{\underline{x}}_2 = f(\underline{x}_1,\underline{x}_2) + g(\underline{x}_2)\,\underline{u} := h\left(\underline{v}\right) \tag{E.60}$$

Assuming that $h\left(\underline{v}\right)$ is a continuous vector field, the qth order approximation of the vector field around any equilibrium point $\underline{v}_e$ is achieved through the following Taylor series expansion according to (E.11):

$$h\left(\underline{v}_e\right) = \underline{0} \Rightarrow \dot{\underline{x}}_2 = h\left(\underline{v}\right) \approx \sum_{k=1}^{q} \frac{1}{k!}\, \partial h^{[k]}\Big|_{\underline{v}=\underline{v}_e}\, \delta\underline{v}^{[k]} \tag{E.61}$$

where:

$$\delta\underline{v} := \underline{v} - \underline{v}_e = \begin{bmatrix} \underline{x} - \underline{x}_e \\ \underline{u} - \underline{u}_e \end{bmatrix} := \begin{bmatrix} \delta\underline{x} \\ \delta\underline{u} \end{bmatrix} \tag{E.62}$$

Now the results of section E.4 can be applied. Indeed, notice that vector field $h\left(\underline{v}\right) := h\left(\underline{x}, \underline{u}\right)$ is affine in the control $\underline{u}$. Thus terms in factors of $\delta\underline{u}^{[k]}$ are null $\forall k > 1$ and (E.61) can be rewritten as follows:

$$\dot{\underline{x}}_2 \approx \sum_{i=1}^{N} \widetilde{\mathbf{A}}_{1i}\, \delta\underline{x}^{[i]} + \sum_{j=1}^{m} \delta u_j \sum_{i=0}^{N} \widetilde{\mathbf{B}}_{1ij}\, \delta\underline{x}^{[i]} \tag{E.63}$$

Then (E.57) can be approximated as follows around any equilibrium point:

$$\boxed{\delta\dot{\underline{x}} = \begin{bmatrix} \delta\dot{\underline{x}}_1 \\ \delta\dot{\underline{x}}_2 \end{bmatrix} = \begin{bmatrix} \delta\underline{x}_2 \\ \sum_{i=1}^{N} \widetilde{\mathbf{A}}_{1i}\, \delta\underline{x}^{[i]} + \sum_{j=1}^{m} \delta u_j \sum_{i=0}^{N} \widetilde{\mathbf{B}}_{1ij}\, \delta\underline{x}^{[i]} \end{bmatrix}} \tag{E.64}$$

where:

$$\begin{bmatrix} \delta\underline{x} \\ \delta\underline{u} \end{bmatrix} := \begin{bmatrix} \underline{x} - \underline{x}_e \\ \underline{u} - \underline{u}_e \end{bmatrix} \tag{E.65}$$

Finally formalism of section E.4 can be applied and leads to Carleman linearization of the form (E.42).

E.6 Control law for bilinear systems

E.6.1 Linear state feedback

We present hereafter the results provided by I. Sánchez and J. Collado[7]. This method considers the following time invariant bilinear system where state vector $\underline{x} \in \mathbb{R}^n$, control $\underline{u} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{B}_i \in \mathbb{R}^{n\times m}$:

$$\dot{\underline{x}} = \mathbf{A}\,\underline{x} + \mathbf{B}_0\,\underline{u} + \left(\sum_{i=1}^{n} x_i \mathbf{B}_i\right)\underline{u} \tag{E.66}$$

Then a locally asymptotically stabilizing linear time invariant state feedback gain where $\mathbf{K} \in \mathbb{R}^{m\times n}$ reads:

$$\underline{u} = \mathbf{K}\,\underline{x} \tag{E.67}$$

Let $\mathbf{R}$ a positive definite matrix such that some positive definite matrix $\mathbf{T} \in \mathbb{R}^{n\times n}$ the following inequality holds:

$$\mathbf{R} \geq \left(\sum_{i=1}^{n} x_i \mathbf{B}_i\right)^T \mathbf{T} \left(\sum_{i=1}^{n} x_i \mathbf{B}_i\right) \tag{E.68}$$

where:

$$\begin{aligned} \left(\textstyle\sum_{i=1}^{n} x_i \mathbf{B}_i\right)^T \mathbf{T} \quad & \left(\textstyle\sum_{i=1}^{n} x_i \mathbf{B}_i\right) \\ &= \left(\begin{bmatrix} \mathbf{B}_1 & \cdots & \mathbf{B}_n \end{bmatrix} \underline{x}\right)^T \mathbf{T} \left(\begin{bmatrix} \mathbf{B}_1 & \cdots & \mathbf{B}_n \end{bmatrix} \underline{x}\right) \\ &= \underline{x}^T \left(\begin{bmatrix} \mathbf{B}_1 & \cdots & \mathbf{B}_n \end{bmatrix}^T \mathbf{T} \begin{bmatrix} \mathbf{B}_1 & \cdots & \mathbf{B}_n \end{bmatrix}\right) \underline{x} \end{aligned} \tag{E.69}$$

Then state feedback gain $\mathbf{K}$ is such that for some positive definite matrix $\mathbf{Q} \in \mathbb{R}^{n\times n}$ there exists a positive definite matrix $\mathbf{P}$ which is solution of the following matrix equation:

$$\mathbf{0} = (\mathbf{A} + \mathbf{B}_0\mathbf{K})^T \mathbf{P} + \mathbf{P}\,(\mathbf{A} + \mathbf{B}_0\mathbf{K}) + \mathbf{K}^T\mathbf{R}\mathbf{K} + \mathbf{P}\mathbf{T}^{-1}\mathbf{P} + \mathbf{Q} \tag{E.70}$$

In the case where $\mathbf{K} = -\mathbf{R}^{-1}\,\mathbf{B}_0^T\,\mathbf{P}$ then $\mathbf{P}$ is the positive definite matrix which is solution of the following algebraic Riccati equation:

$$\mathbf{K} = -\mathbf{R}^{-1}\,\mathbf{B}_0^T\,\mathbf{P} \Rightarrow \mathbf{0} = \mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A} - \mathbf{P}\left(\mathbf{B}_0\mathbf{R}^{-1}\mathbf{B}_0^T - \mathbf{T}^{-1}\right)\mathbf{P} + \mathbf{Q} \tag{E.71}$$

To get this result[7], we start with the following candidate Lyapunov function $V(\underline{x})$ where $\mathbf{P}$ is a positive definite matrix:

$$V(\underline{x}) = \underline{x}^T\mathbf{P}\underline{x} \tag{E.72}$$

Then setting $\underline{u} = \mathbf{K}\,\underline{x}$ and using (E.66), the time derivative of $V(\underline{x})$ reads:

$$\begin{aligned} \dot{V}(\underline{x}) \quad &= \dot{\underline{x}}^T\mathbf{P}\underline{x} + \underline{x}^T\mathbf{P}\dot{\underline{x}} \\ &= \left(\mathbf{A}\,\underline{x} + \mathbf{B}_0\,\mathbf{K}\,\underline{x} + \left(\textstyle\sum_{i=1}^{n} x_i\mathbf{B}_i\right)\mathbf{K}\,\underline{x}\right)^T \mathbf{P}\underline{x} \\ &\quad + \underline{x}^T\mathbf{P}\left(\left(\mathbf{A}\,\underline{x} + \mathbf{B}_0\,\mathbf{K}\,\underline{x} + \left(\textstyle\sum_{i=1}^{n} x_i\mathbf{B}_i\right)\mathbf{K}\,\underline{x}\right)\right) \\ &= \underline{x}^T\left(\mathbf{A}^T\mathbf{P} + \mathbf{K}^T\mathbf{B}_0^T\mathbf{P} + \mathbf{K}^T\left(\textstyle\sum_{i=1}^{n} x_i\mathbf{B}_i\right)^T\mathbf{P}\right)\underline{x} \\ &\quad + \underline{x}^T\left(\mathbf{P}\mathbf{A} + \mathbf{P}\mathbf{B}_0\mathbf{K} + \mathbf{P}\left(\textstyle\sum_{i=1}^{n} x_i\mathbf{B}_i\right)\mathbf{K}\right)\underline{x} \end{aligned} \tag{E.73}$$

[7] I. Sánchez and J. Collado, On a construction of a non-linear control law for non-linear systems through Carleman Bilinearization, 49th IEEE Conference on Decision and Control (CDC), 2010, pp. 4024-4029, doi: 10.1109/CDC.2010.5717625.

