Ilchmann, Joachim :

Contributions To Time-Varying Linear Control Systems

Druckausgabe:
Ilchmann, Achim: Contributions to time-varying linear control systems / Joachim Ilchmann. - Ammersbek b. Hamburg : Verl. an d. Lottbek, 1989
(Wissenschatfliche Beitrage aus europäischen Hochschulen : Reihe 11, Mathemalik ; Bd. 2)
Zugl.: Bremen, Univ., Diss., 1987
ISBN 3-926987-09-X

Wissenschaftliche Beiträge
aus
EUROPÄISCHEN HOCHSCHULEN

REIHE 11

Mathematik

BAND: 2

VaL

Joachim Ilchmann

Contributions To Time-Varying Linear Control Systems

Verlag an der Lottbek
Peter Jensen

VadL

CIP-Titelaufnahme der Deutschen Bibliothek

Ilchmann, Achim:
Contributions to time-varying linear control systems / Joachim Ilchmann.- Ammersbek b. Hamburg : Verl. an d. Lottbek, 1989
(Wissenschaftliche Beiträge aus europäischen Hochschulen : Reihe 11, Mathematik ; Bd. 2)
Zugl.: Bremen, Univ., Diss., 1987
ISBN 3-926987-09-X
NE: Wissenschaftliche Beiträge aus europäischen Hochschulen/11

Ammersbek bei Hamburg
ISBN 3-926987-09-X
ISSN 0934-0505
Druck: Fotodruck an der Uni; Hamburg
Umschlagentwurf: Peter Jensen

Mathematics Subject Classification (1980)

Primary: 93C05, 93C50, 34A30, 34D10

Secondary: 93B05, 93B07,93B10, 93B17, 93D05, 93D15

Key words:
time-varing linear finite dimensional control systems, algebraic description, disturbance decoupling, stabilizability, robustness, differential matrix Riccati equation, controllability, time-varying subspaces, system equivalence, (A,B)-invariance, noninteracting, exponential stability, exponential dichotomy, stability radius

Contents

Introduction

In this thesis linear finite dimensional time-varying control systems in state space form

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) \end{aligned} \tag{0.1}$$

resp. in differential operator description

$$\begin{aligned} P(D)(z) &= Q(D)(u) \\ y &= V(D)(z) + W(D)(u) \end{aligned} \tag{0.2}$$

are studied. Roughly speaking, three different mathematical techniques (algebraic, geometric and analytic) are used to analyse various control theoretic problems of time-varying systems.

Algebraic approach

Rosenbrock (1970) introduced the well-known setting of time-invariant systems in differential operator description (0.2). He studied the question under which conditions two constant control systems represented by (0.2) have the same dynamics and the same input-output behaviour. This is the problem of (strict) *system equivalence*. *Fuhrmann* (1976) and (1977) analyzed this problem via model theoretic tools. This enabled him to associate a canonical state space module with any factorization $V(z)P(z)^{-1}Q(z) + W(z)$ of a proper rational transfer matrix. So far the analysis of the problem of system equivalence for time-invariant systems was done in the *frequency domain*. *Pernebo* (1977) was the first who studied system equivalence in the *time domain*, his basic idea was to consider solution sets of the system equations. This approach was systematically exploited by *Hinrichsen and Prätzel-Wolters* (1980) to obtain a self-contained theory of system equivalence in the time domain. They derived an *algebraic criterion* of system equivalence, defined and characterized *controllability* and *observability*, and presented a *canonical state space model* similar to Fuhrmann's model.
Concerning *time-varying* systems, for a long time there has been a widespread scepticism whether an algebraic treatment in the style of *Kalman* would at all be possible. In particular it was not clear how to extend transform techniques. There were some attempts to study time-varying equations of the form (0.2), cf. *Ylinen* (1975) and (1980), *Kamen* (1976). *Ilchmann, Nürnberger and*

Schmale (1984) were guided by the time-invariant approach of *Hinrichsen and Prätzel-Wolters* (1980) when they generalized the concept of system equivalence for time-varying systems. They considered system matrices defined over a certain ***skew polynomial ring*** and introduced the notion of *"full"* differential operators. This set them in a position to generalize, for a fairly rich class of time-varying systems of the form (0.2), the time-invariant results of *Hinrichsen and Prätzel-Wolters* (1984). These results are presented in the first half of Chapter **2** of this thesis. A module theoretic approach to different definition of ***structural indices*** of time-invariant state space systems was given by *Münzner and Prätzel-Wolters* (1979). Via polynomial modules and their minimal bases they proved the equality of ***controllability indices, minimal indices, geometric indices*** and ***dynamical indices***. *Prätzel-Wolters* (1981) continued this work to characterize Brunovský-equivalence for time-invariant systems of the form (0.2).
Guided by this approach and using the skew polynomial ring introduced in *Ilchmann, Nürnberger and Schmale* (1984) I generalized the results of *Münzner and Prätzel-Wolters* (1979) and *Prätzel-Wolters* (1981) for time-varying systems (see *Ilchmann* (1985a)). The characterization of ***minimal bases*** of ***right skew polynomial modules*** extends a result of *Forney* (1975). It is possible to define a ***transfer matrix*** in the time domain and to use this to characterize system equivalence. Different ***invariants*** with respect to system equivalence resp. similarity were defined and their equality was shown. This is presented in the second half of Chapter **2** of this thesis.

Geometric approach

In the late sixties *Basile and Marro* (1969) and *Wonham and Morse* (1970) developed the concept of *(A,B)*-invariant subspaces to solve decoupling and pole assignment problems for multivariable systems. Later *Wonham* (1974) established the so called geometric approach. This approach was generalized for nonlinear systems by *Hirschhorn* (1981), *Isidori, Krener, Gori-Giori and Monaco* (1981) and for infinite dimensional systems by *Curtain* (1985), (1986), to name a few.
In Chapter **1** of this thesis ***time-varying subspaces*** are studied. This turns out to be the appropriate framework to extend the linear time-invariant geometric approach to *piecewise analytic* state space systems. If only analytic systems are considered this approach is a specialization of the nonlinear setting. However the concept proposed here is more "natural" for time-varying systems (differential geometry is not used) and the class of piecewise analytic systems is richer than the class of analytic systems. In Chapter **3** of this thesis I present the results which were essentially given in *Ilchmann* (1985b) and (1986). The concepts of *(A,B)*- and *(C,A)-invariance* are introduced, characterized and their dual relationships are shown. By using these results the solvability of the ***disturbance decoupling problem*** and the ***noninteracting problem*** is characterized.

Analytic approach

Concerning exponential stability of systems of the form

$$\dot{x}(t) = A(t)x(t), \qquad t \geq 0 \tag{0.3}$$

three aspects are studied in this thesis: for short, sufficient conditions for exponential *stability*; sufficient and necessary conditions for the *stabilizability* of systems (0.1) by state feedback; *robustness* of stability.

It is well-known that if, for all $t \geq 0$, the spectrum of $A(t)$ is lying in the open left half plane and the parameter variation of $A(t)$ is "slow enough", then (0.3) is exponentially stable. See e.g. *Rosenbrock* (1963), *Coppel* (1978). However, these results are qualitative. In *Ilchmann, Owens and Prätzel-Wolters* (1987b) we derived *quantitative* results. This means, upper bounds for the eigenvalues and for the rate of change of $A(t)$ which ensure exponential stability of (0.3) are determined. This is presented in Section 4.1 of the present thesis.

Ikeda, Maeda and Kodama (1972) and (1975) studied the problem to *stabilize* a time-varying system (0.1) by state feedback. Furthermore they gave a sufficient condition which guarantees that (0.1) is stabilizable by deterministic state estimation feedback. In *Ilchmann and Kern* (1987) these problems were analysed in case that the system (0.3) possesses an *exponential dichotomy*. When this is assumed the concept of controllability into subspaces, introduced in Section 1.2, is the appropriate tool to give necessary and sufficient conditions for stabilizability. These results are presented in Section 4.2.

In the remainder of the "analytic chapter" some *robustness* issues concerning the stability of (0.3) are studied. For time-invariant systems there exist two fundamental approaches concerning stability: the successful H^∞-approach (see *Zames*(1981) and *Francis and Zames* (1983)) based on transform techniques and the state space approach (see *Hinrichsen and Pritchard* (1986a,b)) based on the concept of "stability radius". It is not clear how to extend transform techniques to time-varying systems, whereas there are natural extensions in the state space setup. *Hinrichsen and Pritchard* (1986a) defined the (complex) *stability radius* of $A \in \mathbb{C}^{n\times n}$ as the distance of A from the set of unstable matrices in the Euclidean topology. In *Hinrichsen and Pritchard* (1986b) they also treated structured perturbations of the form BDC (B, C are known scaling matrices) and showed that the associated *structural stability radius* $r_{\mathbb{C}}(A; B, C)$ can be determined by the norm of a certain convolution operator (*"perturbation operator"*). Using optimization techniques they proved that $r^2_{\mathbb{C}}(A; B, C)$ is the maximal parameter $\rho \in \mathbb{R}$ for which the *algebraic Riccati equation*

$$A^*P + PA - \rho C^*C - PBB^*P = 0$$

has an Hermitian solution.
In *Hinrichsen, Ilchmann and Pritchard* (1987) these results were partially extended to time-varying systems. A new class of time-varying coordinate transformations (*Bohl transformations*) was introduced and a lower bound for the stability radius $r_{\mathbb{C}}(A; B, C)$ in terms of the norm of perturbation operator was given. Existence of maximal bounded Hermitian solutions of the *differential Riccati equation* parametrized by $\rho \in \mathbb{R}$

$$\dot{P}(t) + A(t)^*P(t) + P(t)A(t) - \rho C(t)^*C(t) - P(t)B(t)B(t)^*P(t) = 0, \qquad t \geq 0$$

was characterized via the norm of the perturbation operator. This is presented in Sections 4.3 to 4.8 of this thesis.

Each chapter has an own detailed introduction. A *subject* and *symbol index* can be found at the end.

Chapter 1

Controllability and Observability for State Space Systems

1.0 Introduction

Controllability and observability are basic concepts in systems and control theory. A first mathematical description was given by *Kalman* (1960). From then on these concepts were studied extensively not only for time-invariant state space systems, but for time-varying systems as well. However, in this chapter some definitions and characterizations concerning controllability are presented which have been not considered before.
The concept of *controllability into subspaces* is introduced, this will become useful when the problem of stabilizabiity of systems which possess an exponential dichotomy is studied in Section 4.2. Extending Rosenbrock's deleting procedure I define controllability indices and use this to derive a canonical form for analytic state space systems. For the study of different structural indices in Section 2.6 this canonical form will become useful. Controllability and observability induce certain *time-varying subspaces*. Time-varying subspaces in general are studied in depth in Section 5. This is a basic tool for the geometric approach of time-varying system presented in Chapter 3.
So far Chapter 1 is a preliminary chapter. I have put together some basic definitions and concept which I will refer to in the following chapters. On the other hand the contents of Chapter 1 have some interest of their own, they serve for a deeper understanding of controllability of time-varying systems.

In Section 1 some notations and certain skew polynomial rings are introduced.
The concept of controllability into a subspace is studied in Section 2.
The dual and adjoint relationships between controllability, reconstructibility, reachability and observability with respect to subspaces are explained in Section 3.
In Section 4 *Rosenbrock's* deleting procedure is generalized and controllability indices for analytic state space systems are defined. They are used to derive a canonical form.
Families of time-varying subspaces are studied in Section 5 and the results are applied to piecewise analytic state space systems.

Section 2 and 3 are mainly based on *Ilchmann and Kern* (1987); Section 4 and Section 5 are based on *Ilchmann* (1985a), (1985b) resp.

1.1 Basic notations and definitions

In this chapter state space systems of the following form will be considered

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) \qquad , t \in \mathbb{R} \end{aligned} \tag{1.1}$$

where $A(\cdot), B(\cdot), C(\cdot)$ are $n \times n$, $n \times m$, $p \times n$ matrices, resp., with entries in a ring $\mathcal{R}$.
$\mathcal{R}$ will be for instance

$\mathcal{C}_p$	the set of piecewise continuous functions $f : \mathbb{R} \to \mathbb{R}$
$\mathcal{C}^k$	the set of k-times continuously differentiable functions $f : \mathbb{R} \to \mathbb{R}$
$\mathcal{A}_p, \mathcal{A}$	the set of (piecewise) real analytic functions $f : \mathbb{R} \to \mathbb{R}$

A function $f : \mathbb{R} \to \mathbb{R}$ is called *piecewise real analytic* if there exists a disjoint partition $\cup_{\nu \in \mathbb{Z}}[a_\nu, a_{\nu+1}) = \mathbb{R}$, $\{a_\nu\}_{\nu \in \mathbb{Z}}$ a discrete set so that each restriction $f(\cdot)\mid_{(a_\nu, a_{\nu+1})}$ is real analytic and has a real analytic extension on some $(a_\nu^l, \; a_{\nu+1}^r)$, $a_\nu^l < a_\nu$, $a_{\nu+1} < a_{\nu+1}^r$. Clearly every function in $\mathcal{A}_p$ is piecewise differentiable.
In this chapter we consider the whole real axis as the time domain of (1.1). Most of the analysis goes through for any subinterval $I \subset \mathbb{R}$ as well.

Throughout the thesis a *fundamental matrix* of

$$\dot{x}(t) = A(t)x(t)$$

is denoted by $X(\cdot)$ and the *transition matrix* by

$$\Phi(t, t_0) = X(t)X(t_0)^{-1}.$$

Suppose

$$T = (t_{ij}) \in GL_n(\mathcal{R}) = \{T \in \mathcal{R}^{n \times n} \mid \exists T^{-1} \in \mathcal{R}^{n \times n} \;\; \forall t \in \mathbb{R} : T(t)T(t)^{-1} = I_n\}$$

and $\dot{T} = (\dot{t}_{ij}) \in \mathcal{R}^{m \times n}$, then the *coordinate transformation*

$$z(t) := T(t)^{-t}x(t)$$

converts the system (1.1) into

$$\begin{aligned} \dot{z}(t) &= A'(t)z(t) + B'(t)u(t) \\ y(t) &= C'(t)z(t) \qquad , t \in \mathbb{R} \end{aligned} \tag{1.2}$$

where

$$\begin{aligned} A' &= T^{-1}AT - T^{-1}\dot{T} && \in \mathcal{R}^{n \times n} \\ B' &= T^{-1}B && \in \mathcal{R}^{n \times m} \\ C' &= T\,C && \in \mathcal{R}^{p \times n} \end{aligned} \tag{1.3}$$

and the transition matrix $\Phi'(t, t_0)$ of (1.2) satisfies

$$\Phi'(t, t_0) = T(t)^{-1}\Phi(t, t_0)T(t_0).$$

In this case (1.1) and (1.2) are called *similar*.
For sake of brevity the tuples

$$(A, B) \in \mathcal{R}^{n \times (n+m)}, \qquad (A, B, C) \in \mathcal{R}^{n \times n} \times \mathcal{R}^{n \times m} \times \mathcal{R}^{p \times n}$$

are associated with the system (1.1).

Remark 1.1 As opposed to time-invariant systems, due to the much richer class of coordinate transformations, a system (A, B) is always similar to a system with constant free motion. More precisely, if $\mathcal{R}$ is $\mathcal{C}_p, \mathcal{C}^k, \mathcal{A}_p$ or $\mathcal{A}$ and $(A, B) \in \mathcal{R}^{n\times(n+m)}$ then the coordinate transformation $T(\cdot) = X(\cdot)$, where$X$ is a fundamental solution of $\dot{x}(t) = A(t)x(t)$, converts (A, B) into $(0, X^{-1}B) \in \mathcal{R}^{n\times(n+m)}$.

Clearly, similarity transformations will not, in general, preserve stability properties of the systems. Additional assumptions have to be imposed. If one requires that $T(\cdot), T(\cdot)^{-1}, \dot{T}(\cdot)$ are uniformly bounded in t one obtains the so called *Lyapunov transformations*, introduced by *Lyapunov* (1893) in his famous memoir, the stability behaviour is not affected. In this case (1.1) and (1.2) are called ***kinematically similar.***

In the remainder of this section certain skew polynomial rings are introduced. They will play an important role for the algebraic description of time-varying systems. The following basic properties of skew polynomial rings can be found, for instance, in *Cohn* (1971) Section 0.8.
Let R be any non-zero ring (not necessarily commutative) with no zero-divisors and x be an indeterminate over R. Then the ring $R[x; \alpha, \delta]$ generated by R and x is called a *right skew polynomial ring* if for some monomorphism $\alpha : R \to R$ and α-derivation $\delta : R \to R$, i.e. $\alpha(r \cdot s) = \alpha(r) \cdot \alpha(s)$ for all $r, s \in R$, the following *commutation rule* is valid

$$r \cdot x = x \cdot \alpha(r) + \delta(r) \qquad \text{for all } r, s \in R$$

Thus every element of $R[x; \alpha, \delta]$ is uniquely expressible in the form

$$r_0 + \ldots + x^n r_n \quad , r_i \in R$$

A left skew polynomial ring is defined analogously with commutation rule $x \cdot r = \alpha(r) \cdot x + \delta(r)$. If α is an automorphism then every left skew polynomial ring is a right one and vice versa.
To introduce certain skew polynomial rings which are important for an algebraic description of time-varying systems some notation is needed. Let $\mathcal{M}$ be the field of fractions of $\mathcal{A}$, i.e.:

$$\mathcal{M} := \{f : \mathbb{R} \to \overline{\mathbb{R}} \mid f \text{ is real meromorphic}\}.$$

By identifying each $f \in \mathcal{M}$ with the multiplication operator $f : g \mapsto fg$, $\mathcal{M}$ is a subring of $end_{\mathbb{R}}(\mathcal{M})$, the ring of $\mathbb{R}$-endomorphisms of $\mathcal{M}$. If

$$\begin{array}{rcl} D : \mathcal{M} & \to & \mathcal{M} \\ f & \mapsto & D(f) = \dot{f} \end{array}$$

denotes the derivation on $\mathcal{M}$ induced by the usual derivative, then $D \in end_{\mathbb{R}}(\mathcal{M})$ as well. The composition of D and f in $end_{\mathbb{R}}(\mathcal{M})$ is

$$(Df)(g) = D(fg) = f\dot{g} + \dot{f}g = (fD + \dot{f})(g) \qquad \text{for all } f, g \in \mathcal{M}$$

and one has the ***multiplication rule***

$$Df = fD + \dot{f} \qquad \text{for all } f \in \mathcal{M} \tag{1.4}$$

since D is algebraically independent over $\mathcal{M}$

$$\mathcal{M}[D] := \{f_0 + \ldots + f_n D^n \mid f_i \in \mathcal{M}, i \in \mathbb{N}_0\} \subset end_{\mathbb{R}}(\mathcal{M})$$

with *commutation rule* (1.4) is a (left and right) skew polynomial ring. Analogously the skew polynomial ring $\mathcal{A}[D]$ is defined. These rings are extensively studied in *Ilchmann, Nürnberger and Schmale* (1984).
A degree function on $\mathcal{A}[D]$ resp. $\mathcal{M}[D]$ is defined as usual. Since these rings do not contain zero divisors they allow a right and left division algorithm. So they are right and left Euclidean domains. Furthermore $\mathcal{A}[D]$ and $\mathcal{M}[D]$ are *simple*, i.e. the only two sided ideal of $\mathcal{A}[D]$ resp. $\mathcal{M}[D]$ are $\{0\}$ and the ring itself, cf. *Cozzens and Faith* (1975) p. 44.