Let vector $\underline{v}$ be defined as follows:

$$\underline{v} = \left(\sum_{i=1}^{n} x_i \mathbf{B}_i \right) \mathbf{K}\, \underline{x} \tag{E.74}$$

Then the time derivative of $V(\underline{x})$ reads:

$$\begin{aligned} \dot{V}(\underline{x}) &= \underline{x}^T \left(\mathbf{A}^T \mathbf{P} + \mathbf{K}^T \mathbf{B}_0^T \mathbf{P} \right) \underline{x} + \underline{v}^T \mathbf{P} \underline{x} \\ &\quad + \underline{x}^T \left(\mathbf{P}\mathbf{A} + \mathbf{P}\mathbf{B}_0\mathbf{K} \right) \underline{x} + \underline{x}^T \mathbf{P} \underline{v} \\ &= \underline{x}^T \left(\left(\mathbf{A} + \mathbf{B}_0\mathbf{K} \right)^T \mathbf{P} + \mathbf{P} \left(\mathbf{A} + \mathbf{B}_0\mathbf{K} \right) \right) \underline{x} + 2\underline{x}^T \mathbf{P} \underline{v} \end{aligned} \tag{E.75}$$

On the other hand, we can write:

$$\begin{aligned} \left(\mathbf{P}\underline{x} - \mathbf{T}\underline{v} \right)^T \mathbf{T}^{-1} \left(\mathbf{P}\underline{x} - \mathbf{T}\underline{v} \right) &= \left(\underline{x}^T \mathbf{P} - \underline{v}^T \mathbf{T} \right) \left(\mathbf{T}^{-1} \mathbf{P} \underline{x} - \underline{v} \right) \\ &= \underline{x}^T \mathbf{P} \mathbf{T}^{-1} \mathbf{P} \underline{x} - 2\underline{x}^T \mathbf{P} \underline{v} + \underline{v}^T \mathbf{T} \underline{v} \end{aligned} \tag{E.76}$$

Thus:

$$2\underline{x}^T \mathbf{P} \underline{v} = \underline{x}^T \mathbf{P} \mathbf{T}^{-1} \mathbf{P} \underline{x} + \underline{v}^T \mathbf{T} \underline{v} - \left(\mathbf{P}\underline{x} - \mathbf{T}\underline{v} \right)^T \mathbf{T}^{-1} \left(\mathbf{P}\underline{x} - \mathbf{T}\underline{v} \right) \tag{E.77}$$

Then $\dot{V}(\underline{x})$ becomes:

$$\begin{aligned} \dot{V}(\underline{x}) &= \underline{x}^T \left(\left(\mathbf{A} + \mathbf{B}_0\mathbf{K} \right)^T \mathbf{P} + \mathbf{P} \left(\mathbf{A} + \mathbf{B}_0\mathbf{K} \right) + \mathbf{P}\mathbf{T}^{-1}\mathbf{P} \right) \underline{x} \\ &\quad - \left(\mathbf{P}\underline{x} - \mathbf{T}\underline{v} \right)^T \mathbf{T}^{-1} \left(\mathbf{P}\underline{x} - \mathbf{T}\underline{v} \right) \\ &\quad + \underline{v}^T \mathbf{T} \underline{v} \\ &= \underline{x}^T \left(\left(\mathbf{A} + \mathbf{B}_0\mathbf{K} \right)^T \mathbf{P} + \mathbf{P} \left(\mathbf{A} + \mathbf{B}_0\mathbf{K} \right) + \mathbf{P}\mathbf{T}^{-1}\mathbf{P} + \mathbf{K}^T \mathbf{R} \mathbf{K} \right) \underline{x} \\ &\quad - \left(\mathbf{P}\underline{x} - \mathbf{T}\underline{v} \right)^T \mathbf{T}^{-1} \left(\mathbf{P}\underline{x} - \mathbf{T}\underline{v} \right) \\ &\quad + \underline{v}^T \mathbf{T} \underline{v} - \underline{x}^T \mathbf{K}^T \mathbf{R} \mathbf{K} \underline{x} \end{aligned} \tag{E.78}$$

Finally using (E.68) we get the following inequality:

$$\underline{v}^T \mathbf{T} \underline{v} = \underline{x}^T \mathbf{K}^T \left(\sum_{i=1}^{n} x_i \mathbf{B}_i \right)^T \mathbf{T} \left(\sum_{i=1}^{n} x_i \mathbf{B}_i \right) \mathbf{K} \underline{x} \leq \underline{x}^T \mathbf{K}^T \mathbf{R} \mathbf{K} \underline{x} \tag{E.79}$$

Thus $\underline{v}^T \mathbf{T} \underline{v} - \underline{x}^T \mathbf{K}^T \mathbf{R} \mathbf{K} \underline{x} \leq 0$ and $\dot{V}(\underline{x}) < 0 \; \forall \underline{x} \neq \underline{0}$ as soon as the following equality holds where $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is some positive definite matrix:

$$\left(\mathbf{A} + \mathbf{B}_0\mathbf{K} \right)^T \mathbf{P} + \mathbf{P} \left(\mathbf{A} + \mathbf{B}_0\mathbf{K} \right) + \mathbf{K}^T \mathbf{R} \mathbf{K} + \mathbf{P}\mathbf{T}^{-1}\mathbf{P} = -\mathbf{Q} < 0 \tag{E.80}$$

This completes the proof. ■

E.6.2 Quadratic state feedback

Consider the following bilinear system where state vector $\underline{x} \in \mathbb{R}^n$, control $\underline{u} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B}_0 \in \mathbb{R}^{n \times m}$ and $\mathbf{B}_j \in \mathbb{R}^{n \times n} \; \forall j > 0$:

$$\dot{\underline{x}} = \mathbf{A}\, \underline{x} + \mathbf{B}_0\, \underline{u} + \left(\sum_{j=1}^{m} u_j \mathbf{B}_j \right) \underline{x} \tag{E.81}$$

Let $\mathbf{B}_0$ be split as follows where $\underline{b}_{j0}$ represents the jth column of $\mathbf{B}_0$:

$$\mathbf{B}_0 = \left[\ \underline{b}_{10} \mid \cdots \mid \underline{b}_{m0}\ \right] \tag{E.82}$$

Following Gutman[8], assume that there exists a positive definite matrix $\mathbf{P} = \mathbf{P}^T > 0$ such that in the set $\left\{\underline{x} \mid \underline{x}^T \left(\mathbf{A}^T\mathbf{P} + \mathbf{PA}\right) \underline{x} > 0\right\}$ the following relation holds:

$$\begin{bmatrix} (\underline{b}_{10} + \mathbf{B}_1\, \underline{x})^T\ \mathbf{P}\, \underline{x} \\ \vdots \\ (\underline{b}_{m0} + \mathbf{B}_m\, \underline{x})^T\ \mathbf{P}\, \underline{x} \end{bmatrix} \neq \mathbf{0} \tag{E.83}$$

Then there exists an $\alpha > 0$ such that following control will stabilize the bilinear system (E.81):

$$u_j = -\alpha\ \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)^T\ \mathbf{P}\, \underline{x} \qquad j = 1, \cdots, m \tag{E.84}$$

Thanks to this control, the closed loop reads:

$$\begin{aligned} \dot{\underline{x}} &= \dot{\underline{x}} = \mathbf{A}\, \underline{x} + \mathbf{B}_0\, \underline{u} + \left(\textstyle\sum_{j=1}^m u_j \mathbf{B}_j\right)\ \underline{x} \\ &= \mathbf{A}\, \underline{x} + \textstyle\sum_{j=1}^m u_j\, \underline{b}_{j0} + \left(\textstyle\sum_{j=1}^m u_j \mathbf{B}_j\right)\ \underline{x} \\ &= \mathbf{A}\, \underline{x} + \textstyle\sum_{j=1}^m \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)\ u_j \\ &= \mathbf{A}\, \underline{x} - \alpha\ \textstyle\sum_{j=1}^m \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)\ \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)^T\ \mathbf{P}\, \underline{x} \end{aligned} \tag{E.85}$$

Then the following candidate Lyapunov function $V\left(\underline{x}\right)$ where $\mathbf{P}$ is a positive definite matrix is introduced:

$$V\left(\underline{x}\right) = \underline{x}^T \mathbf{P} \underline{x} \tag{E.86}$$

Then the time derivative of $V\left(\underline{x}\right)$ reads:

$$\begin{aligned} \dot{V}\left(\underline{x}\right) &= \dot{\underline{x}}^T \mathbf{P} \underline{x} + \underline{x}^T \mathbf{P} \dot{\underline{x}} \\ &= \left(\mathbf{A}\, \underline{x} - \alpha\ \textstyle\sum_{j=1}^m \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)\ \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)^T\ \mathbf{P}\, \underline{x}\right)^T \mathbf{P} \underline{x} \\ &\quad + \underline{x}^T \mathbf{P} \left(\mathbf{A}\, \underline{x} - \alpha\ \textstyle\sum_{j=1}^m \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)\ \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)^T\ \mathbf{P}\, \underline{x}\right) \\ &= \underline{x}^T \left(\mathbf{A}^T\mathbf{P} + \mathbf{PA}\right) \underline{x} \\ &\quad - 2\, \alpha\, \underline{x}^T\, \mathbf{P}\ \left(\textstyle\sum_{j=1}^m \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)\ \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)^T\right)\ \mathbf{P}\, \underline{x} \\ &= \underline{x}^T \left(\mathbf{A}^T\mathbf{P} + \mathbf{PA}\right) \underline{x} - 2\, \alpha\ \textstyle\sum_{j=1}^m \left(\underline{x}^T\, \mathbf{P}\ \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)\right)^2 \end{aligned} \tag{E.87}$$

Alternatively, we can write:

$$\dot{V}\left(\underline{x}\right) = \underline{x}^T \left(\mathbf{A}^T\mathbf{P} + \mathbf{PA} - \mathbf{P}\, \mathbf{W}(\underline{x})\, \mathbf{P}\right) \underline{x} \tag{E.88}$$

where:

$$\mathbf{W}(\underline{x}) = 2\, \alpha\ \left(\sum_{j=1}^m \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)\ \left(\underline{b}_{j0} + \mathbf{B}_j\, \underline{x}\right)^T\right) \tag{E.89}$$

[8] P.-O. Gutman, Stabilizing controllers for bilinear systems, IEEE Transactions on Automatic Control, 1981, Volume: 26, Issue: 4

E.6.3 Optimal control

We consider the following the following time invariant single input bilinear system where state vector $\underline{x} \in \mathbb{R}^n$, control $\underline{u} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{B}_i \in \mathbb{R}^{n\times 1}$:

$$\dot{\underline{x}} = \mathbf{A}\,\underline{x} + \mathbf{B}_0\,\underline{u} + \left(\sum_{i=1}^{n} x_i \mathbf{B}_i\right)\,\underline{u} \tag{E.90}$$

We also consider the following cost function to be minimized where $\mathbf{Q}$ is a semi-definite matrix and $\mathbf{R}$ a positive definite matrix:

$$J(\underline{u}) = \frac{1}{2}\int_0^{\infty} \left(\underline{x}^T\mathbf{Q}\underline{x} + \underline{u}^T\mathbf{R}\underline{u}\right)\,dt \tag{E.91}$$

Denoting $\underline{\lambda}$ the costate variable, the Hamiltonian H related to this optimal control problem reads:

$$H = \frac{1}{2}\left(\underline{x}^T\mathbf{Q}\underline{x} + \underline{u}^T\mathbf{R}\underline{u}\right) + \underline{\lambda}^T\left(\mathbf{A}\,\underline{x} + \mathbf{B}_0\,\underline{u} + \left(\sum_{i=1}^{n} x_i\mathbf{B}_i\right)\,\underline{u}\right) \tag{E.92}$$

The optimal control is determined by necessary optimality condition $\frac{\partial H}{\partial \underline{u}} = \underline{0}$. We get:

$$\frac{\partial H}{\partial \underline{u}} = \underline{0} \Rightarrow \mathbf{R}\underline{u} + \mathbf{B}_0^T\,\underline{\lambda} + \left(\sum_{i=1}^{n} x_i\,\mathbf{B}_i^T\right)\,\underline{\lambda} = \underline{0} \tag{E.93}$$

Thus we get:

$$\begin{aligned} \underline{u} &= -\mathbf{R}^{-1}\left(\mathbf{B}_0^T + \textstyle\sum_{i=1}^{n} x_i\,\mathbf{B}_i^T\right)\,\underline{\lambda} \\ &= -\mathbf{R}^{-1}\left(\mathbf{B}_0 + \textstyle\sum_{i=1}^{n} x_i\,\mathbf{B}_i\right)^T\,\underline{\lambda} \end{aligned} \tag{E.94}$$

Furthermore costate variable $\underline{\lambda}$ satisfies $\frac{\partial H}{\partial \underline{x}} = -\dot{\underline{\lambda}}$. We get:

$$\begin{aligned} \tfrac{\partial H}{\partial \underline{x}} &= -\dot{\underline{\lambda}} \Rightarrow \mathbf{Q}\underline{x} + \left(\mathbf{A}^T + \begin{bmatrix} \mathbf{B}_1 & \cdots & \mathbf{B}_n \end{bmatrix}^T\,\underline{u}\right)\,\underline{\lambda} = -\dot{\underline{\lambda}} \\ \Leftrightarrow \dot{\underline{\lambda}} &= -\mathbf{Q}\underline{x} - \left(\mathbf{A}^T + \begin{bmatrix} \mathbf{B}_1 & \cdots & \mathbf{B}_n \end{bmatrix}^T\,\underline{u}\right)\,\underline{\lambda} \\ &= -\mathbf{Q}\underline{x} - \mathbf{A}^T\underline{\lambda} - \underline{u}\begin{bmatrix} \mathbf{B}_1 & \cdots & \mathbf{B}_n \end{bmatrix}^T\,\underline{\lambda} \end{aligned} \tag{E.95}$$

Control $\underline{u}$ in (E.94) is then inserted into state equation (E.90) and costate equation (E.95). In order to stay in close proximity to the Riccati approach in the linear-quadratic optimization, those equations are written in the same form as in the linear case, namely[9]:

$$\begin{cases} \dot{\underline{x}} = \widetilde{\mathbf{A}}\,\underline{x} - \widetilde{\mathbf{B}}\,\mathbf{R}^{-1}\,\widetilde{\mathbf{B}}^T\,\underline{\lambda} \\ \dot{\underline{\lambda}} = -\widetilde{\mathbf{Q}}\,\underline{x} - \widetilde{\mathbf{A}}\,\underline{\lambda} \end{cases} \tag{E.96}$$

[9] E. Hofer and B. Tibken, An iterative method for the finite-time bilinear quadratic control problem, J. Optim. Theory Applications, vol. 57, pp. 41 1-427, 1988

where:

$$\begin{cases} \widetilde{\mathbf{A}}_{ij} = \mathbf{A}_{ij} - \frac{1}{2}\left[\left(\mathbf{B}_j\mathbf{R}^{-1}\mathbf{B}_0^T + \mathbf{B}_0\mathbf{R}^{-1}\mathbf{B}_j^T\right)\underline{\lambda}\right]_i & i,j = 1,\cdots,n \\ \widetilde{\mathbf{Q}}_{ij} = \mathbf{Q}_{ij} - \frac{1}{2}\underline{\lambda}^T\left(\mathbf{B}_i\mathbf{R}^{-1}\mathbf{B}_j^T + \mathbf{B}_j\mathbf{R}^{-1}\mathbf{B}_i^T\right)\underline{\lambda} & i,j = 1,\cdots,n \\ \widetilde{\mathbf{B}}\,\mathbf{R}^{-1}\,\widetilde{\mathbf{B}}^T = \left(\mathbf{B}_0 + \left(\sum_{i=1}^n x_i\mathbf{B}_i\right)\right)\mathbf{R}^{-1}\left(\mathbf{B}_0 + \left(\sum_{i=1}^n x_i\mathbf{B}_i\right)\right)^T & \\ \qquad -\frac{1}{2}\left(\left(\sum_{i=1}^n x_i\mathbf{B}_i\right)\mathbf{R}^{-1}\mathbf{B}_0^T + \mathbf{B}_0\mathbf{R}^{-1}\left(\sum_{i=1}^n x_i\mathbf{B}_i\right)^T\right) & \end{cases} \tag{E.97}$$

Then an iteration procedure in close proximity to the Riccati approach can be used to get the control[9].

Appendix F

Introduction to fractional systems

F.1 Pre-filtering

A prefilter $C_{pf}(s)$ is a controller which is situated outside the feedback loop as shown in Figure F.1.