Using this operational setup and the multiplication rule (1.4) one obtains

Lemma 1.2 If $(A,B) \in \mathcal{A}^{n\times(n+m)}$ and $T \in GL_n(\mathcal{A})$ then for $i \in \mathbb{N}$ and $(A',B') \in \mathcal{A}^{n\times(n+m)}$ satisfying (1.3) we have

$$\begin{aligned} T^{-1}[DI_n - A]T &= DI_n - A' && (1.5)\\ T^{-1}(DI_n - A)^i(B) &= (DI_n - A')^i(B') && (1.6)\\ (DI_n - A)^i(B) &= X(X^{-1}B)^i && (1.7) \end{aligned}$$

where

$$(DI_n - A)^i(B) := (DI_n - A)((DI_n - A)^{i-1}(B))$$

Proof: (1.5) is an immediate consequence of (1.3) and (1.4). We prove (1.6) by induction on i. For $i = 0$ it holds true by (1.3). If (1.6) is true for i, we conclude

$$T^{-1}(DI_n - A)^{i+1}(B) = T^{-1}(DI_n - A)T(T^{-1}(DI_n - A)^i(B)) = (DI_n - A')(DI_n - A')^i(B).$$

(1.7) is also easily shown by induction. □

1.2 Controllability into subspaces

Throughout this section state space systems $(A,B) \in \mathcal{C}_p^{n\times(n+m)}$ are considered.

The following generalization of the usual controllability concept will in particular be useful when systems of exponential dichotomy are analysed (compare Section 4.2).

Suppose that for some fundamental matrix $X(\cdot)$ of $\dot{x}(t) = A(t)x(t)$ the function space of free motions is decomposed into the direct sum

$$X(\cdot)\mathbb{R}^n = \mathcal{V}_1(\cdot) \oplus \mathcal{V}_2(\cdot) \tag{2.1}$$

where

$$\mathcal{V}_i(t) = X(t)P_i\mathbb{R}^n \qquad \text{for } i = 1,2 \ , t \in \mathbb{R}$$

and $P_1, P_2 \in \mathbb{R}^{n\times n}$ are mutually complementary projections, i.e. $P_1^2 = P_1$, $P_2 = I_n - P_1$. *Daleckii and Krein* (1974) p. 160 have proved that the projection associated with the linear subspace $\mathcal{V}_i(t)$ is similar to P_i and satisfies

$$P_i(t) = X(t)P_iX(t)^{-1} \qquad \text{for } i = 1,2 \text{ resp.,} \quad t \in \mathbb{R} \tag{2.2}$$

In particular it follows that if $x_0 \in \mathcal{V}_i(t_0)$ at some time t_0, then the free trajectory $\Phi(\cdot, t_0)x_o = X(\cdot)X(t_0)^{-1}x_0$ going through x_0 at t_0 belongs to $\mathcal{V}_i(\cdot), i = 1, 2$.

Definition 2.1 Suppose $(A, B) \in \mathcal{C}_p^{n\times(n+m)}$. The free trajectory $\Phi(\cdot, t_0)x_0$ is called *controllable at time t_0 into $\mathcal{V}_1$* if there exists some $t_1 \geq t_0$, $u(\cdot) \in (\mathcal{C}_p)^m$ and $x_1(\cdot) \in \mathcal{V}_1(\cdot)$ (all depending on t_0, x_0) so that

$$x(t) = \begin{cases} \Phi(t, t_0)x_0 & \text{for } t \leq t_0 \\ \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, s)B(s)u(s)ds & \text{for } t_0 \leq t \leq t_1 \\ x_1(t) & \text{for } t_1 \leq t \end{cases}$$

is a solution of $\dot{x}(t) = A(t)x(t) + B(t)u(t)$.
If this is true for every $x_0 \in \mathbb{R}^n$ we say that (A, B) is *completely controllable into $\mathcal{V}_1$ at time t_0*. If (A, B) is completely controllable into $\mathcal{V}_1$ at any time t_0 then (A, B) is called *completely controllable into $\mathcal{V}_1$*.

This definition does not say that every state in $\mathcal{V}_2(t_0)$ can be controlled to zero, but every free motion can be forced in finite time into a free motion of $\mathcal{V}_1$. If $\mathcal{V}_1 = \{0\}$ the above concept coincides with the well-known concept of controllability. In this case we omit "into $\mathcal{V}_1$" and speak only of controllability.

For later purposes it is often necessary to choose the input space $(\mathcal{A}_p)^m$ instead of $(\mathcal{C}_p)^m$. That this is not a restriction is a consequence of the following propostion.

Proposition 2.2 Consider $(A, B) \in \mathcal{C}_p^{n\times(n+m)}$. If the state $x_0 \in \mathbb{R}^n$ at time t_0 can be controlled to $x_1 \in \mathbb{R}^n$ at time $t_1 > t_0$ by $u(\cdot) \in (\mathcal{C}_p)^m$, this can also be achieved by some $\tilde{u}(\cdot) \in (\mathcal{A}\mid_{\overline{I}})^m$, where $I = (t_0, t_1)$.

Proof: By Remark 1.1 and Lemma 1.2 it can be easily seen that it is sufficient to assume $A(\cdot) \equiv 0$. Then it remains to prove that $imG \subset imH$, where

$$\begin{aligned} G : (\mathcal{C}_p\mid_{\overline{I}})^m &\rightarrow \mathbb{R}^r \\ u(\cdot) &\mapsto \int_{t_0}^{t_1} B(s)u(s)ds \end{aligned}$$

and

$$\begin{aligned} H : (\mathcal{A}\mid_{\overline{I}})^m &\rightarrow \mathbb{R}^r \\ \tilde{u}(\cdot) &\mapsto \int_{t_0}^{t_1} B(s)\tilde{u}(s)ds \end{aligned}$$

Let $g_1, \ldots, g_k$ denote a basis of $imG \subset \mathbb{R}^r$ and choose $u_i(\cdot) \in (\mathcal{C}_p\mid_{\overline{I}})^m$ such that $G(u_i(\cdot)) = g_i$ for $i \in \underline{k}$. $\mathcal{C}_{\overline{I}}$ lies dense in $\mathcal{C}_p\mid_{\overline{I}}$ with respect to the L_1-norm and, by the Weierstraß-Theorem, the set of real polynomials restricted to I lies dense in $\mathcal{C}_{\overline{I}}$. Thus for every $\delta > 0$ there exists $\tilde{u}(\cdot) \in (\mathcal{A}\mid_{\overline{I}})^m$ such that

$$\| u_i(\cdot) - \tilde{u}_i(\cdot) \|_{L_1} < \delta \qquad \text{for } i \in \underline{k}$$

Clearly, for $\varepsilon > 0$ small enough and $\tilde{g}_i \in im\ G$ "$\| \tilde{g}_i - g_i \| < \varepsilon$ for $i \in \underline{k}$" implies that $\tilde{g}_1, \ldots, \tilde{g}_k$ is also a basis of imG. Now, by continuity of H, choose $\delta > 0$ suficiently small such that $\| H(\tilde{u}_i) - g_i \| < \varepsilon$ for $i \in \underline{k}$. Then $H(\tilde{u}_1), \ldots, H(\tilde{u}_k)$ is a basis of $im\ G$ and this completes the proof. □

Instead of the *controllability Gramian*

$$W(t_0, t_1) = \int_{t_0}^{t_1} \Phi(t_0, s) B(s) B^T(s) \Phi^T(t_0, s) ds$$

of $(A, B) \in C_p^{n \times (n+m)}$ the *induced controllability Gramian*

$$W_2(t_0, t_1) = P_2(t_0) W(t_0, t_1) P_2^T(t_0)$$

will become an important tool to characterize controllability into subspaces. We call $W_2(t_0, t_1)$ *positive definite on* $P_2^T \mathbb{R}^n$ if for all non-trivial vectors $q \in P_2^T(t_0)\mathbb{R}^n$ we have $q^T W_2(t_0, t_1) q > 0$.

The following invariance properties are easily verified.

Remark 2.3 Suppose $(A, B), (A', B') \in C_p^{n \times (n+m)}$ are similar via $T \in GL_n(C_p)$. If the projection $P_2'(\cdot)$ associated with $\dot{z}(t) = A'(t) z(t)$ is defined via the fundamental matrix $T^{-1}X$ then

$$\begin{aligned} P_2'(t) &= T(t)^{-1} P_2(t) T(t) \\ W_2'(t_0, t_1) &= T(t_0)^{-1} W_2(t_0, t_1) T(t_0)^{-1^T} \end{aligned}$$

and

(i) (A, B) is completely controllable into $\mathcal{V}_1$ iff (A', B') is completely controllable into $T^{-1}(\cdot)\mathcal{V}_1(\cdot)$.

(ii) $W_2(t_0, t_1)$ is positive definite on $P_2^T(t_0)\mathbb{R}^n$ iff $W_2'(t_0, t_1)$ is positive definite on $P_2'^T(t_0)\mathbb{R}^n$.

The main result of this section are the following various characterizations of controllability into subspaces.

Theorem 2.4 For the system $(A, B) \in C_p^{n \times (n+m)}$ the following are equivalent:

(i) (A, B) is completely controllable into $\mathcal{V}_1$ at time t_0.

(ii) There exists some $t_1 > t_0$ such that

$$W_2(t_0, t_1) \text{ is positive definite on } P_2^T(t_0)\mathbb{R}^n$$

resp.

$$imW_2(t_0, t_1) = P_2(t_0)\mathbb{R}^n$$

resp.

the map

$$\begin{array}{rcll} \varphi_{t_1} & : & C_p^m & \to & P_2(t_0)\mathbb{R}^n \\ & & u(\cdot) & \mapsto & \int_{t_0}^{t_1} X(t_0)P_2X^{-1}(s)B(s)u(s)ds \end{array}$$

is surjective.

(iii) Every non trivial solution

$$y(\cdot) = \Phi^T(t_0, \cdot)P_2^T(t_0)q \qquad ,q \in \mathbb{R}^n$$

of the adjoint equation of $\dot{x}(t) = A(t)x(t)$

$$\dot{y}(t) = -A^T(t)y(t)$$

has the property

$$y^T(\cdot)B(\cdot)\,|_{[t_0,\infty)} \not\equiv 0.$$

If $rkP_1 = k$ and A, B have entries in $\mathcal{C}^{n-k-1}, \mathcal{C}^{n-k}$ resp. then (i) is a consequence of (iv).

(iv) There exists some $t_1 > t_0$ such that

$$rk[P_2(t)B(t), \ldots, (D - A(t))^{n-k}(P_2(t)B(t))] = n - k$$

for some $t \in (t_0, t_1)$.

If $(A, B) \in \mathcal{A}^{n\times(n+m)}$ then (i) is equivalent to (iv) and to

(v) There exist $U(D) \in \mathcal{M}[D]^{n\times n}$, $V(D) \in \mathcal{M}[D]^{m\times n}$ such that for all $t \in \mathbb{R}$

$$P_2(t) = [DI_n - A(t)] \cdot U(D) + P_2(t)B(t) \cdot V(D).$$

Proof: Note that (1.6) also holds true for $i \in \underline{n-k}$ if $(A, B) \in (\mathcal{C}^{n-k-1})^{n\times n} \times (\mathcal{C}^{n-k})^{n\times m}$ Thus, using Remark 1.1, it is easily seen that it is sufficient to prove the theorem for the case $A(\cdot) \equiv 0$ and $X(\cdot) \equiv I_n$.
The equivalence of the three statements in (ii) is proved analogously to the usual situation where $P_2 = I_n$, cf. *Knobloch and Kappel* (1974) p. 103. We omit the proof.
To simplify the proof a further condition is introduced:

(iii') For every $t_0 \in \mathbb{R}$ there exists $t_1 > t_0$ so that

$$y(\cdot) \equiv \Phi(t_0, \cdot)^T P_2^T(t_0)q \not\equiv 0 \;\Rightarrow\; y^T(t)B(t) \neq 0 \quad \text{for some } t \in [t_0, t_1]$$

and we proceed as follows

$$\text{(i)} \Longleftrightarrow \text{(ii)} \Longrightarrow \text{(iii)} \Longrightarrow \text{(iii')} \Longrightarrow \text{(ii)}$$

(i) $\Rightarrow$ (ii) : For $t_0 \in \mathbb{R}$ and $x_0 \in \mathbb{R}^n$ there exists $t_1 \geq t_0$ and an input vector $u(\cdot)$ such that

$$x_1 = x_0 + \int_{t_0}^{t_1} B(s)u(s)ds \in P_1\mathbb{R}^n.$$

Then

$$P_2(x_1 - x_0) = -P_2x_0 = \int_{t_0}^{t_1} P_2B(s)u(s)ds.$$

which proves the second statement in (ii).

(ii) $\Rightarrow$ (i): It suffices to determine for arbitrary $t_0 \in \mathbb{R}, x_0 \in \mathbb{R}^n$ some $t_1 > t_0$ and a control function $u(\cdot)$ such that

$$x(t_1) = P_1x_0 + P_2x_0 + \int_{t_0}^{t_1} B(s)u(s)ds \in P_1\mathbb{R}^n.$$

By assumption there exists $t_1 > t_0$ such that $W_2(t_0, t_1)$ is positive definite on $P_2^T\mathbb{R}^n$. Defining

$$u(t) = \begin{cases} -B^T(t)W_2^{-1}(t_0,t_1)P_2x_0 & \text{for } t_0 \leq t \leq t_1 \\ 0 & \text{for } t > t_1 \end{cases}$$

gives

$$P_2(P_2x_0 + \int_{t_0}^{t_1} B(s)u(s)ds) = P_2x_0 - \int_{t_0}^{t_1} P_2B(s)B^T(s)P_2^T ds \;\; W_2^{-1}(t_0,t_1)P_2x_0 = 0.$$

This proves $x(t_1) \in P_1\mathbb{R}^n$.

(ii) $\Rightarrow$ (iii): by contradiction. Assume that for some $P_2^T q \neq 0$

$$y^T(t)B(t) = q^TP_2B(t) = 0 \quad \text{for all } t \geq t_0.$$

Then

$$q^TW_2(t_0,t_1)q = \int_{t_0}^{t_1} q^TP_2B(s)B^T(s)P_2^Tq \; ds = 0 \quad \text{for all } t_1 > t_0$$

which contradicts (ii).

(iii) $\Rightarrow$ (iii') is proved analogously as in *Knobloch and Kwakernaak* (1985) p.33.

(iii') $\Rightarrow$ (ii): It suffices to prove that for arbitrary $t_0 < t_1$ the implication

$$y(\cdot) \equiv P_2^Tq \neq 0 \;\Rightarrow\; y^T(t)B(t) \neq 0 \quad \text{for some } t \in [t_0, t_1]$$

implies that $W_2(t_0,t_1)$ is positive definite on $P_2^T\mathbb{R}^n$. The proof is immediate by contradiction.

In order to prove (iv) $\Rightarrow$ (ii) the following notation is used.
Let

$$S^{-1}P_1S = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} \text{ for some } S \in GL_n(\mathbb{R})$$

Define

$$F(t) := \begin{bmatrix} 0 \\ \vdots \\ 0 \\ f_{k+1}(t) \\ \vdots \\ f_n(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & I_{n-k} \end{bmatrix} S^{-1}B(t) \in \mathbb{R}^{n\times m},$$

$$\overline{F}(t) := \begin{bmatrix} f_{k+1}(t) \\ \vdots \\ f_n(t) \end{bmatrix} \in \mathbb{R}^{(n-k)\times m}.$$

At first it is shown that for arbitrary $t_0 < t_1$ the following are equivalent:

(α) $W_2(t_0,t_1)$ is positive definite on $P_2^T\mathbb{R}^n$.

(β) $rk \int_{t_0}^{t_1} \overline{F}(s)\overline{F}^T(s)ds = n-k$

(γ) The row vector functions $f_{k+1}(t),\dots,f_n(t)$ are linearly independent on $[t_0.t_1]$.

$(\alpha) \Leftrightarrow (\beta)$: We have

$$S^{-1}W_2(t_0,t_1)S = \int_{t_0}^{t_1} F(s)F^T(s)ds.$$

Since

$$F(t)F^T(t) = \begin{bmatrix} 0_{k\times k} & 0_{k\times(n-k)} \\ 0_{(n-k)\times k} & \overline{F}(t)\overline{F}^T(t) \end{bmatrix}$$

the equivalence is obvious.

$(\beta) \Leftrightarrow (\gamma)$ is a consequence of Gram's criterium, cf. *Gantmacher* (1959) p. 247.

Since

$$\begin{aligned} rk[P_2(t)B(t),\dots,(P_2(t)B(t))^{n-k}] &= rk[S(t)F(t),\dots,S(t)F^{(n-k)}(t)] = rk[F(t),\dots,F^{(n-k)}(t)] \\ &= rk[\overline{F}(t),\dots,\overline{F}^{n-k}(t)] \end{aligned}$$

it follows that (iv) is equivalent to

$$rk[\overline{F}(t),\dots,\overline{F}^{(n-k)}(t)] = n-k \quad \text{for some } t \in (t_0,t_1)$$

Now by Lemma 1 in *Silverman and Meadows* (1967) this condition gives (γ) and thus (iv) $\Rightarrow$ (ii) is proved.
If (A,B) is an analytic system then due to the Identity Theorem (iv) is equivalent to

$$rk[\overline{F}(t),\dots,\overline{F}^{(n-k)}(t)] = n-k \text{ for a set of points dense in } (t_0,t_1) \tag{2.3}$$

By Lemma 3 in *Silverman and Meadows* (1967) this condition coincides with (γ). Hence (ii) $\Leftrightarrow$ (iv) is proved.
It remains to prove (iv) $\Leftrightarrow$ (v) in the analytic situation. Since (iv) $\Leftrightarrow$ (2.3) it is sufficient to

show that (2.3) is equivalent to the existence of some $\hat{U} \in \mathcal{M}[D]^{(n-k)\times(n-k)}$, $\hat{V} \in \mathcal{M}[D]^{m\times(n-k)}$ such that

$$I_{n-k} = DI_{n-k} \cdot \hat{U} + \overline{F} \cdot \hat{V}. \tag{2.4}$$

This equation is valid iff

$$S \begin{bmatrix} 0 & 0 \\ 0 & I_{n-k} \end{bmatrix} S^{-1} = S \begin{bmatrix} DI_k & 0 \\ 0 & DI_{n-k} \end{bmatrix} S^{-1} \cdot S \begin{bmatrix} 0 \\ \hat{U} \end{bmatrix} S^{-1} + S \begin{bmatrix} 0 \\ \overline{F} \end{bmatrix} \cdot \hat{V} S^{-1}.$$

Since

$$P_2 = S \begin{bmatrix} 0 & 0 \\ 0 & I_{n-k} \end{bmatrix} S^{-1}, \ DI_n = S \begin{bmatrix} DI_k & 0 \\ 0 & DI_{n-k} \end{bmatrix} S^{-1}, \ P_2 B(t) = S \begin{bmatrix} 0 \\ \overline{F}(t) \end{bmatrix}$$

(2.4) is equivalent to (v) with $A(\cdot) \equiv 0$.
For the following proof of (2.3) $\Leftrightarrow$ (2.4) compare *Ilchmann, Nürnberger and Schmale* (1984) pp. 357/8. Suppose (2.3), then there exist $Y_i \in \mathcal{M}^{m\times(n-k)}, i = 0, \ldots, n-k$ such that

$$[\overline{F}, \ldots, \overline{F}^{(n-k)}] \begin{bmatrix} Y_0 \\ \vdots \\ Y_{n-k} \end{bmatrix} = I_{n-k}.$$

Using the multiplication rule (1.4) it is easily proved by induction that

$$\overline{F} D^i = \sum_{\lambda=0}^{i-1} \binom{i}{\lambda} (-1)^\lambda D^{i-\lambda} \overline{F}^{(\lambda)} + (-1)^i \overline{F}^{(i)} = D \cdot M_i(D) + (-1)^i \overline{F}^{(i)} \tag{2.5}$$

where

$$M_i(D) := \sum_{\lambda=0}^{i-1} \binom{i}{\lambda} (-1)^\lambda D^{i-\lambda-1} \overline{F}^{(\lambda)}$$