What is the purpose of the prefilter ? Once the controller $C(s)$ is designed the poles of the feedback loop transfer function $\frac{Y(s)}{R_{pf}(s)}$ are set. Nevertheless the numerator of this transfer function is not mastered and the zeros of $\frac{Y(s)}{R_{pf}(s)}$ may cause undesirable overshoots in the transient response of the closed loop system. The purpose of the prefilter $C_{pf}(s)$ is to reduce or eliminate such overshoots in the closed loop system.

Let $N_{cl}(s)$ be the numerator of transfer function $\frac{Y(s)}{R_{pf}(s)}$ and $D_{cl}(s)$ its denominator:

$$\frac{Y(s)}{R_{pf}(s)} = \frac{N_{cl}(s)}{D_{cl}(s)} \tag{F.1}$$

The prefilter $C_{pf}(s)$ is then designed such that its poles cancel the zeros of the closed loop system, that are the roots of $N_{cl}(s)$. Furthermore the numerator of the prefilter is usually set to be a constant K_{pf} such that the transfer function of the full system reads:

$$\frac{Y(s)}{R(s)} = \frac{K_{pf}}{D_{cl}(s)} \tag{F.2}$$

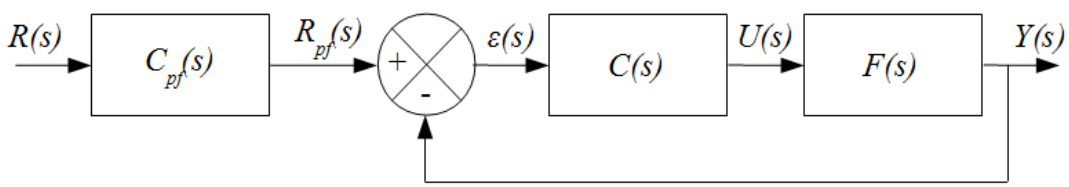

Figure F.1: Feedback loop design with prefilter

As a consequence the transfer function of the prefilter reads:

$$C_{pf}(s) = \frac{K_{pf}}{N_{cl}(s)} \tag{F.3}$$

Usually constant K_{pf} is set such that the static gain of $\frac{Y(s)}{R(s)}$ is unitary, meaning that the position error is zero:

$$\left.\frac{Y(s)}{R(s)}\right|_{s=0} = 1 \Rightarrow K_{pf} = D_{cl}(0) \tag{F.4}$$

F.2 Design steps

The general scheme for the controlled system is provided in Figure F.1 where $C_{pf}(s)$ is the transfer function of the prefilter.

The design philosophy is to set the transfer function $C(s)$ of the controller and the transfer function $C_{pf}(s)$ of the prefilter in order to force the transfer function of the full system to have the following expression where K_{pf} is a constant gain and $D_{cl}(s)$ a polynomial formed with the desired closed loop poles:

$$\frac{Y(s)}{R(s)} = \frac{K_{pf}}{D_{cl}(s)} \tag{F.5}$$

The design steps of the control loop are the following:

- Design the controller $C(s)$ such that transfer function of feedback loop without prefiltering ($C_{pf}(s) = 1$) has the desired denominator $D_{cl}(s)$. In other words controller $C(s)$ is used to set the poles of the controlled system.

- Design the prefilter $C_{pf}(s)$ such that transfer function of the full system does not have any zero:

$$\frac{Y(s)}{R(s)} = \frac{K_{pf}}{D_{cl}(s)} \tag{F.6}$$

  In other words prefilter $C_{pf}(s)$ is used to shape the numerator of the transfer function of the controlled system.

Example F.1. *Consider a plant with the following transfer function:*

$$F(s) = \frac{1}{s\,(s-2)} \tag{F.7}$$

Obviously the plant is not stable, indeed there is one pole at $+2$. *In order to stabilize the plant we decide to use the following PD controller (we do not use an integral action because the plant* $F(s)$ *has already an integral term):*

$$C(s) = K_p + K_d\, s \tag{F.8}$$

Furthermore we set the targeted transfer function of the controlled system as follows (see Figure F.1):

$$\frac{Y(s)}{R(s)} = \frac{2}{s^2 + 5\,s + 2} \tag{F.9}$$

The first step of the design is to set the values K_p and K_d of the PD controller such that the denominator of the targeted transfer function is achieved. Transfer function $\frac{Y(s)}{R_{pf}(s)}$ (no prefilter here) reads:

$$\begin{aligned} \frac{Y(s)}{R_{pf}(s)} &= \frac{C(s)F(s)}{1+C(s)F(s)} \\ &= \frac{K_p+K_d\,s}{K_p+K_d\,s+s\,(s-2)} \\ &= \frac{K_p+K_d\,s}{s^2+s\,(K_d-2)+K_p} \end{aligned} \tag{F.10}$$

The actual denominator will be equal to the targeted denominator as soon as K_p and K_d are set as follows:

$$s^2 + s\,(K_d - 2) + K_p = s^2 + 5\,s + 2 \Rightarrow \begin{cases} K_d = 7 \\ K_p = 2 \end{cases} \tag{F.11}$$

Thus transfer function $\frac{Y(s)}{R_{pf}(s)}$ (no prefilter here) reads:

$$\frac{Y(s)}{R_{pf}(s)} = \frac{K_p + K_d\,s}{s^2 + s\,(K_d - 2) + K_p} = \frac{2 + 7\,s}{s^2 + 5\,s + 2} \tag{F.12}$$

Taking now into account prefilter $C_{pf}(s)$ transfer function $\frac{Y(s)}{R(s)}$ reads:

$$\frac{Y(s)}{R(s)} = \frac{R_{pf}(s)}{R(s)}\frac{Y(s)}{R_{pf}(s)} = C_{pf}(s)\frac{2 + 7\,s}{s^2 + 5\,s + 2} \tag{F.13}$$

Thus transfer function of the controlled system will read $\frac{Y(s)}{R(s)} = \frac{2}{s^2+5\,s+2}$ as soon as prefilter $C_{pf}(s)$ is set as follows:

$$\frac{Y(s)}{R(s)} = \frac{2}{s^2 + 5\,s + 2} \Rightarrow C_{pf}(s) = \frac{2}{2 + 7\,s} \tag{F.14}$$

■

F.3 Pre-filtering design for non-minimum phase feedback loop

Sometimes the numerator of the feedback loop transfer function $\frac{Y(s)}{R_{pf}(s)}$ has zeros with positive real part. Such transfer functions with zeros in the right half plane in the complex plane, that is with positive real part, are called non-minimum phase transfer functions. As far as the denominator of the prefilter is set to the numerator of the feedback loop transfer function this leads to an unstable prefilter, which is not acceptable.

Suppose that feedback loop transfer function $G(s) = \frac{Y(s)}{R_{pf}(s)}$ has a positive real zero of order one at $s = z > 0$, that is $G(z) = 0$ and $\left.\frac{dG(s)}{ds}\right|_{s=z} \neq 0$. Such a transfer function can be decomposed as follows[1] where $G_{mp}(s)$ is a minimum phase transfer function:

$$G(s) = \left(1 - \frac{s}{z}\right) G_{mp}(s) \tag{F.15}$$

Using remarkable identity $1 - x^2 = (1 - x)(1 + x)$, the term $1 - \frac{s}{z}$ in the preceding equation can be expanded using fractional powers of s as follows:

$$\begin{aligned} 1 - \tfrac{s}{z} &= \left(1 - \left(\tfrac{s}{z}\right)^{1/2}\right)\left(1 + \left(\tfrac{s}{z}\right)^{1/2}\right) \\ &= \left(1 - \left(\tfrac{s}{z}\right)^{1/4}\right)\left(1 + \left(\tfrac{s}{z}\right)^{1/4}\right)\left(1 + \left(\tfrac{s}{z}\right)^{1/2}\right) \\ &= \left(1 - \left(\tfrac{s}{z}\right)^{1/8}\right)\left(1 + \left(\tfrac{s}{z}\right)^{1/8}\right)\left(1 + \left(\tfrac{s}{z}\right)^{1/4}\right)\left(1 + \left(\tfrac{s}{z}\right)^{1/2}\right) \\ &= \cdots \end{aligned} \tag{F.16}$$

That is:

$$1 - \frac{s}{z} = \left(1 - \left(\frac{s}{z}\right)^{1/M}\right) \prod_{k=0}^{\log_2(M/2)} \left(1 + \left(\frac{s}{z}\right)^{2^k/M}\right) \tag{F.17}$$

where $\log_2(M/2)$ is the base 2 logarithm of $M/2$ and M is any number multiple of 2.