Thus

$$\begin{aligned}
\overline{F}[Y_0 - DY_1 + \ \ldots \ &+ (-1)^{n-k} D^{n-k} Y_{n-k}] \\
&= \overline{F} Y_0 - [DM_1(D) - \dot{\overline{F}}] Y_1 + \ldots + (-1)^{n-k} [DM_{n-k}(D) + (-1)^{n-k} \overline{F}^{(n-k)}] Y_{n-k} \\
&= \overline{F} Y_0 + \dot{\overline{F}} Y_1 + \ldots + \overline{F}^{(n-k)} Y_{n-k} + \sum_{\lambda=1}^{n-k} (-1)^\lambda D M_\lambda(D) Y_\lambda \\
&= [\overline{F}, \ldots, \overline{F}^{(n-k)}] \begin{bmatrix} Y_0 \\ \vdots \\ Y_{n-k} \end{bmatrix} + D \sum_{\lambda=1}^{n-k} (-1)^\lambda M_\lambda(D) Y_\lambda \\
&= I_{n-k} + D \sum_{\lambda=1}^{n-k} (-1)^\lambda M_\lambda(D) Y_\lambda
\end{aligned}$$

and (2.4) is proved.
Finallay suppose (2.3) holds true and

$$\hat{V} = \sum_{i=0}^{q} D^i \hat{V}_i, \ \hat{V}_i \in \mathcal{M}^{m\times(n-k)} \ \text{for} \quad i = 0, \ldots, q$$

By (2.5) one obtains

$$
\begin{aligned}
I_{n-k} &= D\hat{U} + \overline{F} \cdot \hat{V} \\
&= D\hat{U} + \sum_{i=0}^{q} \sum_{\lambda=0}^{i} (-1)^i \begin{pmatrix} i \\ \lambda \end{pmatrix} D^{i-\lambda} \overline{F}^{(\lambda)} \hat{V}_i \\
&= D\hat{U} + \sum_{i=0}^{q} [\sum_{\lambda=0}^{i-1} (-1)^i \begin{pmatrix} i \\ \lambda \end{pmatrix} D^{i-\lambda} \overline{F}^{(\lambda)} + (-1)^i \overline{F}^{i}]\, \hat{V}_i \\
&= D\hat{U} + D[\sum_{i=0}^{q} \sum_{\lambda=0}^{i-1} (-1)^i \begin{pmatrix} i \\ \lambda \end{pmatrix} D^{i-\lambda-1} \overline{F}^{(\lambda)} \hat{V}_i] + \sum_{i=0}^{q} \overline{F}^{(i)} (-1)^i \hat{V}_i
\end{aligned}
$$

and comparing the coefficients yields

$$
\sum_{i=0}^{q} \overline{F}^{i} (-1)^i \hat{V}_i = I_{n-k}
$$

Since $rk_{\mathcal{M}}[\overline{F}, \ldots, \overline{F}^{(i)}]$ considered as a function of i can only be strictly monotonic within the set $\{0, \ldots, n-k\}$, (2.3) follows and the proof is complete. □

Suppose the system (A, B) is controllable at time t_0, i.e. $P_1 = 0$. Then condition (ii) of Theorem 2.4 is proved in *Kalman, Ho and Narenda* (1963); (iii) is a recent result of *Knobloch and Kwackernaak* (1985); (iv) was shown by *Silverman and Meadows* (1967);(v) is proved by *Ilchmann, Nürnberger and Schmale* (1984).
If $P_0 = 0$ and additionally $(A, B) \in \mathbb{R}^{n\times(n+m)}$, then (iv) is known as the rank condition of the controllability matrix derived by *Kalman* (1960), (v) represents the left coprimeness of $sI_n - A$ and B, see *Rosenbrock* (1970).

Remark 2.5
(i) Suppose $(A, B) \in C_p^{n\times(n+m)}$ is completely controllable into $\mathcal{V}_1$ at time t_0. Then a control which forces a free trajectory $\Phi(\cdot, t_0)x_0$ from time t_0 into $\mathcal{V}_1$ at time t_1 is given by $u(t) = -B^T(t)W_2(t_o, t_1)^{-1}P_2(t_0)$ (compare the proof of "(ii) $\Rightarrow$ (i)" in Theorem 2.4). Since there exists a minimal time $t_1 > t_0$ such that $W_2(t_0, t_1)$ is positive definite on $P_2^T(t_0)\mathbb{R}^n$, every $x_0 \in \mathbb{R}^n$ can be forced into $\mathcal{V}_1(t_1)$ is finite time $\delta = t_1 - t_0$, t_1 does not depend on x_0.

(ii) Since $im\varphi_{t_1} = W_2(t_0, t_1)$, where φ_{t_1} is the map given in the third condition of Theorem 2.4 (ii), we have

$$
imW_2(t_0, t_1) \subset imW_2(t_0, t_1') \quad \text{for } t_1' \geq t_1 > t_0.
$$

Thus the function $t \mapsto rkW_2(t_0, t)$ is monotonically increasing on (t_0, ∞) . If $(A, B) \in \mathcal{A}^{n\times(n+m)}$ then the entries of $W_2(t_0, \cdot)$ are analytic as well and by the Identity-Theorem of analytic functions $rkW_2(t_0, \cdot)$ is constant on (t_0, ∞). Therefore, if (A, B) is analytic and completely controllable into $\mathcal{V}_1$, then every state $x_0 \in \mathbb{R}^n$ at time t_0 can be forced into $\mathcal{V}_1(t_1)$ in *arbitrary* short time $t_1 - t_0 > 0$.

The next proposition will show that for (piecewise) analytic systems $imW_2(t_0, t_1)$ can be computed in terms of A and B . The knowledge of the transition matrix is not necessary.

Proposition 2.6 Suppose $(A,B) \in \mathcal{A}^{n\times(n+m)}$ and $t_0 < t_1$. Then

$$imW_2(t_0,t_1) = \Phi(t_0,t)\sum_{i\geq 0} im(DI_n - A(t))^i(P_2(t)B(t)) \qquad \text{for all } t \in [t_0,t_1] \qquad (2.6)$$

Suppose $(A,B) \in \mathcal{A}_p^{n\times(n+m)}$ and $\mathbb{R} = \cup_{\nu\in\mathbb{Z}}[a_\nu, a_{\nu+1})$ is a partition such that $A(\cdot)$ and $B(\cdot)$ are analytic on every $(a_\nu, a_{\nu+1})$. Then for $t_0 \in [a_0, a_1)$ and $t_1 \in [a_N, a_{N+1})$ one obtains

$$imW(t_0,t_1) = \sum_{i\geq 0} im(DI_n - A(t_0))^i(B(t_0)) + \sum_{\nu=1}^{N}\sum_{i\geq 0} im(DI_n - A(a_\nu))^i(B(a_\nu)) \qquad (2.7)$$

Proof: Using Remark 1.1 and (1.6) it is easily seen that without loss of generality one may assume $A = 0, X = I_n$. So it remains to prove

$$imW_2(t_0,t_1) = \sum_{i\geq 0} im(P_2B(t))^{(i)} \qquad \text{for all } t \in [t_0,t_1]$$

which is equivalent to

$$kerW_2(t_0,t_1) = [\sum_{i\geq 0} im(P_2B(t))^{(i)}]^\perp \qquad \text{for all } t \in [t_0,t_1]$$

Due to the properties of analytic functions it is easily seen that

$$\begin{aligned} & q \in W_2(t_0,t_1) \\ \Leftrightarrow \quad & \\ & [P_2B(t)]^T q = 0 \quad \text{for all } t \in [t_0,t_1] \\ \Leftrightarrow \quad & \\ & [P_2B(t)]^{T^{(i)}} q = 0 \quad \text{for some } t \in [t_0,t_1], \ \text{for all } i \geq 0 \\ \Leftrightarrow \quad & \\ & q \in \bigcap_{i\geq 0} ker[P_2B(t)]^{T^{(i)}} = [\sum_{i\geq 0} im(P_2B(t))^{(i)}]^\perp \ \text{for some } t \in [t_0,t_1] \end{aligned}$$

This proves (2.6). (2.7) follows from (2.6) and the fact that

$$imW_2(t_0,t_1) = imW_2(t_0,a_1) + \ldots + imW_2(a_N,t_1)$$

□

Remark 2.7
(i) Set $P_2 = I_n$ and $t = t_0$ in (2.6). Then for time-invariant systems an application of the Cayley-Hamilton-Theorem reduces (2.6) to the well-known fact that the *controllable space* is given by

$$imB + imAB + \ldots + imA^{n-1}B$$

(ii) In general it is not possible to restrict the sum in (2.6) independently of t_0 to only finitely many summands. See an example in *Kamen* (1979) p. 871.

(iii) If $A(\cdot)$ and $B(\cdot)$ are defined over $\mathbb{R}[t]$ it can be shown that the sum in (2.6) can be restricted to finitely many summands. Cf. the subclass of constant rank systems considered in *Silverman* (1971) and *Kamen* (1979).

It is well-known that for many control problems uniformity constraints are necessary. Uniform controllability as introduced by *Kalman* (1960) is extended in the present set-up as follows:

Definition 2.8 The system $(A, B) \in C_p^{n\times(n+m)}$ is called *uniformly completely controllable into* $\mathcal{V}_1$ if there exist $\sigma, a, b > 0$ such that

$$aI_n \leq W_2(t, t+\sigma) \leq bI_n \quad \text{on } P_2^T(t)\mathbb{R}^n \qquad \text{for all } t \in \mathbb{R} \tag{2.8}$$

Remark 2.9
(i) Suppose $A(\cdot)$ is bounded, i.e. there exists $c > 0$ such that $\| A(t) \| \leq c$ for all $t \in \mathbb{R}$. Since

$$\dot{x}(t) = A(t)x(t), \qquad x(t_0) = x_0$$

is equivalent to

$$x(t) = x_0 + \int_{t_0}^{t} A(s)x(s)ds$$

one obtains

$$\Phi(t, t_0) = I_n + \int_{t_0}^{t} A(s)\Phi(s, t_0)ds$$

and thus

$$\| \Phi(t, t_0) \| \leq 1 + \int_{t_0}^{t} c \| \Phi(s, t_0) \| ds$$

Now an application of Gronwall's Lemma yields

$$\| \Phi(t, t_0) \| \leq e^{c(t-t_0)} \quad \text{for all } t \geq t_0$$

Using this fact it is easily shown that an upper bound in (2.8) always exists if $A(\cdot)$ and $B(\cdot)$ are bounded.

(ii) A straightforward calculation shows that uniform complete controllability into $\mathcal{V}_1$ is preserved if a kinematical similarity transformation is applied to the system (A, B).

For later use we state the following lemma.

Lemma 2.10 Suppose $(A, B) \in C_p^{n\times(n+m)}$ is bounded and the matrices $F(\cdot) \in C_p^{n\times m}, E(\cdot) \in C_p^{m\times m}$ are bounded as well. Then (A, B) is uniformly completely controllable into $\mathcal{V}_1$ iff the system

$$\dot{x}(t) = [A(t) + B(t)F(t)]x(t) + B(t)E(t)u(t)$$

is uniformly completely controllable into $\mathcal{V}_1$.

Proof: The result is proved for uniform complete controllability by *Silvermann and Anderson* (1968). It carries over without any difficulties for uniform complete controllability into $\mathcal{V}_1$. □

1.3 Dual and adjoint relationships between controllability, reconstructibility, reachability and observability

Throughout this section we consider systems $(A, B) \in \mathcal{C}_p^{n\times(n+m)}$.

The well-known concepts of reconstructibility, reachability and observability (cf. *Kalman* (1960), *Knobloch and Kwakernaak* (1985)) are generalized with respect to time-varying subspaces. However, I only concentrate on those definitions and propositions which are of interest in the following. Analogous results as for controllability in Section 2 can be derived without any difficulties.

Following *Kwakernaak and Sivan* (1972) we define

Definition 3.1 The *dual system* of $(A, B) \in \mathcal{C}_p^{n\times(n+m)}$ with respect to some arbitrary fixed time t^* is given by

$$\begin{aligned} \dot{x}(t) &= A^T(t^*-t)x(t) + C^T(t^*-t)u(t) \\ y(t) &= B^T(t^*-t)x(t). \end{aligned} \tag{3.1}$$

The *adjoint system* of (A, B) is defined by

$$\begin{aligned} \dot{x}(t) &= -A^T(t)x(t) - C^T(t)u(t) \\ y(t) &= -B^T(t)x(t). \end{aligned} \tag{3.2}$$

Remark 3.2 It can easily be derived that a fundamental matrix $X^d(\cdot)$ resp. $X^a(\cdot)$ of the dual resp. adjoint system satisfies

$$\begin{aligned} X^d(t) &= [X^T(t^*-t)]^{-1} \\ X^a(t) &= [X^T(t)]^{-1} \end{aligned}$$

the associated transition matrices satisfy

$$\begin{aligned} \Phi^d(t,s) &= \Phi^T(t^*-s, t^*-t) \\ \Phi^a(t,s) &= \Phi^T(s,t) \end{aligned}$$

and the time-varying subspaces are defined by

$$\begin{aligned} \mathcal{V}_i^d(t) &= X^d(t)P_i^T\mathbb{R}^n \quad & i = 1,2 \\ \mathcal{V}_i^a(t) &= X^a(t)P_i^T\mathbb{R}^n \quad & i = 1,2 \end{aligned}$$

To state the dual and adjoint relationship of controllability into a subspace the following definitions are introduced.

Definition 3.3 $(A, B) \in \mathcal{C}_p^{n\times(n+m)}$ is said to be *completely reconstructible* wrt $\mathcal{V}_i$ if for every $t_0 \in \mathbb{R}$ there exists a $t_{-1} < t_0$ such that for every $x_0 \in X(t_0)P_i^T\mathbb{R}^n$ the condition

$$C(\cdot)\Phi(\cdot, t_0)x_0 \mid_{[t_{-1},t_0]} \equiv 0 \tag{3.3}$$

implies $x_0 = 0$, for $i = 1,2$ resp.
The *induced reconstructibility Gramian* is given by

$$H_i(t_{-1}, t_0) = X^{-1^T}(t_0)P_iX^T(t_0) \cdot \int_{t_{-1}}^{t_0} \Phi^T(s,t_0)C^T(s)C(s)\Phi(s,t_0)ds \cdot X(t_0)P_i^TX^{-1}(t_0)$$

for $i = 1,2$ resp.

Definition 3.4 $(A,B) \in C_p^{n\times(n+m)}$ is said to be *completely observable* wrt $\mathcal{V}_i$ if for every $t_0 \in \mathbb{R}$ there exists a $t_1 > t_0$ such that for every $x_0 \in X(t_0)P_i^T\mathbb{R}^n$ the condition

$$C(\cdot)\Phi(\cdot,t_0)x_0 \mid_{[t_0,t_1]} \equiv 0 \tag{3.4}$$

implies $x_0 = 0$, for $i = 1,2$ resp.
The *induced observability Gramian* is given by

$$G_i(t_0,t_1) = X^{-1^T}(t_0)P_iX^T(t_0) \cdot \int_{t_0}^{t_1} \Phi^T(s,t_0)C^T(s)C(s)\Phi(s,t_0)ds \cdot X(t_0)P_i^TX^{-1}(t_0)$$

for $i = 1,2$ resp.

Proposition 3.5 Let $t^* = 0$, then the Gramians associated with the system (A,B), the dual system and the adjoint system, resp. , satisfy

$$W_2(t_0,t_1) = H_2^d(-t_1,-t_0) = G_2^a(t_0,t_1) \tag{3.5}$$

Furthermore the following are equivalent:

(i) (A,B) is completely controllable into $\mathcal{V}_1$.

(ii) The dual system (3.1) wrt $t^* = 0$ is completely reconstructible wrt $\mathcal{V}_2^d$.

(iii) For every $t_0 \in \mathbb{R}$ there exists a $t_{-1} < t_0$ such that $H_2^d(t_{-1},t_0)$ is positive definite on $X^d(t_0)P_2^T\mathbb{R}^n$.

(iv) The adjoint system (3.2) is completely observable wrt $\mathcal{V}_2^a$.

(v) For every $t_0 \in \mathbb{R}$ there exists a $t_1 > t_0$ such that $G_2^a(t_0,t_1)$ is positive definite on $X^a(t_0)P_2^T\mathbb{R}^n$.

Proof: The formulas in Remark 3.2 give

$$\begin{aligned}
W_2(t_0,t_1) &= X(t_0)P_2 \cdot \int_{t_0}^{t_1} X^{-1}(s)B(s)B^T(s)X^{-1^T}(s)ds \cdot P_2^TX^T(t_0) \\
&= X^{d^{-1T}}(-t_0)P_2 \cdot \int_{t_0}^{t_1} X^{dT}(-s)C^{dT}(-s)C^d(-s)X^d(-s)ds \cdot P_2^TX^{d^{-1}}(-t_0) \\
&= X^{d^{-1T}}(-t_0)P_2 \cdot \int_{-t_1}^{-t_0} X^{dT}(\tau)C^{dT}(\tau)C^d(\tau)X^d(\tau)d\tau \cdot P_2^TX^{d^{-1}}(-t_0) \\
&= H_2^d(-t_1,-t_0)
\end{aligned}$$

and

$$\begin{aligned}
W_2(t_0,t_1) &= X^{a^{-1T}}(t_0)P_2 \int_{t_0}^{t_1} X^{aT}(s)(-C^{aT}(s))(-C^a(s))X^a(s)ds\, P_2^TX^{a^{-1}}(t_0) \\
&= G_2^a(t_0,t_1).
\end{aligned}$$

Hence (3.5) is proved. Now (i) $\Leftrightarrow$ (iii) and (i) $\Leftrightarrow$ (v) follows from Theorem 2.4. In order to prove (ii) $\Leftrightarrow$ (iii) put

$$\begin{aligned}
\varphi : X^d(t_0)P_2^T\mathbb{R}^n &\to (C_p[t_{-1},t_0])^m \\
x_0 &\mapsto C^d(\cdot)\Phi^d(\cdot,t_0)x_0
\end{aligned}$$

for some $t_1 < t_0$. Then (3.3) is equivalent to $x_0 \in \ker\varphi$ and this is equivalent to $x \in \ker H_2^d(t_{-1},t_0)$ (see *Knobloch and Kappel* (1974) p.112). The proof of (iv) $\Leftrightarrow$ (v) is analogous, it is omitted. □

Definition 3.6 The system $(A, B) \in C_p^{n\times(n+m)}$ is called *completely reachable from* $\mathcal{V}_2$ if every free trajectory $\Phi(\cdot, t_0)x_0$, for $t_0 \in \mathbb{R}, x_0 \in \mathbb{R}^n$, can be reached from a suitable free trajectory $x_2(\cdot) \in \mathcal{V}_2(\cdot)$, i.e. there exist $t_{-1} \le t_0, u(\cdot) \in C_p^m$ (both depending on t_0, x_0) so that

$$x(t) = \begin{cases} x_2(t) & \text{for } t \le t_{-1} \\ \Phi(t,t_{-1})x_2(t_{-1}) + \int_{t_{-1}}^{t} \Phi(t,s)B(s)u(s)ds & \text{for } t_{-1} \le t \le t_0 \\ \Phi(t,t_0)x_0 & \text{for } t_0 \le t \end{cases}$$

The *induced reachability Gramian* is given by

$$Y_1(t_{-1}, t_0) = P_1(t_0) \cdot \int_{t_{-1}}^{t_0} \Phi(t_0, s)B(s)B^T(s)\Phi^T(t_0, s)ds\ P_1^T(t_0).$$

Analogously to Proposition 3.5 one can prove the following, the proof is omitted.

Proposition 3.7 The Gramians associated with the system (A, B) , the dual system and adjoint system, resp., satisfy

$$Y_1(t_{-1}, t_0) = G_1^d(-t_0, -t_1) = H_1^a(t_{-1}, t_0). \tag{3.6}$$

Furthermore the following are equivalent:
(i) (A, B) is completely reachable from $\mathcal{V}_2$.
(ii) For every $t_0 \in \mathbb{R}$ there exists a $t_{-1} < t_0$ such that $Y_1(t_{-1}, t_0)$ is positive definite on $P_1^T(t_0)\mathbb{R}^n$.
(iii) The dual system (3.1) is complete observable wrt $\mathcal{V}_1^d$.
(iv) The adjoint system (3.2) is completely reconstructible wrt $\mathcal{V}_1^a$.

□

As opposed to time-invariant systems, complete controllability (observability) is not equivalent to complete reachability (reconstructability). For this see the following simple example.

Example 3.8 Put $n = 1, A(\cdot) \equiv 0$ and

$$B(t) = \begin{cases} 0 & , \quad t \le 0 \\ 1 & , \quad t > 0 \end{cases}$$

Then $(0, 1)$ is not reachable from $\mathcal{V}_2 := \{0\}$ however the system is completely controllable into $\mathcal{V}_2$.

As an immediate consequence of Remark 2.5 (ii) and the positive definite conditions on the Gramians we have the following corollary.

Corollary 3.9 If $(A, B) \in \mathcal{A}^{n\times(n+m)}$ then (A, B) is completely controllable into $\mathcal{V}_1$ (observable wrt $\mathcal{V}_2$) iff it is completely reachable from $\mathcal{V}_1$ (reconstructible wrt $\mathcal{V}_2$).

1.4 Controllability indices and a canonical form

Rosenbrock (1970) has introduced the controllability (Kronecker) indices for time-invariant systems. His definition will be extended to time-varying analytic systems $(A, B) \in \mathcal{A}^{n\times(n+m)}$. Set

$$K^i(A, B) = [B, (-1)(DI_n - A)^1(B), \ldots, (-1)^i(DI_n - A)^i(B)], \qquad i \in \mathbb{N}_0$$

where

$$(DI_n - A)^k(B) := (DI_n - A)((DI_n - A)^{k-1}(B)) \quad \text{for } i \in \underline{k}$$

The matrix

$$K(A,B) := K^{n-1}(A,B)$$

is called the *controllability matrix* of the system (A,B). If (A,B) is a constant system then $K(A,B) = [B,\dots,A^{n-1}B]$ is the well-known controllability matrix. For time-varying systems $K(A,B)$ was introduced by *Silverman and Meadows* (1967) in a slightly different form, namely without the factors $(-1)^k$ in it (cf. Theorem 2.4 (iv)), here it is modified for computational reasons.

Clearly, if (A,B) and (A',B') are similar via some $T(\cdot)$ then by (1.6)

$$T^{-1} \cdot K(A,B) = K(A',B') \tag{4.1}$$

and (1.7) yields

$$K(A,B) = [B, (-1)\Phi(\Phi^{-1}B)^{(1)}, \dots, (-1)^{n-1}\Phi(\Phi^{-1}B)^{(n-1)}].$$

Analyticity of the system makes it is possible to define the controllability indices of (A,B) by generalizing *Rosenbrock's* deleting procedure (see *Rosenbrock* (1970) p. 90) as follows:

> Eliminate in the controllability matrix of (A,B)
>
> $$K(A,B) = [B, (-1)\Phi(\Phi^{-1}B)^{(1)}, \dots, (-1)^{n-1}\Phi(\Phi^{-1}B)^{(n-1)}]$$
>
> from left to the right all column vectors which are linearly dependent over $\mathcal{M}$ upon their predecessors.

If the columns of B are denoted by $b_1,\dots,b_m$ one obtains after reordering

$$\begin{aligned} H &= [b_1,\dots,(-1)^{k_1-1}\Phi(\Phi^{-1}b_1)^{(k_1-1)} \quad ,\dots,b_m,\dots,(-1)^{k_m-1}\Phi(\Phi^{-1}b_m)^{(k_m-1)}] \\ &= [b_1,\dots,(-1)^{k_1-1}(DI_n - A)^{k_1-1}(b_1) \quad ,\dots,b_m,\dots,(-1)^{k_m-1}(DI_n - A)^{k_m-1}(b_m)] \in \mathcal{A}^{n\times n'} \end{aligned} \tag{4.2}$$

with $n' \leq n$ and some $k_1,\dots,k_m \in \mathbb{N}$. If $k_i = 0$ then the corresponding columns in H are absent. Note, if $\Phi(\Phi^{-1}b_i)^{(j)}$ is linearly dependent over $\mathcal{M}$ on its predecessors, then $\Phi(\Phi^{-1}b_i)^{(j+1)}$ is as well. The numbers $k_1,\dots,k_m$ are called the *controllability indices* of (A,B). As an immediate consequence of (4.1) they are invariant with respect to analytic similarity transformations.

Example 4.1 Let

$$(A,B) := \left(0_{3\times3}, \begin{bmatrix} e^t & -e^t & 0 \\ t-1 & 1 & t \\ 0 & t & t \end{bmatrix} \right).$$

It is ealisly computed that

$$K(A,B) = \begin{bmatrix} e^t & -e^t & 0 & -e^t & e^t & 0 & e^t & -e^t & 0 \\ t-1 & 1 & t & -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & t & t & 0 & -1 & -1 & 0 & 0 & 0 \end{bmatrix}$$

and

$$H(t) = [b_1(t), \dot{b}_1(t), b_2(t)] = \begin{bmatrix} e^t & -e^t & -e^t \\ t-1 & -1 & 1 \\ 0 & 0 & t \end{bmatrix}.$$

Therefore $(k_1,k_2,k_3) = (2,1,0)$.

As a complete generalization of the time-invariant case we prove the following characterizations of controllability.

Proposition 4.2 Suppose $(A,B) \in \mathcal{A}^{n\times(n+m)}$ has controllability indices $k_1,\ldots,k_m$ and $\ell := rk_{\mathcal{M}}B = \sum_{i:k_i>0} 1$. Then the following are equivalent.

(i) (A,B) is completely controllable

(ii) $rk_{\mathcal{M}}K(A(\cdot),B(\cdot)) = n$

(iii) $rk_{\mathcal{M}}K^{n-\ell}(A(\cdot),B(\cdot)) = n$

(iv) $\sum_{i=1}^{m} k_i = n$

Proof: (i) $\Leftrightarrow$ (ii) follows from Theorem 2.4 (iv). (ii) $\Leftrightarrow$ (iv) and (iii) $\Rightarrow$ (ii) are immediate. It remains to prove (ii) $\Rightarrow$ (iii) : Without restriction of generality assume $k_1 \geq 1,\ldots,k_\ell \geq 1, k_{\ell+1} = \ldots = k_m = 0$. The assumption that there exists $i \in \underline{m}$ such that $k_i > n-l+1$ leads to the contradiction $n = \sum_{i=1}^{l} k_i > l-1+n-l+1 = n$. Therefore $k_i \leq n-l+1$ for $i \in \underline{m}$ and (iii) is proved. $\square$

Brunovský (1970) has introduced a family of indices for time-varying systems $(A,B) \in (C^\infty)^{n\times(n+m)}$ as follows

$$r_i(t) := rk_{\mathbb{R}}K^i(A(t),B(t)) - rk_{\mathbb{R}}K^{i-1}(A(t),B(t)), \qquad i = 0,\ldots,n-1$$

where $K^{-1}(A,B) := 0$. Let $\alpha_i(t)$ denote the number of $r_j(t)'s$ which are bigger or equal to i, i.e.

$$\alpha_i(t) := \sum_{j:r_j(t)\geq i} 1 \quad \text{for } i \in \underline{m}.$$

If A and B are *analytic* matrices then $r_j(\cdot)$ and $\alpha_i(\cdot)$ are constants on $\mathbb{R}\setminus N$ for some discrete set N and

$$\begin{array}{lll} 0 \leq r_{n-1}(t) & \leq \ldots \leq & r_0(t) = rk_{\mathbb{R}}B(t) \leq m \\ 0 \leq \alpha_m(t) & \leq \ldots \leq & \alpha_1(t) \leq n \end{array} \qquad \text{for all } t \in \mathbb{R}\setminus N.$$

The functions $\alpha_1(\cdot),\ldots,\alpha_m(\cdot)$ are called the *geometric indices* of the system (A,B) . Again (4.1) yields that the geometric indices are invariant with respect to an analytic similarity action. If $k_1,\ldots,k_m$ denotes the controllability indices of $(A,B) \in \mathcal{A}^{n\times(n+m)}$ then for every interval $I \subset \mathbb{R}\setminus N$

$$\sum_{i=1}^{m} \alpha_i(t) = \sum_{j=0}^{n-1} \sum_{i:i\leq r_j(t)} 1 = \sum_{j=0}^{n-1} r_j(t) = rk_{\mathbb{R}}K(A(t),B(t)) = \sum_{i=1}^{m} k_i \qquad \text{for all } t \in I \tag{4.3}$$

It is not obvious how the controllability and geometric indices are related. By using the theory of $\mathcal{M}[D]$ right modules it will be shown in Section **2.6** that these families coincide.

Since the geometric indices are time-varying functions they contain more information about the system than the controllability indices do. More information than in the geometric indices is contained in the $r_i(\cdot)'s$. This illustrated in the following example.

Example 4.3 Consider the system given in Example 4.1. Then

$$r_0(t) = rk_{\mathbf{R}}B(t) = \begin{cases} 1 & \text{for} \quad t = 0 \\ 2 & \text{for} \quad t \neq 0 \end{cases}$$

$$\begin{aligned} r_1(t) &= rk_{\mathbf{R}}[B(t), \dot{B}(t)] - rk_{\mathbf{R}}B(t) = 3 - r_0(t) = \begin{cases} 2 & \text{for} \quad t = 0 \\ 1 & \text{for} \quad t \neq 0 \end{cases} \\ r_2(t) &= rk_{\mathbf{R}}K(A(t), B(t)) - rk_{\mathbf{R}}[B(t), \dot{B}(t)] = 0 \end{aligned}$$

and $(\alpha_1(t), \alpha_2(t), \alpha_3(t)) = (2, 1, 0)$.

In order to derive a canonical form for the analytic similarity action on $\mathcal{A}^{n\times(n+m)}$ a second family of indices will be defined. Suppose $(A, B) \in \mathcal{A}^{n\times(n+m)}$ is controllable with controllability indices $k_1, \ldots, k_m$. Then

$$H = [b_1, \ldots, (-1)^{k_1-1}(DI_n - A)^{k_1-1}(b_1), \ldots, b_m, \ldots, (-1)^{k_m-1}(DI_n - A)^{k_m-1}(b_m)] \in GL_n(\mathcal{M})$$

and $U \in \mathcal{M}^{n\times m}$ is uniquely defined by

$$[(DI_n - A)^{k_1}(b_1), \ldots, (DI_n - A)^{k_m}(b_m)] = HU \tag{4.4}$$

It follows from the construction of H, see (4.2), that $U = [u_1, \ldots, u_m]$ has a very special structure with many zero entries in it, namely

$$\left.\begin{aligned} u_i &= (u_1^i, \ldots, u_{i-1}^i, 0, \ldots, 0)^T \\ & u_\lambda^i = 0 \text{ if } k_\lambda = 0, \lambda < i \end{aligned}\right\} \text{ if } k_i = 0 \tag{4.5}$$

$$\left.\begin{aligned} u_i &= (u_{1,0}^i, \ldots, u_{1,k_1-1}^i, \ldots, u_{m,0}^i, \ldots, u_{m,k_m-1}^i)^T \\ & u_{j,\lambda}^i = 0 \begin{cases} \text{if} & \lambda = k_i \quad \text{and} \quad j > i \\ \text{if} & \lambda > k_i \end{cases} \end{aligned}\right\} \text{ if } k_i > 0. \tag{4.6}$$

By Lemma 1.2 U is invariant with respect to coordinate transformations $T \in GL_n(\mathcal{A})$.

Lemma 4.4 Suppose $(A, B) \in \mathcal{A}^{n\times(n+m)}$ is controllable with controllability indices $k_1, \ldots, k_m$ and H is given by (4.2). Then

$$H^{-1}[DI_n - A, -B] \begin{bmatrix} H & 0 \\ 0 & I_m \end{bmatrix} = [DI_n - A_c, B_c]$$

and A_c, B_c are in column form, i.e.

$$A_c = \left[\begin{array}{cccccccccc}
0 & & & * & & & & & & * \\
1 & \ddots & & \vdots & & & 0 & & & \vdots \\
 & \ddots & 0 & \vdots & & & & & & \vdots \\
 & & 1 & * & & & & & & \vdots \\
\cline{1-4}
 & & & * & & & & & & \vdots \\
 & & & \vdots & \ddots & & & & & \vdots \\
 & & & \vdots & & & & & & * \\
\cline{7-10}
 & & & \vdots & & & 0 & & & * \\
 & & & \vdots & & & 1 & \ddots & & \vdots \\
 & 0 & & \vdots & & & & \ddots & 0 & \vdots \\
 & & & * & & & & & 1 & *
\end{array}\right] \in \mathcal{M}^{n\times n}$$

where the diagonal blocks are $k_i \times k_i$ matrices, for $k_i > 0$, and the corresponding $*$-columns coincide with u_i as described in (4.6), resp.,

$$B_c = [b_1^c, \ldots, b_m^c]$$

$$b_i^c = \begin{cases} e_{s_i}, s_i = k_i + \ldots + k_{i-1} + 1, & \text{if } \; k_i > 0 \\ u_i \quad \text{as in (4.5)} & \text{if } \; k_i = 0 \end{cases}$$

where e_j denotes the j-th unit vector of $\mathbb{R}^n$, $k_{-1} := 0$.

Proof: Since $H \in GL_n(\mathcal{M})$

$$\mathcal{B} = \{b_1, \ldots, (-1)^{k_1-1}(DI_n - A)^{k_1-1}(b_1), \ldots, b_m, \ldots, (-1)^{k_m-1}(DI_n - A)^{k_m-1}(b_m)\}$$

is a basis of $\mathcal{M}[D]^n$, viewed as a right $\mathcal{M}[D]$ module. Now

$$[DI_n - A] : \mathcal{M}[D]^n \to \mathcal{M}[D]^n, \;\; v(D) \mapsto [DI_n - A] \cdot v(D)$$

is a $\mathcal{M}[D]$-right linear map and by the mulitplikation rule (1.4) one obtains for $v_i := (-1)^i(DI_n - A)^i(b_i)$

$$[DI_n - A] \cdot v_i = v_i D + \dot{v}_i - Av_i = v_i D + (DI_n - A)(v_i) = v_i D - v_{i+1} \tag{4.7}$$

The linear map $[DI_n - A]$ relative to the basis $\mathcal{B}$ is associated with the matirix

$$H^{-1}[DI_n - A]H = DI_n - A_c$$

and by (4.7) it is immediate that A_c has the form described in (i). It follows from the construction of H that $H^{-1}B = B_c$. $\square$

Note that in general A_c, B_c are not associated with a state space system since H may have meromorphic entries.

Proposition 4.5 Consider the class

$$\sum := \{(A,B) \in \mathcal{A}^{n\times(n+m)} \mid H \text{ defined in (4.2) belongs to } GL_n(\mathcal{A})\}$$

Then every $(A,B) \in \sum$ is analytically similar to $(A_c, B_c) \in \sum$ with the form as in Lemma 4.4. If $(A,B), (A',B') \in \sum$ are corresponding to $(A_c,B_c), (A'_c, B'_c) \in \sum$ resp. then

$$(A,B) \text{ analytically similar to } (A',B') \quad \Leftrightarrow \quad (A_c,B_c) = (A'_c,B'_c).$$

Proof: Since H associated with (A,B) belongs to $GL_n(\mathcal{A})$ it follows from Lemma 4.4 that $(A_c,B_c) \in \mathcal{A}^{n\times(n+m)}$. If (A,B) is similar to (A',B') then by (1.6) and Lemma 4.4 $(A_c,B_c) = (A'_c,B'_c)$. The opposite direction is trivial. This completes the proof. □

The previous proposition says that (A_c,B_c) is a *canonical form* for the analytic similarity action on $\sum$.

Remark 4.6
(i) For time-invariant controllable systems $(A,B) \in \mathbb{R}^{n\times(n+m)}$ *Popov* (1972) derived the analogous result to Proposition 4.5 in a complicated way.

(ii) Supppose $(A,B) \in \mathcal{A}^{n\times(n+m)}$ is controllable and the associated H defined in (4.2) is invertible over $\mathcal{A}$. Then it can be shown that (A,B) is analytically similar to some (A_r,B_r) in row form, cf. *Ilchmann* (1985a) or for an alternative but incomplete (see *Ilchmann* (1987)) proof *Nguyen* (1986). However (A_r,B_r) is not a canonical form. As opposed to the constant case (cf. *Kailath* (1980) Section 6.4) the proof of the row form is by far more tricky. For systems $(A,B) \in (C^\infty)^{n\times(n+m)}$ *Brunovský* (1970) derived the row form in a completely different way.

1.5 Time-varying subspaces, the controllable and the unreconstructible family

In this section time-varying subspaces are studied. This framework will be useful to tackle disturbance decoupling problems of time-varying systems in Chapter 3.

$\mathcal{V} = (\mathcal{V}(t))_{t\in\mathbb{R}}$ is called a *time-varying subspace* if $\mathcal{V}(t)$ is a subspace of $\mathbb{R}^n$ for every $t \in \mathbb{R}$. So $\mathcal{V}$ is a family of subspaces parameterized by $t \in \mathbb{R}$.
$\mathbf{W}_n$ denotes the set of all time-varying subspaces $\mathcal{V} = (\mathcal{V}(t))_{t\in\mathbb{R}}$ where $\mathcal{V}(t)$ is a subspace of $\mathbb{R}^n$ for every $t \in \mathbb{R}$.
If $\mathcal{V}(t)$ is given by

$$\mathcal{V}(t) = V(t)\mathbb{R}^k \quad , t \in \mathbb{R} \text{ where } V \in \mathcal{A}_p^{n\times k}$$

then $\mathcal{V}$ is called the time-varying subspace *generated* by V.

A problem arises: If $\mathcal{V} \in \mathbf{W}_n$ has a generator $V \in \mathcal{A}_p^{n\times k}$ then $\mathcal{V}^\perp := (\mathcal{V}(t)^\perp)_{t\in\mathbb{R}} \in \mathbf{W}_n$ does, in general, not have some piecewise analytic generator $W \in \mathcal{A}_p^{n\times k'}$. Consider for instance

$$\mathcal{V}(t) = t\cdot\mathbb{R}, \text{ then } \mathcal{V}(t)^\perp = \begin{cases} O & \text{if } t \neq 0 \\ \mathbb{R} & \text{if } t = 0 \end{cases}$$

which belongs to $\mathbf{W}_1$ but does not have a piecewise analytic generator. To cope with this equivalence classes are introduced:

Two families $\mathcal{V}_1, \mathcal{V}_2 \in \mathbf{W}_n$ are called *equal almost everywhere* (a.e.) on an interval $I \subset \mathbb{R}$

$$\mathcal{V}_1(t) \overset{ae}{=} \mathcal{V}_2(t) \quad \text{on } I$$

if $\mathcal{V}_1(t) = \mathcal{V}_2(t)$ for all $t \in I \setminus N$, where N denotes some discrete set.
In this sense one obtains for the preceeding example $\mathcal{V}(t)^\perp \overset{ae}{=} \{0\}$.
Analogously, one defines $\mathcal{V}_1$ is *included a.e.* in $\mathcal{V}_2$ on I .
I is omitted if $I = \mathbb{R}$.
The notation

$$\mathcal{V}_1(t) \overset{\subset}{\overset{ae}{=}} \mathcal{V}_2(t) \quad \text{on } I$$

is used if $\mathcal{V}_1(t) \subset \mathcal{V}_2(t)$ for all $t \in I$ and $\mathcal{V}_1(t) \overset{ae}{=} \mathcal{V}_2(t)$ on I .
"a.e. equivalent" is an equivalence relation on $\mathbf{W}_n$ and the equivalence class of $\mathcal{V} \in \mathbf{W}_n$ is denoted by

$$\bar{\mathcal{V}} = \{\mathcal{W} \in \mathbf{W}_n \mid \mathcal{V}(t) \overset{ae}{=} \mathcal{W}(t)\}$$

In order to show basic properties of time-varying subspaces some results concerning divisors and multiples of analytic matrices are proved.
Suppose $P \in \mathcal{A}^{n\times k}, Q \in \mathcal{A}^{n\times \ell}$. Then $G \in \mathcal{A}^{n\times r}$ is called a *greatest common left divisor* of P and Q, $G = gcld(P,Q)$ for short, if for every common left divisor G' of P and Q there exists an analytic matrix R of appropriate size such that $G'R = G$.
$K \in \mathcal{A}^{n\times s}$ is called a *least common right multiple* of P and Q, $K = lcrm(P,Q)$ for short, if for every common right multiple K' of P and Q there exists an analytic matrix S of appropriate size such that $K' = KS$.