The positive real zero z can then be partially compensated through the term $D_M(s) = \prod_{k=0}^{\log_2(M/2)} \left(1 + \left(\frac{s}{z}\right)^{2^k/M}\right)$ which will appear in the denominator of the transfer function of the prefilter. Indeed it can be shown[1] that $G_M(s) = \left(1 - \left(\frac{s}{z}\right)^{1/M}\right) G_{mp}(s)$ has a *weaker* non-minimum phase zero at $s = z$ than $G(s) = \left(1 - \frac{s}{z}\right) G_{mp}(s)$.

The next step consists in approximating the state space fractional system with the following transfer function $G_f(s)$ which will appear in the prefilter:

$$G_f(s) = \frac{1}{\prod_{k=0}^{\log_2(M/2)} \left(1 + \left(\frac{s}{z}\right)^{2^k/M}\right)} \tag{F.18}$$

F.4 Mathematical Definitions of Fractional Integral and Derivative

The most encountered definition of the fractional integral is called Riemann-Liouville (RL) fractional integral and is defined as follows where $q \in \mathbb{R}^+$ is the order of integral and $\Gamma(q)$ is the Gamma function[2]:

$$\begin{cases} \mathcal{D}^{-q} x(t) := \frac{1}{\Gamma(q)} \int_{t_0}^{t} (t - \tau)^{q-1} x(\tau)\, d\tau \\ \Gamma(q) = \int_0^\infty e^{-t}\, t^{q-1}\, dt \end{cases} \tag{F.19}$$

[1] Farshad Merrikh-Bayat, Practical and efficient method for fractional-order unstable pole-zero cancellation in linear feedback systems, https://arxiv.org/abs/1207.6962

[2] Fujio Ikeda & Shigehiro Toyama (2009) A Frequency Domain Approach for Robust Control Design by Fractional Calculus, SICE Journal of Control, Measurement, and System Integration, 2:3, 162-167, DOI: 10.9746/jcmsi.2.162

The above expression is used to define the fractional derivative as follows, where $m \in \mathbb{N}$ satisfies $m-1 < q \leq m$:

$$\mathcal{D}^q x(t) := \frac{d^m}{dt^m}\left(\mathcal{D}^{-(m-q)}x(t)\right) = \frac{1}{\Gamma(m-q)}\frac{d^m}{dt^m}\int_{t_0}^{t}(t-\tau)^{m-q-1}\,x(\tau)\,d\tau \quad \text{(F.20)}$$

The Laplace transform of the Riemann-Liouville (RL) fractional derivative/integral under zero initial conditions for the order q is given by:

$$X(s) = \mathcal{L}\left[x(t)\right] \Rightarrow \mathcal{L}\left[\mathcal{D}^{\pm q}x(t)\right] = s^{\pm q}X(s) \quad \text{(F.21)}$$

F.5 CRONE approximation of fractional derivative

CRONE (*Commande Robuste d'Ordre Non-Entier*: this is the French abbreviation for Non Integer Order Robust Control which was introduced by A. Oustaloup[3]) is a method which enables to approximate the infinite dimension fractional filter s^{α_i}, $\alpha_i \in (0,1)$, by the following finite dimension rational transfer function $G_{\alpha_i}(s)$[4]:

$$s^{\alpha_i} \approx G_{\alpha_i}(s) = \left(\frac{1}{\omega_B}\right)^{\alpha_i}\frac{\prod_{i=-N}^{i=+N}\left(1+\frac{s}{\omega_{z,i}}\right)}{\prod_{i=-N}^{i=+N}\left(1+\frac{s}{\omega_{p,i}}\right)} \quad \text{(F.22)}$$

In order to obtain a good accuracy in a frequency range $(\omega_{min}, \omega_{max})$, poles and zeros are distributed in a broader frequency range (ω_A, ω_B) defined as a function of $(\omega_{min}, \omega_{max})$ and an adjustment coefficient σ which is often chosen to $\sigma = 10$:

$$\begin{cases} \omega_B = \sigma\,\omega_{max} \\ \omega_A = \frac{\omega_{min}}{\sigma} \\ \eta = \left(\frac{\omega_B}{\omega_A}\right)^{(1-\alpha_i)/(2N+1)} \\ \delta = \left(\frac{\omega_B}{\omega_A}\right)^{\alpha_i/(2N+1)} \\ \omega_{z,-N} = \sqrt{\eta}\,\omega_A \\ \omega_{p,i} = \delta\,\omega_{z,i} \qquad i = -N, \cdots, N \\ \omega_{z,i+1} = \eta\,\omega_{p,i} \qquad i = -N, \cdots, N-1 \end{cases} \quad \text{(F.23)}$$

We recall that as far as the approximated transfer function $G_{\alpha_i}(s)$ of s^{α_i} has distinct real poles λ_i, its partial fraction expansion reads:

$$\begin{aligned} s^{\alpha_i} \approx G_{\alpha_i}(s) &= \tfrac{N(s)}{D(s)} + d \\ &= \tfrac{N(s)}{(s-\lambda_1)(s-\lambda_2)\cdots(s-\lambda_n)} + d \\ &= \tfrac{r_1}{s-\lambda_1} + \tfrac{r_2}{s-\lambda_2} + \cdots + \tfrac{r_n}{s-\lambda_n} + d \end{aligned} \quad \text{(F.24)}$$

Number r_i is called the residue of transfer function $G_{\alpha_i}(s)$ in λ_i. When the multiplicity of the pole (or eigenvalue) λ_i is 1 we have seen that residue r_i can be obtained thanks to the following formula:

$$r_i = (s-\lambda_i)G_{\alpha_i}(s)\big|_{s=\lambda_i} \quad \text{(F.25)}$$

[3] A. Oustaloup - La commande CRONE, Hermes, 1991

[4] Mansouri Rachid, Bettayeb Maamar & Djennoune Said, Comparison between two approximation methods of state space fractional systems Signal Processing 91 (2011) 461–469

Now we define constants b_i and c_i such that the product $b_i c_i$ is equal to r_i:

$$r_i = b_i\, c_i \tag{F.26}$$

This leads to the following state-space representation, which is called the diagonal (Jordan) or modal form, corresponding to the approximation of s^{α_i}:

$$s^{\alpha_i} \approx G_{\alpha_i}(s) : \left\{ \begin{array}{l} \dot{\underline{z}}(t) = \mathbf{A}_{\alpha_i}\, \underline{z}(t) + \mathbf{B}_{\alpha_i}\, u(t) \\ y(t) = \mathcal{D}^{\alpha_i} u(t) \approx \mathbf{C}_{\alpha_i}\, \underline{z}(t) + \mathbf{D}_{\alpha_i}\, u(t) \end{array} \right. \tag{F.27}$$

where:

$$\left\{ \begin{array}{l} \mathbf{A}_{\alpha_i} = \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \\ \\ \mathbf{B}_{\alpha_i} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\ \\ \mathbf{C}_{\alpha_i} = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix} \\ \\ \mathbf{D}_{\alpha_i} = d \end{array} \right. \tag{F.28}$$

Similarly the Laplace transform of the fractional integration operator $\mathcal{I}^{\alpha_i}$ is $s^{-\alpha_i}$. The approximation of the fractional integration operation $s^{-\alpha_i}$ can be obtained by exploiting the following equality[5]:

$$s^{-\alpha_i} = \frac{1}{s}\, s^{1-\alpha_i} \tag{F.29}$$

Because $0 \leq 1-\alpha_i \leq 1$ as soon as $\alpha_i \in (0,1)$, fractional integration operation $s^{1-\alpha_i}$ can be approximated by a transfer function similar to (F.22). Then the finite dimension rational model is multiplied by $\frac{1}{s}$ which leads to a strictly proper approximation $\frac{1}{s}\, s^{1-\alpha_i}$ of the fractional order integration operation $s^{-\alpha_i}$:

$$s^{-\alpha_i} = \frac{1}{s}\, s^{1-\alpha_i} \approx \frac{1}{s}\, G_{1-\alpha_i}(s) \tag{F.30}$$

Because $\frac{1}{s}\, s^{1-\alpha_i}$ is a strictly proper transfer function, matrix $\mathbf{D}_{-\alpha_i}$ is null. The state space representation corresponding to the approximation of $s^{-\alpha_i}$ reads:

$$s^{-\alpha_i} \approx \frac{1}{s}\, G_{1-\alpha_i}(s) : \left\{ \begin{array}{l} \dot{\underline{z}}(t) = \mathbf{A}_{-\alpha_i}\, \underline{z}(t) + \mathbf{B}_{-\alpha_i}\, u(t) \\ y(t) = \mathcal{I}^{\alpha_i} u(t) \approx \mathbf{C}_{-\alpha_i}\, \underline{z}(t) \end{array} \right. \tag{F.31}$$

[5] Thierry Poinot & Jean-Claude Trigeassou, A method for modelling and simulation of fractional systems, Signal Processing 83 (2003) 2319-2333

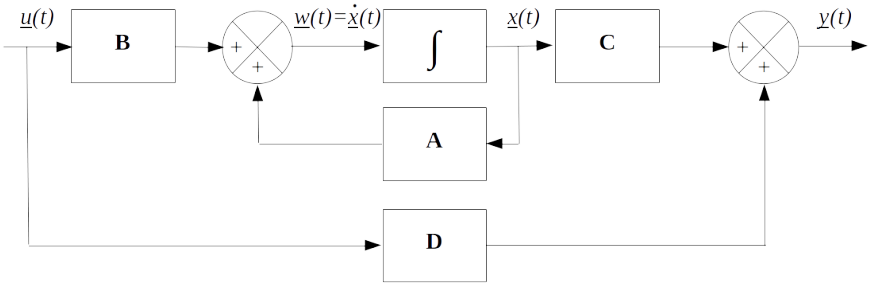

Figure F.2: Block diagram of a state-space representation

F.6 State space fractional systems

Usually the state space model of an integer linear time invariant system is the following:

$$\begin{cases} \dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{F.32}$$

Where:

- $\underline{x}(t)$ is the state vector, which is of dimension n. The number n of the state vector components is called the order of the system;
- $\underline{u}(t)$ is the input of the system;
- $\underline{y}(t)$ is the output of the system.

State vector $\underline{x}(t)$ can be defined as a set of variables such that their knowledge at the initial time $t_0 = 0$, together with knowledge of system inputs $\underline{u}(t)$ at $t \geq 0$ are sufficient to predict the future system state and output $\underline{y}(t)$ for all time $t > 0$.

Both equations in (F.32) have a name:

- Equation $\dot{\underline{x}}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t)$ is named as the state equation;
- Equation $\underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t)$ is named as the output equation.

The state equation and the output equation both constitute the state-space representation of the system.

The block diagram corresponding to the state-space representation (F.32) is shown in Figure F.2.

The corresponding transfer function of the model is given by:

$$\mathbf{G}(s) = \mathbf{C}\left(s\mathbb{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} \tag{F.33}$$

State space representation (F.32) can be extended to fractional case as follows[6]:

$$\begin{cases} \mathcal{D}^{\underline{\alpha}}\underline{x}(t) = \mathbf{A}\underline{x}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\underline{x}(t) + \mathbf{D}\underline{u}(t) \end{cases} \tag{F.34}$$

[6]W. Krajewski & U. Viaro, On the rational approximation of fractional order systems, 2011 16th International Conference on Methods & Models in Automation & Robotics, Miedzyzdroje, 2011, pp. 132-136. doi: 10.1109/MMAR.2011.6031331

where:

$$\mathcal{D}^{\underline{\alpha}}\underline{x}(t) = \left[\begin{array}{ccc} \mathcal{D}^{\alpha_1}x_1(t) & \cdots & \mathcal{D}^{\alpha_n}x_n(t) \end{array}\right]^T \tag{F.35}$$

Denoting by $[\alpha_k]$ the integer part of α_k the Laplace transform of the fractional Caputo derivative $\mathcal{D}^{\alpha_k}x_k(t)$ is[6]:

$$\mathcal{L}\left[\mathcal{D}^{\alpha_k}x_k(t)\right] = s^{\alpha_k}\mathcal{L}\left[x_k(t)\right] - \sum_{i=0}^{[\alpha_k]-1} s^{\alpha_k-i-1}\left.\frac{d^i x_k(t)}{dt^i}\right|_{t=0} \tag{F.36}$$

If all fractional orders are multiples of the same real number $\alpha \in (0,1)$ (commensurate fractional–order systems), operator $\mathcal{D}^{\underline{\alpha}}\underline{x}(t)$ simplifies as follows:

$$\mathcal{D}^{\underline{\alpha}}\underline{x}(t) = \left[\begin{array}{ccc} \mathcal{D}^{\alpha}x_1(t) & \cdots & \mathcal{D}^{\alpha}x_n(t) \end{array}\right]^T \tag{F.37}$$

Example F.2. *The following example presents a fractional state space representation and its corresponding transfer function*[4]*:*

$$\begin{cases} \begin{bmatrix} \mathcal{D}^{1.56}\underline{x}_1(t) \\ \mathcal{D}^{1.13}\underline{x}_2(t) \\ \mathcal{D}^{0.77}\underline{x}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -4 & -20 & -10 \end{bmatrix}\underline{x}(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\underline{u}(t) \\ \underline{y}(t) = \begin{bmatrix} 4 & 1 & 0 \end{bmatrix}\underline{x}(t) \end{cases} \tag{F.38}$$

The corresponding transfer function is the following:

$$G(s) = \frac{s^{1.56}+4}{s^{3.46}+10\,s^{2.69}+20\,s^{1.56}+4} \tag{F.39}$$

Denoting by $y(t)$ the output of the system, it is worth noticing that the components of the state vector are the following where $y^{(i)}(t)$ indicates the ith time derivative of $y(t)$:

$$\underline{x}(t) = \begin{bmatrix} y(t) \\ y^{(1.56)}(t) \\ y^{(2.69)}(t) \end{bmatrix} \tag{F.40}$$

Furthermore $\mathcal{D}^{1.56}$ corresponds to the lower fractional derivative which appears in the denominator of transfer function $G(s)$ whereas the others terms, $\mathcal{D}^{1.13}$ and $\mathcal{D}^{0.77}$ namely, are obtained by subtracting the consecutive fractional derivatives which appears in the denominator of transfer function $G(s)$.

Last but not least the fractional exponents which appear in the numerator of $G(s)$ are the same than those which appear in its denominator.

■

F.7 Approximation of fractional systems based on differentiation operator

The approximation of fractional systems based on differentiation operator is obtained by coupling (F.34) and (F.27):

$$\left\{\begin{array}{l}\left\{\begin{array}{l}\mathcal{D}^{\underline{\alpha}}\underline{x}(t)=\mathbf{A}\underline{x}(t)+\mathbf{B}\underline{u}(t)\\ \underline{y}(t)=\mathbf{C}\underline{x}(t)+\mathbf{D}\underline{u}(t)\end{array}\right.\\ \left\{\begin{array}{l}\dot{\underline{z}}(t)=\mathbf{A}_{\underline{\alpha}}\,\underline{z}(t)+\mathbf{B}_{\underline{\alpha}}\,\underline{x}(t)\\ \mathcal{D}^{\underline{\alpha}}\underline{x}(t)\approx\mathbf{C}_{\underline{\alpha}}\,\underline{z}(t)+\mathbf{D}_{\underline{\alpha}}\,\underline{x}(t)\end{array}\right.\end{array}\right. \tag{F.41}$$

Block-diagonal matrices $\mathbf{A}_{\underline{\alpha}}\in\mathbb{R}^{(2N+1)\cdot n\times(2N+1)\cdot n}$, $\mathbf{B}_{\underline{\alpha}}\in\mathbb{R}^{(2N+1)\cdot n\times n}$, $\mathbf{C}_{\underline{\alpha}}\in\mathbb{R}^{n\times(2N+1)\cdot n}$ and $\mathbf{D}_{\underline{\alpha}}\in\mathbb{R}^{n\times n}$ are obtained from (F.27) as follows:

$$\left\{\begin{array}{l}\mathbf{A}_{\underline{\alpha}}=diag\left[\begin{array}{ccc}\mathbf{A}_{\alpha_1} & \cdots & \mathbf{A}_{\alpha_n}\end{array}\right]\\ \mathbf{B}_{\underline{\alpha}}=diag\left[\begin{array}{ccc}\mathbf{B}_{\alpha_1} & \cdots & \mathbf{B}_{\alpha_n}\end{array}\right]\\ \mathbf{C}_{\underline{\alpha}}=diag\left[\begin{array}{ccc}\mathbf{C}_{\alpha_1} & \cdots & \mathbf{C}_{\alpha_n}\end{array}\right]\\ \mathbf{D}_{\underline{\alpha}}=diag\left[\begin{array}{ccc}\mathbf{D}_{\alpha_1} & \cdots & \mathbf{D}_{\alpha_n}\end{array}\right]\end{array}\right. \tag{F.42}$$