A greatest common left divisor and least common right multiple of matrices over certain rings have been examined by several authors (see for example, *Mac Duffee* (1956)). Unfortunately their results are only valid for Euclidean domains or principal ideal domains; the set of real analytic functions is not a principal ideal domain, however it is a *Bezout ring*, i.e. if $f, g \in \mathcal{A}$ have no common zeros then there exists $a, b \in \mathcal{A}$ so that $af + bg = 1$, see *Narasimhan* (1985) Section 6.4. Nevertheless, the proof of the following lemma is partially based on Mac Duffee's ideas.

Lemma 5.1 Suppose $P \in \mathcal{A}^{n\times k}, Q \in \mathcal{A}^{n\times \ell}$ with $rk_{\mathcal{M}}P = k,\ rk_{\mathcal{M}}Q = \ell,\ rk_{\mathcal{M}}[P,Q] = r$. Then for $s := k + \ell - r$

(i) there exists $G = gcld(P,Q)$ with $rk_{\mathcal{M}}G = r$ which is unique up to multiplication by an invertible matrix from the right. Furthermore there exist analytic matrices U_1, U_3 of appropriate sizes such that

$$G = PU_1 + QU_3$$

and

$$G \cdot \mathcal{A}^r = P \cdot \mathcal{A}^k + Q \cdot \mathcal{A}^\ell$$

(ii) there exists $K = lcrm(P,Q)$ with $rk_{\mathcal{M}}K = s$ which is unique up to multiplication by an invertible matrix from the right and

$$K \cdot \mathcal{A}^s = P \cdot \mathcal{A}^k \cap Q \cdot \mathcal{A}^\ell$$

Proof: (i) By *Silverman and Bucy* (1970) [1] there exists

$$U = \begin{bmatrix} U_1 & U_2 \\ U_3 & U_4 \end{bmatrix} \in GL_{k+\ell}(\mathcal{A})$$

such that

$$[P, Q] \cdot U = [G, 0_{n \times s}], \quad rk_{\mathcal{M}} G = r \tag{5.1}$$

Let $V = \begin{bmatrix} V_1 & V_2 \\ V_3 & V_4 \end{bmatrix}$ be the inverse of U partitioned in such a form that

$$\begin{bmatrix} U_1 & U_2 \\ U_3 & U_4 \end{bmatrix} \begin{bmatrix} V_1 & V_2 \\ V_3 & V_4 \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & I_s \end{bmatrix}$$

Then $P = GV_1, Q = GV_2$ and G is a common left divisor of P and Q.
All matrices used in the following are defined over $\mathcal{A}$ and are of appropriate formats.
Now it is proved that G is a *greatest* common left devisor. Assume $P = \hat{G}W$, $Q = \hat{G}W'$ and $\hat{G} = GR$. Since $rk_{\mathcal{M}}\hat{G} \leq rk_{\mathcal{M}}G$ it is assumed without restriction of generality that $\hat{G}$ is a $n \times r$ matrix. By (5.1)

$$G = \hat{G}S, \quad \text{where } S := WU_1 + W'U_3$$

Thus $rk_{\mathbb{R}}G(t) = rk_{\mathbb{R}}\hat{G}(t)$ for all $t \in \mathbb{R}$. Let $I \subset \mathbb{R}$ be an open interval such that G is left invertible over $\mathcal{A}\,|_I$. Then $G(t) = G(t)R(t)S(t)$ for all $t \in I$ implies $I_r = R(t)S(t)$ for all $t \in I$. Since R and S are analytic $I_r = R(t)S(t)$ holds on $\mathbb{R}$. Therefore $G = gcld(P, Q)$ and the uniqueness statement is proved as well.

(ii) $K := PU_2 = -QU_4$ is a common right multiple of P and Q. At first it is proved that $rk_{\mathcal{M}}K = s$. Assume $rk_{\mathcal{M}}K < s$. Then there exists a $Z \in GL_n(\mathcal{A})$ such that $KZ = [\hat{K}, 0] = PU_2Z = -QU_4Z$. Since P and Q are left invertible on an open interval $I \subset \mathbb{R}$, U_2Z and $-U_4Z$ are of the form $[*, 0]$ on I. Therefore

$$U \cdot \begin{bmatrix} I_k & 0 \\ 0 & Z \end{bmatrix} = \begin{bmatrix} U_1 & [*,0] \\ U_2 & [*,0] \end{bmatrix}$$

which contradicts the invertibility of U on I.
Secondly it is proved that $K = lcrm(P, Q)$. Let

$$K' = PY = QY' \text{ and } \tilde{G} = gcld(K, K'),\ \tilde{G}H = K,\ \tilde{G}H' = K'.$$

Clearly $rk_{\mathcal{M}}\tilde{G} \geq s$. By (i) there exist N and N' such that

$$\tilde{G} = KN + K'N' \text{ and thus } \tilde{G} \cdot \mathcal{A}^{s'} \subset P \cdot \mathcal{A}^k \cap Q \cdot \mathcal{A}^\ell.$$

Since by (i) $\max_{t \in \mathbb{R}} \dim_{\mathbb{R}}[P(t) \cdot \mathbb{R}^k \cap Q(t) \cdot \mathbb{R}^\ell] = s$ we have $rk_{\mathcal{M}}\tilde{G} \leq s$. Therefore $rk_{\mathcal{M}}\tilde{G} = s$ and without restriction of generality let $\tilde{G}$ be a $n \times s$ matrix. From the equations above we compute

$$P[U_2N + YN'] = KN + K'N' = \tilde{G} = Q[-U_4N + Y'N'].$$

Let $E := U_2N + YN'$ and $F := -U_4N + Y'N'$, then

$$-QU_4 = PU_2 = K = \tilde{G}H = PEH = QFH.$$

[1] *Wedderburn* (1915) proves that a matrix over the ring of holomorphic functions can be transformed into a diagonal matrix by unimodular matrix operations, cf. *Narasimhan* (1985). This result is also valid for matrices over the ring of real analytic functions. However, in the following we will quote *Silverman and Bucy* (1970). This weaker result is sufficient for our purposes.

Since Q and P are left invertible on some open interval $I \subset \mathbb{R}$, $-U_4 = FH$ and $U_2 = EH$ on I. Therefore

$$I_s = V_3 U_2 + V_4 U_4 = V_3 EH - V_4 FH = (V_3 E - V_4 F)H \qquad \text{on } I$$

and since all involved matrices are analytic H is invertible over $\mathcal{A}$. Thus $K' = \tilde{G}H' = KH^{-1}H'$. Using similar arguments one can prove that also H' is invertible over $\mathcal{A}$, whence $K = \tilde{G}H = K'H'^{-1}H$. This completes the proof. □

Remark 5.2 It is also possible to define and to show the existence of a $gcld_{\mathcal{A}_p}(P,Q)$ and a $lcrm_{\mathcal{A}_p}(P,Q)$ for matrices P and Q defined over $\mathcal{A}_p$ instead of $\mathcal{A}$.
This is demonstrated for a $gcld_{\mathcal{A}_p}$ of $p, q \in \mathcal{A}_p$. Suppose $\mathbb{R} = \cup_{\nu \in \mathbb{Z}}[a_\nu, a_{\nu+1})$ is a disjoint partition such that $p_\nu, q_\nu \in \mathcal{A}\mid_{[a_\nu,a_{\nu+1})}$ have analytic extensions on both sides of $(a_\nu, a_{\nu+1})$, see Section 1.1. For short, put

$$f_\nu := f\mid_{[a_\nu,a_{\nu+1})} \text{ for } f = g \text{ or } f = h.$$

Let

$$g_\nu := gcrd(p_\nu, q_\nu) \in \mathcal{A}\mid_{[a_\nu,a_{\nu+1})} \text{ and } g_\nu = p_\nu c_\nu + q_\nu d_\nu \text{ for } a_\nu, d_\nu \in \mathcal{A}\mid_{[a_\nu,a_{\nu+1})}, \nu \in \mathbb{Z}.$$

Now it is straightforward to prove that g defined by $g\mid_{[a_\nu,a_{\nu+1})} := g_\nu$ is a $gc\ell d_{\mathcal{A}_p}(p,q) \in \mathcal{A}_p$.

Using Remark 5.2 it is immediate that the statements of Lemma 5.1 can be extended to piecewise analytic matrices as follows

Lemma 5.3 Suppose $P \in \mathcal{A}_p^{n\times k}$ and $\mathbb{R} = \cup_{\nu\in\mathbb{Z}}[a_\nu, a_{\nu+1})$ is a disjoint partition so that $P\mid_{[a_\nu,a_{\nu+1})}$, $Q\mid_{[a_\nu,a_{\nu+1})}$ have real analytic extensions on some $(a_\nu^\ell, a_{\nu+1}^r)$, $a_\nu^\ell < a_\nu$, $a_{\nu+1} < a_{\nu+1}^r$. Then

(i) there exists $G = gcld_{\mathcal{A}_p}(P,Q) \in \mathcal{A}_p^{n\times n}$ so that

$$G\mid_{[a_\nu,a_{\nu+1})} \text{ is of the form } [*, 0_{n\times s_0}]$$

where

$$rk_{\mathcal{M}} P\mid_{[a_\nu,a_{\nu+1})} = k_\nu,\ \ rk_{\mathcal{M}} Q\mid_{[a_\nu,a_{\nu+1})} = \ell_\nu,\ \ rk_{\mathcal{M}}[P,Q]\mid_{[a_\nu,a_{\nu+1})} = r_\nu$$

and $s_\nu := k_\nu + \ell_\nu - r_\nu$.
Furthermore there exist $U_1 \in \mathcal{A}_p^{k\times n}, U_3 \in \mathcal{A}_p^{\ell\times n}$ so that

$$G = PU_1 + QU_3.$$

(ii) there exists $K = lcrm_{\mathcal{A}_p}(P,Q) \in \mathcal{A}_p^{n\times n}$ with $rk_{\mathcal{M}} K\mid_{[a_\nu,a_{\nu+1})} = s_\nu$ and

$$K \cdot \mathcal{A}_p^n = P \cdot \mathcal{A}_p^k + Q \cdot \mathcal{A}_p^\ell$$

Lemma 5.4 Let $C \in \mathcal{A}^{p\times n}$, $V \in \mathcal{A}^{n\times k}$, $\mathcal{V}(t) = V(t)\mathbb{R}^k$.
Then there exist real analytic matrices $\hat{V}, \hat{U}, \hat{C}, \hat{W}$ of formats $n\times k$, $n\times(n-\ell)$, $n\times s$, $n\times s'$, resp. which have constant ranks and satisfy

(i) $\mathcal{V}(t) \overset{\subset}{\underset{ae}{=}} \hat{V}(t)\mathbb{R}^\ell \qquad , \ell = rk_{\mathcal{M}} V$

(ii) $\mathcal{V}(t)^{\perp} \overset{\supset}{\stackrel{ae}{=}} \hat{U}(t)\mathbb{R}^{n-\ell}$

(iii) $kerC(t) \overset{\supset}{\stackrel{ae}{=}} \hat{C}(t)\mathbb{R}^{s} \qquad , s = n - rk_{\mathcal{M}}C = rk_{\mathcal{M}}\hat{C}$

(iv) $\mathcal{V}(t) \cap kerC(t) \stackrel{ae}{=} \hat{W}(t)\mathbb{R}^{s'} \qquad , s' = rk_{\mathcal{M}}\hat{C} + rk_{\mathcal{M}}V - rk_{\mathcal{M}}[\hat{C}, V]$

Proof: (i) By *Silverman and Bucy* (1970) there exists $S \in GL_n(\mathcal{A})$ so that

$$V^T \cdot S = [V_1, 0], \ V_1 \in \mathcal{A}^{k \times \ell}, \ rk_{\mathcal{M}}V_1 = \ell$$

Therefore

$$(V(t)\mathbb{R}^k)^{\perp} \overset{\supset}{\stackrel{ae}{=}} S(t) \begin{bmatrix} 0 \\ I_{n-\ell} \end{bmatrix} \mathbb{R}^{n-\ell}$$

and

$$\hat{V}(t) := S^{T^{-1}}(t) \begin{bmatrix} I_\ell \\ 0 \end{bmatrix}$$

has constant rank and satisfies (i).
(ii) is valid for $\hat{U}(t) := S(t) \begin{bmatrix} 0 \\ I_{n-\ell} \end{bmatrix}$.
(iii) Let $R \in GL_n(\mathcal{A})$ so that

$$C(t)R(t) = [C_1(t), 0], \ C_1 \in \mathcal{A}^{p \times (n-s)}, \ rk_{\mathcal{M}}C_1 = n - s$$

Clearly,

$$\hat{C}(t) := R(t) \begin{bmatrix} 0 \\ I_s \end{bmatrix}$$

satisfies (iii).
(iv) Use Lemma 5.1 (ii) to determine $W := lcrm(\hat{C}, V)$ with $rk_{\mathcal{M}}W = s'$. Now by (i) one can choose $\hat{W} \in \mathcal{A}^{n \times s}$ so that (iv) holds true. □

In order to characterize when the rank of $V(\cdot) \in \mathcal{A}^{n \times k}$ is constant in t the following definition is needed.

Definition 5.5 For a family $\mathcal{V} \in \mathbf{W}_n$ let

$$P(t) : \mathbb{R}^n \to \mathcal{V}(t)$$

be the orthogonal projector on $\mathcal{V}(t)$ along $\mathcal{V}(t)^{\perp}$. $\mathcal{V}$ is called an *analytic family* if $P \in \mathcal{A}^{n \times n}$ resp. a *piecewise analytic* (p.a.)*family* if $P \in \mathcal{A}_p^{n \times n}$.

Note that analyticity of $V \in \mathcal{A}^{n \times k}$ does not ensure that the family $\mathcal{V}$ generated by V is an analytic family, consider for instance $V(t) = t$.

Proposition 5.6 If $\mathcal{V} \in \mathbf{W}_n$ is generated by $V \in \mathcal{A}^{n \times k}$ then $\mathcal{V}$ is an analytic family if and only if

$$rk_{\mathbb{R}}V(t) = \text{ const.} \quad \text{for all } t \in \mathbb{R}.$$

Proof: If the orthogonal projector $P(t)$ on $\mathcal{V}(t)$ is real analytic in $t \in \mathbb{R}$ then by Corollary A.5 in *Gohberg, Lancaster and Rodman* (1983) the function $t \mapsto rk_{\mathbf{R}}V(t)$ is constant (continuity of $P(\cdot)$ is already sufficient). Conversely, if $t \mapsto rk_{\mathbf{R}}V(t)$ is constant on $\mathbb{R}$ then by Proposition A.11 in *Gohberg, Lancaster and Rodman* (1983) $\mathcal{V}$ is an analytic family. □

Proposition 5.6 will be extended to the piecewise analytic situation. For this a definition is necessary.

Definition 5.7 $V \in \mathcal{A}_p^{n\times k}$ is said to have *piecewise constant (p.c.) rank* if there exists a disjoint partition $\mathbb{R} = \cup_{\nu\in\mathbb{Z}}[a_\nu, a_{\nu+1})$ so that each restriction

$$\begin{aligned}&V\mid_{(a_\nu,a_{\nu+1})} \text{ is real analytic}\\ &\text{and has a real analytic extension}\\ &V_\nu \text{ on some } (a_\nu^\ell, a_{\nu+1}^r),\ a_\nu^\ell < a_\nu,\ a_{\nu+1} < a_{\nu+1}^r\\ &\text{and}\\ &rk_{\mathbf{R}}V_\nu(t) = \text{ const. for all } t \in (a_\nu^\ell, a_{\nu+1}^\ell).\end{aligned}$$

Proposition 5.8 If $\mathcal{V} \in \mathbf{W}_n$ is generated by $V \in \mathcal{A}_p^{n\times k}$ then $\mathcal{V}$ is a p.a. family if and only if V has p.c. rank.

Proof: If $\mathcal{V}$ is a p.a. family then there exists a partition $\mathbb{R} = \cup_{\nu\in\mathbb{Z}}[a_\nu, a_{\nu+1})$ so that each restriction $V\mid_{[a_\nu,a_{\nu+1})}, P\mid_{[a_\nu,a_{\nu+1})}$ is real analytic and has a real analytic extension V_ν, P_ν resp. on some $(a_\nu^\ell, a_{\nu+1}^r)$ where $a_\nu^\ell < a_\nu$, $a_{\nu+1} < a_{\nu+1}^r$. Now it follows from Proposition 5.6 that for each $\nu \in \mathbb{Z}$, $rk_{\mathbf{R}}V_\nu(t) = \text{const.}$ for all $t \in (a_\nu^\ell, a_{\nu+1}^r)$. This proves that V has p.c. rank.
The opposite direction follows by reversing the foregoing arguments. □

Proposition 5.9 Let $C \in \mathcal{A}_p^{p\times n}, V \in \mathcal{A}_p^{n\times k}$. Then there exist $\hat{V} \in \mathcal{A}_p^{n\times k}$ and $\hat{U}, \hat{C}, \hat{W} \in \mathcal{A}_p^{n\times n}$ with p.c. ranks so that

(i) $V(t)\mathbb{R}^k \overset{\subseteq}{\underset{ae}{=}} \hat{V}(t)\mathbb{R}^n$

(ii) $(V(t)\mathbb{R}^k)^\perp \overset{\supseteq}{\underset{ae}{=}} \hat{U}(t)\mathbb{R}^n$

(iii) $kerC(t) \overset{\supseteq}{\underset{ae}{=}} \hat{C}(t)\mathbb{R}^n$

(iv) $V(t)\mathbb{R}^k \cap kerC(t) \overset{ae}{=} \hat{W}(t)\mathbb{R}^{s'}$

Proof: To prove (i) choose an interval $[a_\nu, a_{\nu+1})$ so that $V\mid_{[\,a_\nu,a_{\nu+1})}$ is real analytic and has a real analytic extension V_ν on $(a_\nu^\ell, a_{\nu+1}^r)$. Then by Lemma 5.4 (i) there exists $\hat{V}_\nu \in \mathcal{A}\mid^{n\times n}_{(a_\nu^\ell,a_{\nu+1}^r)}$ with constant rank so that

$$V_\nu(t)\mathbb{R}^k \overset{\subseteq}{\underset{ae}{=}} \hat{V}_\nu(t)\mathbb{R}^n \qquad \text{on } (a_\nu^\ell, a_{\nu+1}^r)$$

Since this can be done for every interval of the partition corresponding to $V \in \mathcal{A}_p^{n\times k}$ (i) is proved. For the proof of (ii) - (iv) use the similar arguments. □

Proposition 5.10

(i) Suppose $\mathcal{V} \in \mathbf{W}_n$ is generated by $V \in \mathcal{A}^{n\times k}$ and $rk_{\mathbb{R}}V(t)$ is constant in $t \in \mathbb{R}$. If $v \in \mathcal{A}^n$ satisfies

$$v(t) \in \mathcal{V}(t) \quad \text{for all } t \in \mathbb{R} \setminus N, \text{ where } N \text{ is a discrete set} \tag{5.2}$$

then there exists $r \in \mathcal{A}^k$ so that

$$v(t) = V(t)r(t) \quad \text{for all } t \in \mathbb{R} \tag{5.3}$$

and thus

$$v(t) \in \mathcal{V}(t) \quad \text{for all } t \in \mathbb{R}.$$

(ii) Suppose $\mathcal{V} \in \mathbf{W}_n$ is generated by $V \in \mathcal{A}_p^{n\times k}$ and V has p.c. rank. If $v \in \mathcal{A}_p^n$ satisfies (5.2) then (5.3) is valid for some $r \in \mathcal{A}_p^k$.

Proof: (i) Let $\ell \in \mathbb{N}$ so that $rk_{\mathbb{R}}V(t) = \ell$ for all $t \in \mathbb{R}$. Then by *Silverman and Bucy* (1970) there exists $S \in GL_k(\mathcal{A})$ such that

$$VS^{-1} = [W, 0] \quad \text{for some } W \in \mathcal{A}^{n\times \ell} \text{ with } rk_{\mathbb{R}}W(t) = \ell \quad \text{for all } t \in \mathbb{R}.$$

Put

$$r = S^{-1}\begin{bmatrix} r' \\ 0_{k-\ell} \end{bmatrix}$$

where $r' := W^T(WW^T)^{-1}v$ then r satisfies (5.3).