Equating the expression of $\mathcal{D}^{\underline{\alpha}}\underline{x}(t)$ in both equations of (F.41) yields to the expression of the state space vector $\underline{x}(t)$ of the fractional model:

$$\begin{array}{rl}\mathbf{A}\,\underline{x}(t)+\mathbf{B}\,\underline{u}(t) & \approx\mathbf{C}_{\underline{\alpha}}\,\underline{z}(t)+\mathbf{D}_{\underline{\alpha}}\,\underline{x}(t)\\ \Leftrightarrow\underline{x}(t) & \approx\left(\mathbf{A}-\mathbf{D}_{\underline{\alpha}}\right)^{-1}\left(\mathbf{C}_{\underline{\alpha}}\,\underline{z}(t)-\mathbf{B}\,\underline{u}(t)\right)\\ & \approx\left(\mathbf{A}-\mathbf{D}_{\underline{\alpha}}\right)^{-1}\mathbf{C}_{\underline{\alpha}}\,\underline{z}(t)+\left(\mathbf{D}_{\underline{\alpha}}-\mathbf{A}\right)^{-1}\mathbf{B}\,\underline{u}(t)\end{array} \tag{F.43}$$

Finally the approximation of fractional system (F.41) based on differentiation operator reads:

$$\begin{array}{l}\left\{\begin{array}{l}\dot{\underline{z}}(t)=\mathbf{A}_D\,\underline{z}(t)+\mathbf{B}_D\,\underline{u}(t)\\ \underline{y}(t)\approx\mathbf{C}_D\underline{z}(t)+\mathbf{D}_D\underline{u}(t)\end{array}\right.\\ \text{where}\left\{\begin{array}{l}\mathbf{A}_D=\mathbf{A}_{\underline{\alpha}}+\mathbf{B}_{\underline{\alpha}}\left(\mathbf{A}-\mathbf{D}_{\underline{\alpha}}\right)^{-1}\mathbf{C}_{\underline{\alpha}}\\ \mathbf{B}_D=\mathbf{B}_{\underline{\alpha}}\left(\mathbf{D}_{\underline{\alpha}}-\mathbf{A}\right)^{-1}\mathbf{B}\\ \mathbf{C}_D=\mathbf{C}\left(\mathbf{A}-\mathbf{D}_{\underline{\alpha}}\right)^{-1}\mathbf{C}_{\underline{\alpha}}\\ \mathbf{D}_D=\mathbf{C}\left(\mathbf{D}_{\underline{\alpha}}-\mathbf{A}\right)^{-1}\mathbf{B}+\mathbf{D}\end{array}\right.\end{array} \tag{F.44}$$

F.8 Approximation of fractional systems based on integration operator

The inspection of Figure F.2 shows that we may also choose the integral operator input $\underline{x}(t)=\int_0^t\underline{w}(\tau)d\tau$ as the *state vector* vector and write the corresponding state space model using the integral function instead of the derivative one. This yields to[4]:

$$\left\{\begin{array}{l}\underline{w}(t)=\mathbf{A}\int_0^t\underline{w}(\tau)d\tau+\mathbf{B}\underline{u}(t)\\ \underline{y}(t)=\mathbf{C}\int_0^t\underline{w}(\tau)d\tau+\mathbf{D}\underline{u}(t)\\ \underline{x}(t)=\int_0^t\underline{w}(\tau)d\tau\end{array}\right. \tag{F.45}$$

The corresponding transfer function of the model then given by

$$\mathbf{G}(s)=\mathbf{C}\frac{\mathbb{I}}{s}\left(\mathbb{I}-\mathbf{A}\frac{\mathbb{I}}{s}\right)^{-1}\mathbf{B}+\mathbf{D} \tag{F.46}$$

Denoting by $\mathcal{I}^{\alpha_i}$ the fractional integration operator, state space representation (F.45) can be extended to fractional case as follows[4]:

$$\begin{cases} \underline{w}(t) = \mathbf{A}\mathcal{I}^{\underline{\alpha}}\underline{w}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\mathcal{I}^{\underline{\alpha}}\underline{w}(t) + \mathbf{D}\underline{u}(t) \\ \underline{x}(t) = \mathcal{I}^{\underline{\alpha}}\underline{w}(t) \end{cases} \tag{F.47}$$

where:

$$\mathcal{I}^{\underline{\alpha}}\underline{w}(t) = \left[\ \mathcal{I}^{\alpha_1}w_1(t) \quad \cdots \quad \mathcal{I}^{\alpha_n}w_n(t)\ \right]^T \tag{F.48}$$

The approximation of fractional systems based on integration operator is obtained by coupling (F.47) and (F.31):

$$\begin{cases} \begin{cases} \underline{w}(t) = \mathbf{A}\mathcal{I}^{\underline{\alpha}}\underline{w}(t) + \mathbf{B}\underline{u}(t) \\ \underline{y}(t) = \mathbf{C}\mathcal{I}^{\underline{\alpha}}\underline{w}(t) + \mathbf{D}\underline{u}(t) \end{cases} \\ \begin{cases} \dot{\underline{z}}(t) = \mathbf{A}_{-\underline{\alpha}}\,\underline{z}(t) + \mathbf{B}_{-\underline{\alpha}}\,\underline{w}(t) \\ \underline{x}(t) = \mathcal{I}^{\underline{\alpha}}\underline{w}(t) \approx \mathbf{C}_{-\underline{\alpha}}\,\underline{z}(t) \end{cases} \end{cases} \tag{F.49}$$

Block-diagonal matrices $\mathbf{A}_{-\underline{\alpha}} \in \mathbb{R}^{(2N+2)\cdot n\times(2N+2)\cdot n}$, $\mathbf{B}_{-\underline{\alpha}} \in \mathbb{R}^{(2N+2)\cdot n\times n}$ and $\mathbf{C}_{-\underline{\alpha}} \in \mathbb{R}^{n\times(2N+2)\cdot n}$ are obtained from (F.31) as follows:

$$\begin{cases} \mathbf{A}_{-\underline{\alpha}} = diag\left[\ \mathbf{A}_{-\alpha_1} \quad \cdots \quad \mathbf{A}_{-\alpha_n}\ \right] \\ \mathbf{B}_{-\underline{\alpha}} = diag\left[\ \mathbf{B}_{-\alpha_1} \quad \cdots \quad \mathbf{B}_{-\alpha_n}\ \right] \\ \mathbf{C}_{-\underline{\alpha}} = diag\left[\ \mathbf{C}_{-\alpha_1} \quad \cdots \quad \mathbf{C}_{-\alpha_n}\ \right] \end{cases} \tag{F.50}$$

Using the expression of $\underline{w}(t)$ provided in the first equation of (F.49) within the expression of $\dot{\underline{z}}(t)$ in the third equation of (F.49) and using the approximation $\mathcal{I}^{\underline{\alpha}}\underline{w}(t) \approx \mathbf{C}_{-\underline{\alpha}}\,\underline{z}(t)$ provided in the fourth equation yields:

$$\begin{aligned} \underline{w}(t) &= \mathbf{A}\mathcal{I}^{\underline{\alpha}}\underline{w}(t) + \mathbf{B}\underline{u}(t) \\ \Rightarrow \dot{\underline{z}}(t) &= \mathbf{A}_{-\underline{\alpha}}\,\underline{z}(t) + \mathbf{B}_{-\underline{\alpha}}\,\underline{w}(t) \\ &= \mathbf{A}_{-\underline{\alpha}}\,\underline{z}(t) + \mathbf{B}_{-\underline{\alpha}}\left(\mathbf{A}\mathcal{I}^{\underline{\alpha}}\underline{w}(t) + \mathbf{B}\underline{u}(t)\right) \\ &= \mathbf{A}_{-\underline{\alpha}}\,\underline{z}(t) + \mathbf{B}_{-\underline{\alpha}}\mathbf{A}\mathcal{I}^{\underline{\alpha}}\underline{w}(t) + \mathbf{B}_{-\underline{\alpha}}\mathbf{B}\underline{u}(t) \\ &\approx \mathbf{A}_{-\underline{\alpha}}\,\underline{z}(t) + \mathbf{B}_{-\underline{\alpha}}\mathbf{A}\mathbf{C}_{-\underline{\alpha}}\,\underline{z}(t) + \mathbf{B}_{-\underline{\alpha}}\mathbf{B}\underline{u}(t) \\ &= \left(\mathbf{A}_{-\underline{\alpha}} + \mathbf{B}_{-\underline{\alpha}}\mathbf{A}\mathbf{C}_{-\underline{\alpha}}\right)\underline{z}(t) + \mathbf{B}_{-\underline{\alpha}}\mathbf{B}\underline{u}(t) \end{aligned} \tag{F.51}$$

Finally the approximation of fractional system (F.49) based on integration operator reads:

$$\begin{cases} \dot{\underline{z}}(t) = \mathbf{A}_I\,\underline{z}(t) + \mathbf{B}_I\,\underline{u}(t) \\ \underline{y}(t) \approx \mathbf{C}_I\underline{z}(t) + \mathbf{D}_I\underline{u}(t) \end{cases}$$
$$\text{where} \begin{cases} \mathbf{A}_I = \mathbf{A}_{-\underline{\alpha}} + \mathbf{B}_{-\underline{\alpha}}\mathbf{A}\mathbf{C}_{-\underline{\alpha}} \\ \mathbf{B}_I = \mathbf{B}_{-\underline{\alpha}}\mathbf{B} \\ \mathbf{C}_I = \mathbf{C}\,\mathbf{C}_{-\underline{\alpha}} \\ \mathbf{D}_I = \mathbf{D} \end{cases} \tag{F.52}$$

Example F.3. *Coming back to Figure F.1, we consider the following non-minimum phase transfer function:*

$$G(s) = \frac{Y(s)}{R_{pf}(s)} = \frac{10s - 1}{s^2 + 1.4s + 1} \tag{F.53}$$

It is clear that in order to obtain no static error, that is $\frac{Y(s)}{R(s)} = \frac{1}{s^2+1.4s+1}$, *we shall choose the prefilter* $C_{pf}(s)$ *as follows:*

$$\frac{Y(s)}{R(s)} = \frac{1}{s^2+1.4s+1} = G(s)C_{pf}(s) \Rightarrow C_{pf}(s) = \frac{1}{10s-1} = \frac{-1}{1-\frac{s}{0.1}} \tag{F.54}$$

Obviously $C_{pf}(s)$ *is not a stable system and this prefilter cannot be implemented.*

Alternatively we can write $G(s)$ *as follows where* $G_{mp}(s)$ *is a minimum phase transfer function:*

$$G(s) = \left(1-\frac{s}{z}\right) G_{mp}(s) = \left(1-\frac{s}{0.1}\right) \frac{-1}{s^2+1.4s+1} \tag{F.55}$$

Then, choosing for example $M = 4$ *we write:*

$$\begin{aligned} 1-\tfrac{s}{0.1} &= \left(1-\left(\tfrac{s}{0.1}\right)^{1/M}\right) \prod_{k=0}^{\log_2(M/2)} \left(1+\left(\tfrac{s}{0.1}\right)^{2^k/M}\right) \\ &= \left(1-\left(\tfrac{s}{0.1}\right)^{0.25}\right)\left(1+\left(\tfrac{s}{0.1}\right)^{0.25}\right)\left(1+\left(\tfrac{s}{0.1}\right)^{0.5}\right) \end{aligned} \tag{F.56}$$

From this decomposition prefilter $C_{pf}(s)$ *now reads:*

$$\begin{aligned} C_{pf}(s) &= \frac{-1}{\left(1+\left(\frac{s}{0.1}\right)^{0.25}\right)\left(1+\left(\frac{s}{0.1}\right)^{0.5}\right)} \\ &= \frac{-(0.1)^{0.25}\,(0.1)^{0.5}}{(s^{0.25}+(0.1)^{0.25})(s^{0.5}+(0.1)^{0.5})} \\ &= \frac{-(0.1)^{0.75}}{s^{0.75}+(0.1)^{0.25}s^{0.5}+(0.1)^{0.5}s^{0.25}+(0.1)^{0.75}} \end{aligned} \tag{F.57}$$

The fractional state space representation corresponding to this transfer function reads:

$$\begin{cases} \begin{bmatrix} \mathcal{D}^{0.25}\underline{x}_1(t) \\ \mathcal{D}^{0.25}\underline{x}_2(t) \\ \mathcal{D}^{0.25}\underline{x}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -(0.1)^{0.75} & -(0.1)^{0.5} & -(0.1)^{0.25} \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \underline{u}(t) \\ \underline{y}(t) = \begin{bmatrix} -(0.1)^{0.75} & 0 & 0 \end{bmatrix} \underline{x}(t) \end{cases} \tag{F.58}$$

Figure F.3 shows Bode plots of the two approximating methods for fractional prefilter $C_{pf}(s)$ *with the following setting:*

$$\begin{cases} N = 4 \\ \omega_{min} = 10^{-2} \\ \omega_{max} = 10^{2} \end{cases} \tag{F.59}$$

Figure F.4 shows the step response of the plant with the rational prefilter $C_{pf}(s)$*: it can be seen that the non-minimum phase effect has been reduced but the time response has been highly increased compared with the result obtained with a static prefilter* $C_{pf}(s) = -1$ *which leads to* $\frac{Y(s)}{R(s)} = -\frac{10s-1}{s^2+1.4s+1} = -G(s)$

■

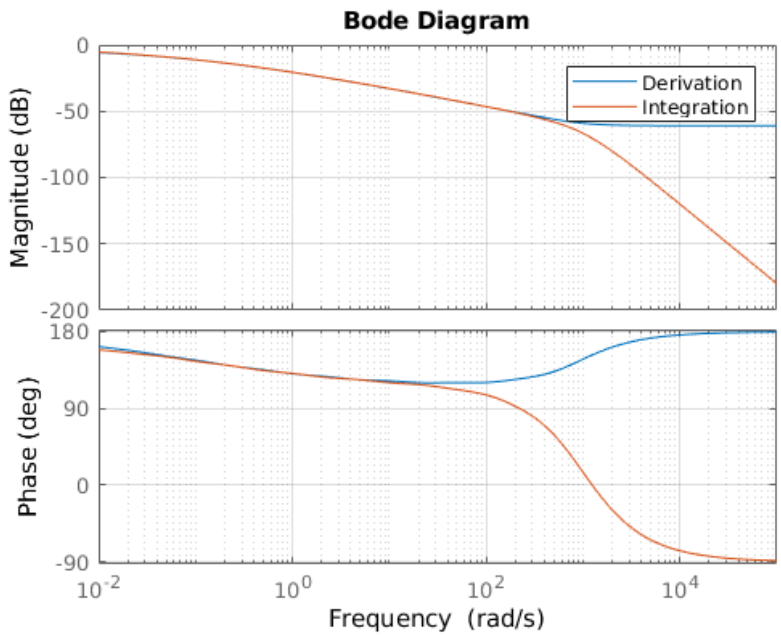

Figure F.3: Bode plots of approximated fractional prefilter $C_{pf}(s)$

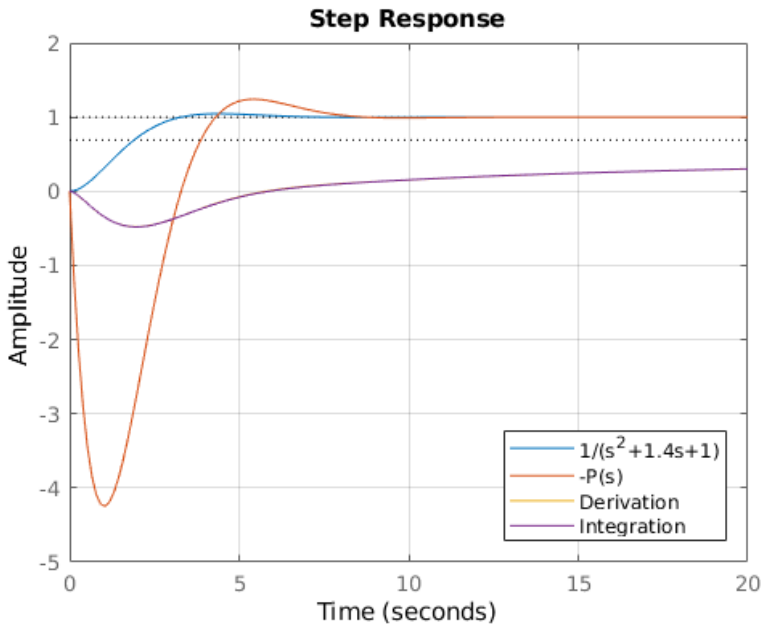

Figure F.4: Step response with approximated fractional prefilter $C_{pf}(s)$