(ii) Use the notation of Definition 5.7. It is sufficient to prove the assertion on some $(a_\nu^\ell, a_{\nu+1}^\ell)$ where $rk_{\mathbb{R}}V_\nu(t)$ is constant. Then (ii) follows from (i). □

Time-varying subspaces arise when controllability subspaces of time-varying systems are considered. This will be described in the remainder of this section.

For systems $(A, B) \in \mathcal{C}_p^{n\times(n+m)}$ the following is well-known (see e.g. *Kalman* (1960)): There exists a control $u \in \mathcal{C}_p^m$ which forces the state $x_0 \in \mathbb{R}^n$ at time t_0 to zero in time $t_1 - t_0 > 0$, i.e.

$$\Phi(t_1, t_0)x_0 + \int_{t_0}^{t_1} \Phi(t_1, s)B(s)u(s)ds = 0,$$

if and only if $x_0 \in imW(t_0, t_1)$. In terms of Definition 2.1 this means that the free trajectory $\Phi(\cdot, t_0)x_0$ is controllable at time t_0 into $\mathcal{V}_1 = \{0\}$. Thus

$$\Re(t_0) := \bigcup_{t_1 > t_0} imW(t_0, t_1)$$

is the vector space of all states which can be controlled at time t_0 to zero in finite time. We call

$$\Re = (\Re(t))_{t\in\mathbb{R}}$$

the *controllable family* of the system $(A, B) \in \mathcal{C}_p^{n\times(n+m)}$.
Clearly,

$$\Re(t_0) \subset \Phi(t_0, t_{-1})\Re(t_{-1}) \quad \text{for } t_{-1} \leq t_0.$$

If $(A, B) \in \mathcal{A}^{n\times(n+m)}$ is a real analytic system then

$$rk_{\mathbb{R}} W(t_0, t) = \text{ const.} \quad \text{for all } t \in \mathbb{R} \setminus \{t_0\}$$

Thus

$$\Re(t) = \Phi(t, t_0)\Re(t_0) \quad \text{for all } t, t_0 \in \mathbb{R}. \tag{5.4}$$

Now Proposition 5.6 and formulae (5.4) and (2.6) yield

Remark 5.11 The controllable family $\Re = (\Re(t))_{t\in\mathbb{R}}$ of an analytic system $(A, B) \in \mathcal{A}^{n\times(n+m)}$ is an analytic family given by

$$\Re(t) = \sum_{i\geq 0} im(DI_n - \mathcal{A}(t))^i(B(t)) \qquad , t \in \mathbb{R}. \tag{5.5}$$

It is also well-known that the state $x_0 \in \mathbb{R}^n$ at time t_0 is unreconstructible iff $x_0 \in \ker H(t_{-1}, t_0)$ for all $t_{-1} < t_0$. Thus

$$\mathcal{B}(t_0) = \bigcap_{t_{-1}<t_0} \ker H(t_{-1}, t_0)$$

denotes the vector space of the unreconstructible states at time t_0. We have the following dual relationsships.

Proposition 5.12 Let $\Re^d(t_0)$ resp. $\mathcal{B}^d(t_0)$ denote the controllable resp. unreconstructible subspace of the dual system of $(A, B) \in \mathcal{C}_p^{n\times(n+m)}$ with respect to $t^* = 2t_0$. Then

(i) $\Re(t_0)^\perp = \mathcal{B}^d(t_0)$

(ii) $\mathcal{B}(t_0)^\perp = \Re^d(t_0)$

Proof: Only (i) is proved , the proof of (ii) is entirely similar. Since

$$\Phi^d(t, t_0) = \Phi^T(t^* - t_0, t^* - t)$$

is the transition matrix of the dual system, for arbitrary $t_1 > t_0$ and $x \in \mathbb{C}^n$ we have

$$\begin{array}{lll}
& x \perp \Re(t_0) & \\
\Leftrightarrow & & \\
& x^T \int_{t_0}^{t_1} \Phi(t_0, s)B(s)u(s)ds = 0 & \text{for all } u(\cdot) \in \mathcal{C}_p^m \\
\Leftrightarrow & & \\
& B^T(-s)\Phi^T(t_0, -s)x = 0 & \text{for all } s \in [-t_1, -t_0] \\
\Leftrightarrow & & \\
& B^T(t^* - s)\Phi^d(s, t^* - t_0)x = 0 & \text{for all } s \in [t^* - t_1, t_0] \\
\Leftrightarrow & & \\
& x \in H^d(2t_0 - t_1, t_0) &
\end{array}$$

□

Remark 5.13 (i) If $(A,B) \in \mathcal{A}^{n\times(n+m)}$ then by Proposition 5.12 and (5.5) one obtains the simple presentation

$$\begin{aligned}\mathcal{B}(t_0) &= \Re^d(t_0)^{\perp} \\ &= [\sum_{i\geq 0} im(DI_n - A^T(2t_0 - t_0))^i(C^T(2t_0 - t_0))\,]^{\perp} \\ &= \bigcap_{i\geq 0}[\,im(DI_n - A^T(t_0))^i(C^T(t_0))]^T \\ &= \bigcap_{i\geq 0}\ker[(DI_n - A^T(t_0))^i(C^T(t_0))]^T \end{aligned} \tag{5.6}$$

(ii) For time-invariant systems (5.6) reduces to the well-known result that the unreconstructible resp. unobservable subspace is given by

$$\bigcap_{i\geq 0}\ker CA^{i-1}.$$

Chapter 2

Differential Polynomial Matrix Systems - An Algebraic Approach

2.0 Introduction

Equations of a physical system are usually not in state space form and it may not be obvious how they can be brought to this form. For this reason *Rosenbrock* (1970) proposed the well-known setting of systems in differential operator description

$$P(D)(z) = Q(D)(u) \qquad (0.1)$$

$$y = V(D)(z) + W(D)(u) \qquad (0.2)$$

where the entries of the matrices are polynomials in D (the usual differential operator) with *real* coefficients. Since there is some free choice in selecting the internal variables z of such a system the question arises under which conditions two systems of the form (0.1) have the same dynamics and the same input-output behaviour. This is the problem of (strict) *system equivalence* already studied by *Rosenbrock* (1970). *Wolovich* (1974) further developed the polynomial approach. Via module theoretic tools *Fuhrmann* (1976) and (1977) was able to associate a canonical state space model with any factorization $V(z)P(z)^{-1}Q(z) + W(z)$ of a proper rational transfer matrix. So far the analysis of the problem of system equivalence for time-invariant systems was done in the *frequency domain*. *Pernebo* (1977) was the first who studied system equivalence in the *time domain*, his basic idea was to consider solution sets of the system equations. This approach was systematically exploited by *Hinrichsen and Prätzel-Wolters* (1980) to obtain a self-contained theory of system equivalence in the time domain. They derived an *algebraic criterion* of system equivalence, defined and characterized *controllability* and *observability*, and presented a *canonical state space model* similar to Fuhrmann's model.

For a long time there has been a widespread scepticism whether an algebraic treatment in the style of *Kalman*, i.e. a module theoretic framework, would at all be possible for time-varying systems. In the second half of the seventies there were some attempts to introduce *time-varying* systems of the form (0.1), where the entries of the differential polynomial matrices are usually elements of some skew polynomial ring $\mathcal{M}[D]$ and the coefficients belong to some differentially closed ring of functions $\mathcal{M}$ or generalizations of such a ring. The choice of $\mathcal{M}$ represents a main decision with regard to the chances for a successful treatment of systems described by (0.1) and to the applicability of the results.
Ylinen (1975) collected basic algebraic results necessary for an analysis of equation (0.1) in case where $\mathcal{M}$ is a ring of endomorphisms. He also discussed basic system theoretic problems. However, concrete results suffer from restrictive assumptions which in situations of interest turn out

to be unrealistic in the time-varying case.
Kamen (1976) assumed for his main result, that $\mathcal{M}$ is Noetherian. Under this hypothesis he constructed a state space representation for (0.1) with monic $P(D)$. The Noether condition seems to be rather restrictive (see examples given by *Kamen* (1976)). The ring of analytic functions is not Noetherian.
In another report of *Ylinen* (1980) he concentrated mainly on the situation where $\mathcal{M}$ is a subring of C^∞. He also treated controllability, a coprimeness criterion similar to the one known from the time-invariant case was approached and partially established. The main restriction required for his substantial results are: $\mathcal{M}$ must not contain zero-divisors of C^∞ and the composite matrix $[P(D), -Q(D)]$ and all its right factors of the same format must be row equivalent to a matrix in upper triangular form with coefficients also in $\mathcal{M}$ and monic diagonal elements.
In *Ilchmann, Nürnberger and Schmale* (1984) we were guided by the time-invariant approach of *Hinrichsen and Prätzel-Wolters* (1980). We chose $\mathcal{M}$ to be the field of fraction of real meromorphic functions and considered *"full"* operators $P(D)$, i.e. every local analytic solution f of $P(D)(f) = 0$ can be continued to a global solution of $P(D)(f) = 0$. Analytic systems considered by *Ylinen* (1980) and constant systems in differential operator descriptions introduced by *Rosenbrock* (1970) fulfill these assumptions. Furthermore the assumptions set us in a position to present a far reaching algebraic analysis of systems of the form (0.1). The results of *Hinrichsen and Prätzel-Wolters* (1980) were generalized. This is presented in the first half of the present chapter.
A different algebraic approach to various definitions of *structural indices* of time-invariant state space systems was introduced by *Münzner and Prätzel-Wolters* (1979). Using polynomial modules and their minimal bases they proved the equality of *controllability indices, minimal indices, geometric indices* and *dynamical indices. Prätzel-Wolters* (1981) continued this approach to characterize Brunovský-equivalence for time-invariant systems of the form (0.1), (0.2). Guided by this approach and using the skew polynomial ring introduced in *Ilchmann, Nürnberger and Schmale* (1984) I generalized the results of *Münzner and Prätzel-Wolters* (1979) and *Prätzel-Wolters* (1981) for time-varying systems (see *Ilchmann* (1985a)). The characterization of *minimal bases* of *right skew polynomial modules* extended a result of *Forney* (1975). It is possible to define a *transfer matrix* in the time domain and to use this to characterize system equivalence. Different *invariants* with respect to system equivalence resp. similarity were defined and their equality was shown. This is presented in the second half of this chapter.

In Section 1 matrices over the skew polynomial ring $\mathcal{M}[D]$ are analysed and the lattice of full polynomial matrices is established. The basic idea of considering matrices defined over $\mathcal{M}[D]$, where $\mathcal{M}$ is the field of real meromorphic functions, and assuming that $P(D)$ is full, makes an algebraic study of systems of the form (0.1), (0.2)possible.
In Section 2 solution vector spaces associated with (0.1) are studied. Using this, system equivalence is defined and algebraically characterized. It is shown that every system of the form (0.1), (0.2) is system equivalent to an analytic state space system.
The results of Section 1 to 3 are complete generalizations of the time-invariant case, see *Hinrichsen and Prätzel-Wolters* (1980).
Although for time-varying systems there is no transform technique, in Section 4 a formal transfer matrix is defined as a matrix over the left-skew field of fractions of $\mathcal{M}[D]$. This matrix is as powerful as the input-output map in the time-domain.
Instead of the differential equation (0.1), *Münzner and Prätzel-Wolters* (1979) considered in the time-invariant case the algebraic equation $P(D)z(D) = Q(D)u(D)$. In Section 5 this is extended to the present setting. $\mathcal{M}[D]$–right modules of $\mathcal{M}[D]^r$ and their minimal bases (see *Forney* (1975) for commutative rings $F[D]$, F a field) are analysed. In particular the input module of a system (0.1) is studied. This module is invariant with respect to system equivalence.

In Section 6 the question (posed in Section 1.4), how the controllability - and geometric indices are related, is answered. Dynamical indices are defined via the formal transfer matrix (see *Forney* (1975) for the constant case) and by use of the input module it will be proved that all indices (roughly speaking) coincide.
In Section 7 system equivalence is characterized via the input module and the formal transfer matrix.

2.1 Differential polynomial matrices

For an algebraic study of time-invariant polynomial matrix systems the solution module $\ker P(D)$, $P(D) \in \mathbb{R}[D]^{r\times r}$, turned out to be very useful, see *Hinrichsen and Prätzel-Wolters* (1980). In order to extend this approach to time-varying systems I introduce

$$\ker_{\mathcal{F}_I} P(D) = \{f \in \mathcal{F}_I^r \mid P(D)(f) = 0\}, \qquad P(D) \in \mathcal{M}[D]^{r\times r}$$

where $I \subset \mathbb{R}$ is some open interval and $\mathcal{F}_I(\mathcal{F} = \mathcal{A}$ or $\mathcal{M})$ denotes the algebra of real-analytic or meromorphic functions on I. We omit I if $I = \mathbb{R}$.
Firstly the *scalar* case is discussed. There are considerable differences to time-invariant polynomials. If $p(D) \in \mathbb{R}[D]$, i.e. $p(D)$ has *constant* coefficients, it is well-known that $\dim \ker_{\mathcal{F}_I} p(D) = \deg p(D)$. This is, in general, not true for polynomials $p(D) \in \mathcal{M}[D]$. Consider for instance $p_1(D) = tD + 1$, then $\ker_{\mathcal{M}} p_1(D) = < 1/t >_{\mathbb{R}}$ and $\ker_{\mathcal{A}} p_1(D) = \{0\}$. Moreover there are polynomials for which even the dimension of the kernel over $\mathcal{M}$ does not coincide with the degree of the polynomial: if $p_2(D) = t^2D + 1$ and $0 \notin I$ then $\ker_{\mathcal{M}_I} p_2(D) = < e^{1/t} >_{\mathbb{R}}$ whereas $\ker_{\mathcal{M}} p_2(D) = \{0\}$. Since for every $p(D) = p_0 + \ldots + p_nD^n \in \mathcal{M}[D]$ there exists an interval $I \subset \mathbb{R}$ such that the numerators and denumerators of the p_i's do not have zeros on I one obtains $\dim \ker_{\mathcal{A}_I} p(D) = \deg p(D)$. By enlarging the interval I one might loose a meromorphic solution, as illustrated by the preceeding example. So in general

$$\dim \ker_{\mathcal{F}_I} p(D) \leq \deg p(D) \tag{1.1}$$

This leads to the following definition.

Definition 1.1 $p(D) \in \mathcal{M}[D]$ is called full wrt $\mathcal{F}$ ($\mathcal{F}$ or $\mathcal{A}$ or $\mathcal{M}$) if $p \neq 0$ and $\dim \ker_{\mathcal{F}} p(D) = \deg p(D)$.

It is immediate from the definition that the concept of full polynomials can be characterized as follows.

Proposition 1.2 $p \in \mathcal{M}[D], p \neq 0$ is full wrt $\mathcal{F}$ iff the map

$$\begin{array}{rcl} \pi_I : \ker_{\mathcal{F}} p(D) & \to & \ker_{\mathcal{F}_I} p(D) \\ f & \mapsto & f\,|_I \end{array}$$

is an isomorphism for every open interval $I \subset \mathbb{R}$.

Thus a polynomial $p(D)$ is full wrt $\mathcal{F}$ if any local solution f of $p(D)(f) = 0$ on I can be analytically resp. meromorphically continued to a global solution. Every $p(D) \in \mathbb{R}[D]$ or monic $p \in \mathcal{A}[D]$ is a full polynomial *wrt*$\mathcal{A}$. There are non-monic polynomials $p \in \mathcal{A}[D]$ which are full, consider for example $p(D) = tD - 1$ with $\ker_{\mathcal{A}} p(D) = < t >_{\mathbb{R}}$.

Proposition 1.3 Suppose $p, q, g \in \mathcal{M}[D]$ satisfy $p = qg$ and p is full wrt $\mathcal{F}$. The g is full wrt $\mathcal{F}$ and q is full wrt $\mathcal{M}$.

Proof: Choose $I \subset \mathbb{R}$ sufficiently small so that p, q, g are full $wrt\mathcal{A}_I$ and $\mathcal{A}_I \subset img$. Put $\ker_{\mathcal{A}_I} p(D) = \ker_{\mathcal{A}_I} g(D) \oplus V$, where V is some complementing vector space. Now $g(D)(V) = \ker_{\mathcal{A}_I} q(D)$ and g is injective on V. Since p is full $wrt\mathcal{F}$ all solutions of $\ker_{\mathcal{A}_I} g(D)$ extend to solutions of $\ker_{\mathcal{F}} g(D)$ and all solutions in $g(D)(V)$ extend to solutions in $\ker_{\mathcal{M}} q(D)$. This completes the proof. □

The set of full polynomials does not form a multiplicative semigroup. Consider for example $p(D) = tD + 1$ and $q(D) = D$ which are full $wrt\mathcal{M}$ since its solutions are $1/t$ resp. 1 . However $\ker_{\mathcal{M}_I} pq = \ker_{\mathcal{M}_I} (D^2 + \frac{t^2-1}{t} D) = < 1, \ell n \mid t \mid >_{\mathbb{R}}$ for every interval I with $0 \notin I$.

Since there exist a left and a right division algorithm for polynomials in $\mathcal{M}[D]$, it can be shown (see *Ore* (1933) pp.483) that for any $p, q \in \mathcal{M}[D]$ there exist a *greatest common right divisor* $g = gcrd(p, q) \in \mathcal{M}[D]$ and a *least common left multiple* $\ell = \ell c\ell m(p, q) \in \mathcal{M}[D]$ ($gcrd$ and $\ell c\ell m$ over $\mathcal{M}[D]$ are defined analogously as over $\mathcal{A}^{n \times k}$, see Section 1.5). g and ℓ are unique if they are required to be monic. *Ore* (1933) has also proved the existence of $a, b \in \mathcal{M}[D]$ such that

$$g = ap + bq$$

and

$$\deg\, p + \deg\, q = \deg\, \ell + \deg\, g \tag{1.2}$$

This is an extension of the results known for $\mathbb{R}[D]$ since for every $p, q \in \mathbb{R}[D]$ we can show that the *greatest common divisor* and the *least common multiple* of p and q denoted by $gcd_{\mathbb{R}[D]}(p, q)$ resp. $\ell cm_{\mathbb{R}[D]}(p, q)$ coincide with $gcrd(p, q)$ and $\ell c\ell m(p, q)$, resp. Put $p = p'g,\ q = q'g$ such that $g = gcd(p', q') \in \mathbb{R}[D]$ and p', q' are coprime over $\mathbb{R}[D]$. Then p', q' satisfy the *Bezout equation*, i.e.

$$1 = ap' + bq' \quad \text{for some} \quad a, b \in \mathbb{R}[D],$$

hence $p'q'$ are right coprime over $\mathcal{M}[D]$ as well and $g = gcrd(p', q')$. The same holds true for $\ell cm_{\mathbb{R}[D]}(p, q)$. This yields $\ell cm_{\mathbb{R}[D]}(p, q) = gp'q'$. Since

$$\deg\, \ell c\ell m(p, q) = \deg\, p + \deg\, q - \deg\, g = \deg\, (gp'q') = \deg\, \ell cm_{\mathbb{R}[D]}(q, q)$$

it follows that $gp'q' = \ell c\ell m(p, q)$.

Before further properties of the $gcrd$ and $\ell c\ell m$ are stated a basic lemma is shown. This was already known to *Schlesinger* (1895) p. 81.

Lemma 1.4 Let $p(D) \in \mathcal{M}[D]$ and $0 \neq f \in \ker_{\mathcal{M}} p(D)$. Then there exist a $r(D) \in \mathcal{M}[D]$ such that $p(D) = r(D) \cdot (fD - \dot{f})$.

Proof: (i) The right Euclidean algorithm leads to

$$p(D) = r(D) \cdot (fD - \dot{f}) + s \quad \text{for some} \quad s \in \mathcal{M}, r(D) \in \mathcal{M}[D].$$

Now $p(D)(f) = 0$ implies $s = 0$. □

Proposition 1.5 For $p, q \in \mathcal{M}[D]$ and $g = gcrd(p, q)$, $\ell = \ell c\ell m(p, q)$ we have

(i) $\ker_{\mathcal{F}} g = \ker_{\mathcal{F}} p \cap \ker_{\mathcal{F}} q$

(ii) $\ker_{\mathcal{F}}\ \ell = \ker_{\mathcal{F}}\ p + \ker_{\mathcal{F}}\ q$ if p and q are full wrt $\mathcal{F}$

(iii) If p is full wrt $\mathcal{F}$ and $\ker_{\mathcal{F}} p(D) = < f_1, \ldots, f_n >_{\mathbb{R}}$ then

$$p(D) = u \cdot \ell c\ell m\{(f_i D - \dot{f}_i),\ i \in \underline{n}\} \ \text{ for some } \ u \in \mathcal{M}^*.$$

(iv) If p, q are full $wrt\mathcal{F}$ then g and ℓ are full $wrt\mathcal{F}$ as well.

Proof:
(i) is obvious from Lemma 1.4 and the definition of g.

(ii) Since the inclusion " $\supset$ " is immediate it suffices to prove that dim $\ker_{\mathcal{F}}\ \ell \leq$ dim $(\ker_{\mathcal{F}}\ p + \ker_{\mathcal{F}}\ q)$. Now by (i), (1.2) and (1.1) one obtains

$$\begin{aligned}
\dim\,(\ker_{\mathcal{F}}\ p + \ker_{\mathcal{F}}\ q) &= \dim\ \ker_{\mathcal{F}}\ p + \dim\ \ker_{\mathcal{F}}\ q - \dim\,(\ker_{\mathcal{F}}\ p \cap \ker_{\mathcal{F}}\ q) \\
&\geq \deg\ p + \deg\ q - \deg\ g \\
&= \deg\ \ell \\
&\geq \dim\ \ker_{\mathcal{F}}\ \ell
\end{aligned}$$

(iii) Suppose $f_1, \ldots, f_n$ are linearly independent. Then by Lemma 1.4 $\ell := \ell c\ell m\{(f_i D - \dot{f}_i)\ , i \in \underline{n}\}$ is a right factor of $p(D)$ with $n \geq$ deg ℓ. By (1.1) deg $\ell \geq$ dim $\ker_{\mathcal{F}} \ell \geq n$. Thus deg $p =$ deg ℓ and the proof is complete.

(iv) By Proposition 1.3 g is full $wrt\mathcal{F}$. Using (ii), (i) and (1.2) yields that

$$\begin{aligned}
\dim\ \ker_{\mathcal{F}}\ \ell &= \dim\,(\ker_{\mathcal{F}}\ p + \ker_{\mathcal{F}}\ q) \\
&= \dim\ \ker_{\mathcal{F}}\ p + \dim\ \ker_{\mathcal{F}}\ q - \dim\,(\ker_{\mathcal{F}}\ p \cap \ker_{\mathcal{F}}\ q) \\
&= \deg\ p + \deg\ q - \deg\ g \\
&= \deg\ \ell.
\end{aligned}$$

Thus by Definition 1.1 ℓ is full wrt $\mathcal{F}$. □

Proposition 1.6 Let $p, q \in \mathcal{M}[D]$ and suppose p is full *wrt F*. Then

$$\ker_{\mathcal{F}}\ p(D) \subset \ker_{\mathcal{F}}\ q(D) \ \text{ iff } \ q(D) = r(D) \cdot p(D) \ \text{ for some } \ r(D) \in \mathcal{M}[D].$$

Proof: Sufficiency is obvious. If the inclusion of the kernels is valid Proposition 1.5 (iii) gives a representation for $p(D)$. Thus by Lemma 1.4 $p(D)$ must right divide $q(D)$. □

We are now in a position to show an important result: The lattice (*wrt gcrd* and *ℓcℓm*) of left ideals $\mathcal{M}[D]p(D)$ generated by full (*wrt* $\mathcal{F}$) polynomials $p(D) \in \mathcal{M}[D]$ is antiisomorphic to the lattice of finite dimensional $\mathbb{R}$–subspaces of $\mathcal{F}$. Since this result is included in the matrix case, see Proposition 1.2.1, it is not proved here.

In the following a canonical form for matrices over $\mathcal{M}[D]$ with respect to multiplication by matrices of $GL_n(\mathcal{M}[D])$ from the left and the right is presented. For this a definition is needed.

Definition 1.7 $p, q \in \mathcal{M}[D]$ are called *similar* if they can be put in a *coprime relation*, i.e. if $pa = bq$ for some $a, b \in \mathcal{M}[D]$ and the only common left (right) divisors of $p, b(a, q)$ are units. p, q are called *associated* if $pu = vq$ for some units $u, v \in \mathcal{M}$.

Clearly, associated elements are similar. The more general notion of similarity is needed for normal forms over non commutative principal ideal domains. In a moment it will be shown that if p, q belong to the commutative ring $\mathbb{R}[D]$ and are similar *wrt* $\mathbb{R}[D]$ then p, q are necessarily associated.

First, let us note that similarity is an *equivalence* relation. This is due to the fact (see *Cohn* (1971) Section 3.2) that $p, q \in \mathcal{M}[D]$ are similar if and only if

$\mathcal{M}[D]/_{p\mathcal{M}[D]}$, $\mathcal{M}[D]/_{q\mathcal{M}[D]}$ are isomorphic as $\mathcal{M}[D]$ right modules

and this holds true if and only if

$\mathcal{M}[D]/_{\mathcal{M}[D]p}$, $\mathcal{M}[D]/_{\mathcal{M}[D]q}$ are isomorphic as $\mathcal{M}[D]$ left modules.

Thus similar polynomials have necessarily the same degree.

Assume $p, q \in \mathbb{R}[D]$ are similar. Then

$$p\mathbb{R}[D] = Ann(\mathbb{R}[D]/_{p\mathbb{R}[D]}) = Ann(\mathbb{R}[D]/_{q\mathbb{R}[D]}) = q\mathbb{R}[D],$$

where $Ann(\mathbb{R}[D]/_{p\mathbb{R}[D]}) := \{x \in \mathbb{R}[D] \mid \bar{r}x = 0 \; \forall\, \bar{r} \in \mathbb{R}[D]/_{p\mathbb{R}[D]}\}$, and thus p and q are associated.

Proposition 1.8 Suppose $p, q \in \mathcal{M}[D]$ are similar. Then they can be put in a coprime relation $pa' = b'q$ with $\deg a' = \deg b' < \deg p = \deg q$ for some $a', b' \in \mathcal{M}[D]$.

Proof: If $pa = bq$ is a coprime relation for some $a, b \in \mathcal{M}[D]$, then by the right Euclidean algorithm there exist $r, a' \in \mathcal{M}[D]$ such that $a = rq + a'$, $\deg a' < \deg q$. The coprime relation $pa = bq$ is equivalent to $pa' = b'q$ where $a' = a - rq$, $b' = b - pr$. It is easily seen that $pa' = b'q$ is coprime as well. Since $\deg b' = \deg a' < \deg q = \deg p$ the proof is complete. □

As a consequence of Proposition 1.8 similar polynomials of degree 1 are associated. This holds in general not true for polynomials of degree greater than 1. Consider for instance $p(D) = D^2$ and $q(D) = D^2 + 1$. It is easily seen that they are not associated. However $pa = bq$ with

$$a(D) := (t \sin t + 2 \cos t)D + 2 \sin t - t \cos t \text{ and } b(D) := (t \sin t + 2 \cos t)D + t \cos t$$

is a coprime relation. To see this assume that a and q have a common right divisor. By Lemma 1.4 this divisor can be assumed to be of the form $fD - \dot{f}$. Since $f \in \ker_{\mathcal{F}} q$ there exist $x, y \in \mathbb{R}$ such that $f = x \sin t + y \cos t$. From $a(D)(f) = 0$ it follows that $x = y = 0$. By using the same arguments the left coprimeness of p and b is shown.

If two full polynomials $p, q \in \mathcal{M}[D]$ are in a coprime relation $pa = bq$ then a, b are not necessarily full. To see this let $q(D) = D^2$ and $a(D) = t^2 D + 1$. a is not full *wrt* $\mathcal{M}$. However $pa = bq = \ell c\ell m(a, q)$ is a coprime relation and p is full *wrt* $\mathcal{M}$.

In order to characterize the equivalence classes of full polynomials a lemma is needed.

Lemma 1.9 Let $f_1, \ldots, f_n \in \mathcal{M}$ be linearly independent over $\mathbb{R}$ and $h_1, \ldots, h_n \in \mathcal{M}$. Then there exists $a \in \mathcal{M}[D]$ with $\deg a \leq n - 1$ so that $af_i = h_i$ for $i \in \underline{n}$.

Proof: Put $\ell := \ell c\ell m_{i \in \underline{n}}\{f_i D - \dot{f}_i)\}$. Then the Wronskian of $\ell(D)(f) = 0$ is given by

$$W = \begin{bmatrix} f_1 & \cdots & f_n \\ \vdots & & \vdots \\ f_1^{(n-1)} & \cdots & f_n^{(n-1)} \end{bmatrix}$$

and W is invertible over $\mathcal{M}$ since $t_1, \ldots, t_n$ are linearly independent, see *Coddington and Levinson* (1955) p. 83. Thus we can determine the coefficients of $a(D) := a_{n-1}D^{n-1} + \ldots + a_0$ as a solution of

$$\begin{bmatrix} a(D)(t_1) \\ \\ a(D)(t_n) \end{bmatrix} = W^T \begin{bmatrix} a_0 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix}$$

and the proof is complete. □

Theorem 1.10 Suppose $p \in \mathcal{M}[D]$ is full *wrt* $\mathcal{F}$ and deg $p = n$. Then the similarity class of p consists of all full polynomials *wrt* $\mathcal{M}$ of degree n .

Proof: Assume $pa = bq$ is a coprime relation and q is full . Since a and q are right coprime a acts as a monomorphism on $\ker_{\mathcal{M}} q$, use Lemma 1.4. Therefore $\dim \ker_{\mathcal{F}} p \geq \dim \ker_{\mathcal{M}} q$. Since p and q are full and of the same degree we obtain $\deg p = \dim \ker_{\mathcal{M}} q = n$.
So it remains to show that any full polynomial q of degree n can be put into a coprime relation with p. Let $f_1, \ldots, f_n$ and $h_1, \ldots, h_n$ be a basis of $\ker_{\mathcal{M}} q$ and $\ker_{\mathcal{F}} p$, resp. Since by Lemma 1.9 there exists $a(D) \in \mathcal{M}[D]$ with deg $a \neq n-1$ so that $a(D)(f_i) = h_i$ one obtains $\ker_{\mathcal{F}} q(D) \subset \ker_{\mathcal{M}} p(D)a(D)$. By Proposition 1.6 there exists $b \in \mathcal{M}[D]$ such that $pa = bq$. By construction a and q are right coprime. Suppose $p = up'$, $b = ub'$ for some $u, p', b' \in \mathcal{M}[D]$ such that p', b' are left coprime. Since a acts as a monomorphism on $\ker_{\mathcal{M}} q$, $p'a = b'q$ yields as in the first part of the proof $\deg p' = n$. Thus $u \in \mathcal{M}$ and the proof is complete. □

Now we are in a position to generalize the concept of full polynomials to the matrix case. Firstly a normal form for matrices over $\mathcal{M}[D]$ is given. $P, Q \in \mathcal{M}[D]^{m\times n}$ are called *equivalent* if $P = UQV$ for some $U \in \mathcal{M}[D]^{m\times m}, V \in \mathcal{M}[D]^{n\times n}$ invertible over $\mathcal{M}[D]$.

Proposition 1.11 Suppose $P(D) \in \mathcal{M}[D]^{m\times n}$. Then $P(D)$ is equivalent to some

$$P_c(D) = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & p(D) & & \\ & & & & 0 & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} \in \mathcal{M}[D]^{m\times n}$$

where $p(D)$ is uniquely determined up to similarity.
Let ℓ denote the number of non-zero entries in $P_c(D)$. If $\ell > 1$ then $p(D)$ can be chosen arbitrarily within its similarity class.

Proof: *Cohn* (1971) p. 288 proves the normal form for a more general ring. Since the ring $\mathcal{M}[D]$ is *simple*, i.e. the only two-sided ideals of $\mathcal{M}[D]$ are the trivial ones $\{0\}$ and $\mathcal{M}[D]$, the result simplifies considerably. To prove the last statement of the proposition it is sufficient to consider the case $m = n = 2$. Suppose p, q are full and are in a coprime relation $pa = bq$. Then by *Cohn* (1971) p.89 there exist $r, s, v, w \in \mathcal{M}[D]$ such that the inverse over $\mathcal{M}[D]$ of

$$U = \begin{bmatrix} p & b \\ r & s \end{bmatrix}$$

is given by

$$U^{-1} = \begin{bmatrix} v & -a \\ w & q \end{bmatrix}.$$

Now $-ra + sq = 1$ yields

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p & b \\ r & s \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & q \end{bmatrix} \begin{bmatrix} 1 & -a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -r & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & p \end{bmatrix}$$

and the proof is complete . □

Definition 1.12 Let $P(D) \in \mathcal{M}[D]^{r \times r}$ be equivalent to some $P_c(D)$ as in Proposition 1.11. Then $P(D)$ is called *non-singular* if no zeros occur in the diagonal of $P_c(D)$. The degree of $p(D)$ is called the *order* of $P(D)$, for short *ord* P.

A non-singular $P(D) \in \mathcal{M}[D]^{r \times r}$ is called *full wrt* $\mathcal{F}$ if the map

$$\begin{array}{rcl} \pi_I : \ker_{\mathcal{F}} P(D) & \to & \ker_{\mathcal{F}_I} P(D) \\ f & \mapsto & f\,|_I \end{array}$$

is an isomorphism for every open interval $I \subset \mathbb{R}$.

Lemma 1.13 If $P \in \mathcal{M}[D]^{r \times r}$ is non-singular then

$$Px = 0 \text{ for every } x \in \mathcal{M}[D]^r \quad \Rightarrow \quad x = 0$$

Proof: Use the normal form and the fact that $\mathcal{M}[D]$ does not contain zero divisors. □

Proposition 1.14 Suppose $P(D) \in \mathcal{M}[D]^{r \times r}$ is non-singular and equivalent to $P_c(D) = diag(1, \ldots, 1, p(D))$. Then

$$P(D) \text{ is full wrt } \mathcal{F} \text{ iff } \dim \ker_{\mathcal{F}} P(D) = \deg p(D) = ord\ P.$$

Proof: Clearly for every open interval $I \subset \mathbb{R}$

$$\deg p \geq \dim \ker_{\mathcal{M}_I} p = \dim \ker_{\mathcal{M}_I} P_c = \dim \ker_{\mathcal{M}_I} P \geq \dim \ker_{\mathcal{F}_I} P.$$

Since for sufficiently small I equality holds in the above inequality the proposition follows. □

Theorem 1.9 and Proposition 1.14 immediately give

Corollary 1.15 Every full $P \in \mathcal{M}[D]^{r \times r}$ is equivalent to

$$P_c(D) = \begin{bmatrix} 1 & & & \\ & \ddots & & 0 \\ & & 1 & \\ 0 & & & D^n \end{bmatrix} \in \mathbb{R}[D]^{r \times r}, \text{ where } n = ord\ P.$$

Example 1.16
(i) In the time-invariant polynomial framework polynomial matrices over $\mathbb{R}[D]$ are extensively studied, cf. *Rosenbrock* (1970), *Wolovich* (1974). If $P \in \mathbb{R}[D]^{r\times r}$ is transformed into a Smith form $P_s = diag(p_1, \ldots, p_r)$ by unimodular (over $\mathbb{R}[D]$) matrices then

$$\dim \ker_{\mathcal{A}} P = \dim \ker_{\mathcal{A}} P_s = \sum_{i=1}^{r} \deg p_i.$$

Thus every non-singular $P \in \mathbb{R}[D]^{r\times r}$ is full wrt $\mathcal{A}$ and $ord\, P = \deg\, det\, P$.

(ii) Every monic $P \in (\mathcal{A}^{r\times r})[D]$ is full wrt $\mathcal{A}$. To see this reduce $P(D)(f) = 0$ to $(DI_{rn} - B)(g) = 0$ where $B \in \mathcal{A}^{rn\times rn}$ and $P = \sum_{i=0}^{n} P_i D^i, P_n \neq 0$. (This is done in the same way as a nth-order differential equation is reduced to a first order matrix equation, see e.g. *Coddington and Levinson* (1955) p. 21.) Since the solution spaces of $P(D)(f) = 0$ and $(DI_{rn} - B)(g) = 0$ are isomorphic the claim follows.

(iii) Every non-singular $P \in \mathcal{A}[D]^{r\times r}$ which is in normed upper triangular form (as considered in *Ylinen* (1980)) is full wrt $\mathcal{A}$. P is called in *normed upper triangular form* if (1) it is an upper triangular matrix and (2) if $(0, \ldots, 0, p_{i,j_0}, *, \ldots, *)$ denotes the i-th row of P so that $p_{i,j_0} \neq 0$ is monic and $p_{i+1,j_0} = \ldots = p_{r,j_0} = 0$, then $\deg\, p_{\lambda,j_0} < \deg\, p_{i,j_0}$ for all $\lambda \in \underline{i-1}$.
Let $U \in GL_r(\mathcal{A}[D])$ so that the entries of $P' = PR$ satisfy $p'_{ij} = 0$ if $i > j, p'_{ii} = p_{ii}, \deg\, p'_{ij} < min\,(\deg\, p'_{ij}, \deg\, p'_{jj})$ if $i \neq j$. Put

$$Q = diag(D^{s_0 - s_1}, \ldots, D^{s_0 - s_r}) \text{ where } s_0 := max_{i\in \underline{r}} \deg\, p'_{ii},\ s_i = \deg\, p'_{ii} \text{ for } i \in \underline{r}.$$

Then $Q\,P\,U$ is a monic element of $\mathcal{A}^{r\times r}[D]$ which is full by (ii). This implies fullness of P .

As a generalization of the scalar case one obtains

Proposition 1.17 Suppose $P, Q, G \in \mathcal{M}[D]^{r\times r}$ and $P = Q\,G$. Then
(i) G is full *wrt* $\mathcal{F}$ and Q is full *wrt*$\mathcal{M}$ if P is full *wrt* $\mathcal{F}$

(ii) $ord\, P = ord\, Q + ord\, G$.

Proof: (i) is a straigtforward generalization of the proof of Proposition 1.3. To prove the order formula note that for every interval $I \subset \mathbb{R}$

$$\dim \ker_{\mathcal{F}_I} P \geq \dim \ker_{\mathcal{M}_I} Q + \dim \ker_{\mathcal{F}_I} G \tag{1.3}$$

Now for I sufficiently small $\ker_{\mathcal{M}_I} Q \subset im(G\,|_I)$. Thus equality holds in (1.3). Choosing I eventually smaller one can achieve that $\dim \ker_{\mathcal{A}_I} P = ord\, P$ and the analogous statement for Q and G . This proves the order formula. □

The following proposition extends Proposition 1.6 to the matrix case.

Proposition 1.18 If $P \in \mathcal{M}[D]^{r\times r}$ is full wrt $\mathcal{F}$ and $Q \in \mathcal{M}[D]^{r_1\times r}$ then

$$\ker_{\mathcal{F}} P(D) \subset \ker_{\mathcal{F}} Q(D) \text{ iff } Q = RP \text{ for some } R \in \mathcal{M}[D]^{r_1\times r}.$$

Proof: Only necessity has to be shown. By Corollary 1.15 there exist $U, V \in GL_n(\mathcal{M}[D])$ so that $P = UP_cV$ where $P_c = diag(1, \ldots, 1, D^n)$. Since P is full, $V^{-1}(f) \in \ker_{\mathcal{F}} P$ for every $f \in \ker_{\mathcal{F}} P_c$. By assumption $\ker_{\mathcal{F}} P_c \subset \ker_{\mathcal{F}} QV^{-1}$. Since

$$\ker_{\mathcal{F}} D^n \subset \ker_{\mathcal{F}} (QV^{-1})_{ir} \text{ for } i \in \underline{r_1},$$

by Proposition 1.6 there exist $t_1, \dots, t_{r_1} \in \mathcal{M}[D]$ such that $(QV^{-1})_{ir} = t_i D^n$. Let $(QV^{-1})_j, j \in \underline{r}$, denote the colums of QV^{-1}. Then

$$\underbrace{[(QV^{-1})_1, \dots, (QV^{-1})_{r-1}, (t_1, \dots, t_{r_1})^T]}_{=:\tilde{T}} diag(1, \dots, 1, D^n) = QV^{-1}$$

and thus $(\tilde{T}U^{-1})UP_cV = Q$. This completes the proof. □

In the time-varying setup we have the nice result that for every finite dimensional $\mathbb{R}$-linear subspace $\mathcal{V}$ of $\mathcal{F}^r$ one can find a full polynomial matrix $P \in \mathcal{M}[D]^{r \times r}$ which annules exactly this subspace $\mathcal{V}$. This is proved in the following proposition and extends the scalar case considered in Proposition 1.5 (iii).

Proposition 1.19 Suppose $\mathcal{V} = < f_1, \dots, f_n >_{\mathbb{R}}$ is an n-dimensional subspace of $\mathcal{F}^r$. Then there exists a $P \in \mathcal{M}[D]^{r \times r}$ full *wrt* $\mathcal{F}$ such that $\ker_{\mathcal{F}} P(D) = \mathcal{V}$.

Proof: Denote $f_i = (f_{i1}, \dots, f_{ir})^T$ for $i \in \underline{n}$ and

$$A := \begin{bmatrix} f_{11} & \cdots & f_{n1} \\ \vdots & & \vdots \\ f_{1r} & \cdots & f_{nr} \end{bmatrix}$$

Without restriction assume that the first row of A is non zero otherwise multiply A from the left by an invertible matrix. Choose a $\mathbb{R}$-basis of the first row entries and multiplication from the right by some $U_1 \in GL_n(\mathbb{R})$ yields

$$AU_1 = \begin{bmatrix} g_{11} & \cdots & g_{i_1 1} & 0 & \cdots & 0 \\ g_{12} & & \cdots & & & g_{n2} \\ \vdots & & & & & \vdots \\ g_{1r} & & \cdots & & & g_{nr} \end{bmatrix}$$

with $g_{11}, \dots, g_{i_1,1}$ linearly independent. The columns of AU_1 are still a basis of $\mathcal{V}$. By Lemma 1.9 there exists $p_2 \in \mathcal{M}[D]$ such that $p_2 g_{k1} = g_{k2}$ for $k \in \underline{i_1}$.
Therefore with

$$P_2 = \begin{bmatrix} 1 & & & & \\ -p_2 & \ddots & & & 0 \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & 1 \end{bmatrix} \in \mathcal{M}[D]^{r \times r}$$

one obtains

$$P_2AU_1 = \begin{bmatrix} g_{11} & \cdots & g_{i1} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & g_{i_1+1,2} & \cdots & g_{n2} \\ & & * & & & \end{bmatrix}$$

Defining $p_3, \dots, p_r \in \mathcal{M}[D]$ and $P_3, \dots, P_r \in \mathcal{M}[D]^{r \times r}$ in a similar way gives

$$P_r \dots P_2AU_1 = \begin{bmatrix} g_{11} & \cdots & g_{i_1 1} & 0 & \cdots & 0 \\ & & & g_{i_1+1,2} & \cdots & g_{n2} \\ & 0 & & \vdots & & \vdots \\ & & & g_{i_1+1,r} & \cdots & g_{nr} \end{bmatrix}$$

Applying this procedure successively on the remaining submatrices we finally obtain $\tilde{P} \in GL_r(\mathcal{M}[D])$ and $\tilde{U} \in GL_n(\mathbb{R})$ such that

$$\tilde{P}A\tilde{U} = \begin{bmatrix} g_{11} & \cdots & g_{i_1 1} & 0 & \cdots & & & 0 \\ & & & g_{12}^{(2)} & \cdots & g_{i_1 2}^{(2)} & 0 & \\ & & 0 & & \ddots & & \ddots & \\ & & & & & g_{1k}^{(k)} & \cdots & g_{i_k k}^{(k)} \\ 0 & & & \cdots & & & & 0 \\ \vdots & & & & & & & \vdots \\ 0 & & & \cdots & & & & 0 \end{bmatrix}$$

for $k \in \underline{r}$ and the elements $g_{c,j}^{(j)}$ in every row are linearly independent. Now define

$$q_j = \ell c\ell m_{1 \geq c \geq i_j}(g_{cj}^{(j)} D - \dot{g}_{cj}^{(j)}) \qquad \text{for } j \in \underline{k}$$

For $Q := diag(q_1, \ldots, q_k, 1, \ldots, 1) \in \mathcal{M}[D]^{r \times r}$ we have $Q\tilde{P}A\tilde{U} = 0$. Since Q is full *wrt* $\mathcal{F}$ the product $P = Q\tilde{P}$ is also full *wrt* $\mathcal{F}$ and $\ker_{\mathcal{F}} P = V$. $\square$

Algebraic properties of common divisors and multiple hold true analogously to the commutative case (cf. for instance *Mac Duffee* (1956)). For $P \in \mathcal{M}[D]^{r \times r}, Q \in \mathcal{M}[D]^{r \times m}$ a *greatest common right divisor* $G = gcrd(P,Q)$ and a *least common left multiple* $L = \ell c\ell m(P,Q)$ are defined analogously as for analytic matrices in Section 1.5.

Lemma 1.20 For $P \in \mathcal{M}[D]^{r \times r}, Q \in \mathcal{M}[D]^{r_1 \times r}$ we have

(i) There exists $G = gcrd(P,Q)$ and $A \in \mathcal{A}[D]^{r \times r}, B \in \mathcal{A}[D]^{r \times r_1}$ such that

$$G = AP + BQ.$$

If P is non-singular then G is unique up to left multiplication by an invertible matrix. If P is full wrt $\mathcal{F}$, G is full *wrt* $\mathcal{F}$ as well.

(ii) If $r_1 = r$ and both P and Q are non-singular, then there exists $L = \ell c\ell m(P,Q)$ which is unique up to multiplication from the left by an invertible matrix. If P and Q are full wrt $\mathcal{F}$, L is full wrt $\mathcal{F}$ as well.

(iii) P and Q are called *right coprime* if every square common right divisor of P and Q is invertible over $\mathcal{M}[D]$. This is true iff there exists $S \in \mathcal{M}[D]^{r \times r}, T \in \mathcal{M}[D]^{r \times r_1}$ such that

$$I_r = SP + TQ.$$

Proof: The main idea of the proof is as follows: There exist matrices

$$U = \begin{bmatrix} U_1 & U_2 \\ U_3 & U_4 \end{bmatrix}, \; U^{-1} = \begin{bmatrix} V_1 & V_2 \\ V_3 & V_4 \end{bmatrix} \in \mathcal{M}[D]^{(r+r_1)\times(r+r_1)} \tag{1.4}$$

where U_1 is a $r \times r$ matrix and all other matrices have appropriate formats such that

$$U \begin{bmatrix} P \\ Q \end{bmatrix} = \begin{bmatrix} G \\ O \end{bmatrix}, G \in \mathcal{M}[D]^{r \times r}$$

This is proved analogously as in *Newman* (1972) p. 15, he considers matrices over commutative principle ideal domains.

That G is a *gcrd* of P and Q and satisfies the uniqueness statement is proved analogously as in *Mac Duffee* (1956) p. 35. If P is full it follows from Proposition 1.17 that G is full. Using the same computations as in *Mac Duffee* (1956) p.36 yields that

$$L := U_3\,P \;=\; -U_4\,Q \;=\; \ell c\ell m\,(P,Q).$$

To prove that L is non-singular assume $xL = 0$ for some $x \in \mathcal{M}[D]^{1\times r}$. Then $xU_3P = 0 = xU_4Q$, and hence $xU_3 = xU_4 = 0$. Since $U_3V_2 + U_4V_4 = I_r$ by (1.4) it follows that $x = 0$. To prove that L is full *wrt* $\mathcal{F}$ is more difficult. Put $\mathcal{V} = \ker_{\mathcal{F}}\, P + \ker_{\mathcal{F}}\, Q$, then by Proposition 1.16 there exists $L' \in \mathcal{M}[D]^{r\times r}$ full *wrt*$\mathcal{F}$ such that $\ker_{\mathcal{F}}\, L' = \mathcal{V}$. Thus by Proposition 1.16 L' is a common left multiple of P and Q and there exists $E \in \mathcal{M}[D]^{r\times r}$ such that $E\,L = L'$. Since L' is full (ii) follows from Proposition 1.17.
(iii) is a consequence of the presentation of the *gcrd* given in (i). □

Now we are in a in a position to state the main result of this section. It shows in particular that full matrices have a one-to-one correspondence to finite dimensional linear subspaces of $\mathcal{F}^r$.

Theorem 1.21 Let $\mathcal{F} = \mathcal{A}$ or $\mathcal{M}$. The set

$$\mathcal{L}_f := \{\mathcal{M}[D]^{r\times r}\cdot P \mid P \in \mathcal{M}[D]^{r\times r} \text{ full } wrt\,\mathcal{F}\}$$

of left $\mathcal{M}[D]$ modules generated by full matrices is a lattice with respect to the operations

$$\mathcal{M}[D]^{r\times r}\cdot P \;\vee\; \mathcal{M}[D]^{r\times r}\cdot Q \;=\; \mathcal{M}[D]^{r\times r}\cdot gcrd(P,Q)$$

$$\mathcal{M}[D]^{r\times r}\cdot P \;\wedge\; \mathcal{M}[D]^{r\times r}\cdot Q \;=\; \mathcal{M}[D]^{r\times r}\cdot \ell c\ell m(P,Q)$$

The set

$$L_f := \{\mathcal{V} \subset \mathcal{F}^r \mid \mathcal{V} \text{ is a finite dimensional linear subspace of } \mathcal{F}^r\}$$

is a lattice with respect to intersection and sum.
The map

$$\begin{array}{rcll} h & : & \mathcal{L}_f & \to\; L_f \\ & & \mathcal{M}[D]^{r\times r}\cdot P & \mapsto\; \ker_{\mathcal{F}}\, P \end{array}$$

is an anti-isomorphism, where 'anti' means

$$h(\mathcal{M}[D]^{r\times r}\cdot P \,\vee\, \mathcal{M}[D]^{r\times r}\cdot Q) \;=\; h(\mathcal{M}[D]^{r\times r}\cdot P) + h(\mathcal{M}[D]^{r\times r}\cdot Q) \tag{1.5}$$

$$h(\mathcal{M}[D]^{r\times r}\cdot P \,\wedge\, \mathcal{M}[D]^{r\times r}\cdot Q) \;=\; h(\mathcal{M}[D]^{r\times r}\cdot P) \cap h(\mathcal{M}[D]^{r\times r}\cdot Q) \tag{1.6}$$

Proof: By Lemma 1.20 (i) and (ii) the *gcrd* and the *$\ell c\ell m$* of full matrices are full as well, whence $\mathcal{L}_f$ is a lattice. The map h is well-defined by Proposition 1.18. Injectivity and surjectivity of h follows from Proposition 1.18 and 1.19. So it remains to prove (1.5) and (1.6) which are equivalent to

$$\ker_{\mathcal{F}}\, L = \ker_{\mathcal{F}}\, P + \ker_{\mathcal{F}}\, Q \quad ,\text{for } L = \ell c\ell m(P,Q) \tag{1.7}$$

$$\ker_{\mathcal{F}}\, G = \ker_{\mathcal{F}}\, P \cap \ker_{\mathcal{F}}\, Q \quad ,\text{for } G = gcrd(P,Q). \tag{1.8}$$

" $\supset$ " in (1.7) is evident. To prove the converse inclusion note that by Proposition 1.19 there exists $L' \in \mathcal{M}[D]^{r\times r}$ such that

$$\ker_{\mathcal{F}}\, P + \ker_{\mathcal{F}}\, Q = \ker_{\mathcal{F}}\, L'.$$

Now Proposition 1.18 yields $L' = EL$ for some $E \in \mathcal{M}[D]^{r\times r}$ and thus $\ker_{\mathcal{F}} L \subset \ker_{\mathcal{F}} L'$ which proves (1.7). (1.8) is easily proved by using Proposition 1.18, here the assumption that P and Q are full is not necessary. $\square$

For the sake of completeness it is shown that the degree formula (1.2) carries over for matrices as follows.

Remark 1.22 Suppose $P, Q \in \mathcal{M}[D]^{r\times r}$ and $G = gcrd(P,Q)$, $L = \ell c\ell m(P,Q)$. Then by using (1.7) and (1.8) and choosing $I \subset \mathbb{R}$ sufficiently small one obtains

$$\begin{aligned} \dim \ker_{\mathcal{M}_I} P + \ker_{\mathcal{M}_I} Q &= \dim (\ker_{\mathcal{M}_I} P + \ker_{\mathcal{M}_I} Q) + \dim (\ker_{\mathcal{M}_I} P \cap \ker_{\mathcal{M}_I} Q) \\ &= \dim \ker_{\mathcal{M}_I} L + \dim \ker_{\mathcal{M}_I} G. \end{aligned}$$

For I eventually smaller this gives

$$ord\, P + ord\, Q = ord\, L + ord\, G \tag{1.9}$$

2.2 Polynomial matrix systems, solution vector spaces and system equivalence

In this section we will analyse time-varying finite dimensional linear systems in differential operator representation.

$$\begin{aligned} P(D)(z) &= Q(D)(u) \\ y &= V(D)(z) + W(D)(u) \end{aligned} \tag{2.1}$$

where P, Q, V, W are $r\times r, r\times m, p\times r, p\times m$ matrices, resp. defined over $\mathcal{M}[D]$.

$$u \in \mathcal{U}^m := \{u \in (C^\infty)^m \mid \text{supp } u \text{ bounded to the left }\}$$

Additionally it is assumed that

(A1) $P(D)$ is full wrt $\mathcal{A}$.

(A2) $im\ Q \subset im\ P$, i.e. for every $u \in \mathcal{U}^m$ there exists $z \in (C^\infty)^r$ such that the first equation in (2.1) is satisfied

(A3) If $u \in \mathcal{U}^m$ then it follows that $y \in (C^\infty)^p$.

The first assumption yields that every free motion of the first equation in (2.1) is defined on the whole time axis $\mathbb{R}$ and does not have poles. Furthermore the requirement "full" allows, as we will see, an algebraic treatment of systems of the form (2.1). (A 2) is natural from a system theoretic point of view, for every input $u(\cdot)$ the existence of an "internal state" and an output is expected.However, if only (A 2) holds it may happen that for some $u \in \mathcal{U}^m$ the output is of the form

$$\left(\frac{y_1}{\tilde{y}_1}, \ldots, \frac{y_p}{\tilde{y}_p}\right) \quad \text{where } y_i \in C^\infty, \tilde{y}_i \in \mathcal{A}^*,\ i \in \underline{p}.$$

The poles occur from the poles in the coefficients of V and W. Assumption (A 3) ensures that the output does not have poles. If V and W are defined over $\mathcal{A}[D]$ this is clearly satisfied.

In the following the matrix

$$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(r+p)\times(r+m)}$$

is called a *system matrix* if it corresponds to (2.1) and (A1) - (A3) are satisfied.

This class of systems covers in particular

- time-invariant systems in differential operator representation as introduced by *Rosenbrock* (1970).
- systems where $P \in \mathcal{A}[D]^{r\times r}$ is non-singular and in normed upper triangular form as dealt with in *Ylinen* (1980), cf. Example 1.16 (iii).
- state space system of the form

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) + E(D)(u(t)) \end{aligned} \tag{2.2}$$

- where A, B, C are analytic $n \times n, n \times m, p \times n$ matrices, resp. and $E(D) \in \mathcal{A}[D]^{p\times m}$.

The following proposition shows that the *solution* (vector) *space*

$$\mathcal{M}(P,Q) := \{(z,u) \in (C^{\infty})^r \times \mathcal{U}^m \mid P(D)(z) = Q(D)(u)\}$$

can be decomposed into the direct sum of the $\mathbb{R}$-linear subspace of *forced motions starting from zero*

$$\mathcal{M}_{+}(P,Q) := \{(z_u, u) \in \mathcal{M}(P,Q) \cap (\mathcal{U}^r \times \mathcal{U}^m)\},$$

(where throughout this chapter z_u denotes the uniquely defined forced motion starting from zero) and into the $\mathbb{R}$-linear *subspace of free motions*

$$\ker_{\mathcal{A}} \; P \times \{0\} := \{(z',0) \in \mathcal{M}(P,Q)\}.$$

As opposed to time-invariant systems where $\mathcal{M}(P,Q)$ is an $\mathbb{R}[D]$ -module, for time-varying systems of the form (2.1) $\mathcal{M}(P,Q)$ is, in general, only an $\mathbb{R}$–vector space not an $\mathbb{R}[D]$ - or $\mathcal{M}[D]$-module.

Proposition 2.1 Suppose $P \in \mathcal{M}[D]^{r\times r}$ is full *wrt*$\mathcal{A}$ and $Q \in \mathcal{M}[D]^{r\times m}$, then

$$\mathcal{M}(P,Q) \;=\; \mathcal{M}_{+}(P,Q) \;\oplus\; (\ker_{\mathcal{A}} \; P \times \{0\})$$

Proof: That the sum is direct follows from the definition of the vector spaces and the fact that P is full. It remains to prove " $\subset$ ". Let $(z,u) \in \mathcal{M}(P,Q)$. Then $u\,|_I \equiv 0$ for some $I = (-\infty, t_0)$ and $(z,u)\,|_I = (z\,|_{I'}, 0)$. Since P is full there exist $z' \in \ker_{\mathcal{A}} \; P$ such that $z'\,|_I = z\,|_I$. Thus $(z,u) = (z',0) + (z - z', u)$ where $(z - z', u) \in \mathcal{M}_{+}(P,Q)$. □

The next lemma is frequently used in the following.

Lemma 2.2 Suppose $A(D) \in \mathcal{M}[D]^{r\times m}$. Then

$$A(D)(u) = 0 \;\text{ for all } u \in \mathcal{U}^m \;\;\Rightarrow\;\; A(D) = 0$$

Proof: Suppose $A(D) = \sum_{i=0}^{n} A_i D^i$. For $t_0 < t_1$ choose $u \in C^\infty$ such that

$$u(t) = \begin{cases} 0 & t \le t_0 \\ 1 & t_1 \le t \end{cases}$$

If e_j denotes the j-th canonical basis vector in $\mathbb{R}^m$ one obtains

$$A(D)(e_j u(t)) = A_0 e_j = 0 \qquad \text{for } t \ge t_1 \text{ and } j \in \underline{m}.$$

Since A_0 is analytic $A_0 = 0$. Inserting successively $te_j u(\cdot), \ldots, t^n e_j u(\cdot)$ yields

$$A(D)(t^k e_j u(t)) = A_k e_j = 0 \qquad \text{for } t \ge t_1, j \in \underline{m}, k \in \underline{n}$$

Therefore $A_k = 0$ for $k = 0, \ldots, n$. □

Proposition 2.3 Suppose $P_i \in \mathcal{M}[D]^{r_i \times r_i}$, $Q_i \in \mathcal{M}[D]^{r_i \times m}$ for $i = 1, 2$.

(i) If P_1 is full *wrt*$\mathcal{M}$ and $im\ Q_1 \subset im\ P_1$ then

$$\mathcal{M}(P_1, Q_1) \subset \mathcal{M}(P_2, Q_2) \text{ iff } T[P_1, Q_1] = [P_2, Q_2] \text{ for some } T \in \mathcal{M}[D]^{r_2 \times r_1}.$$

(ii) Assume $r_1 = r_2$, P_1 and P_2 are full wrt $\mathcal{M}$ and $im\ Q_1 \subset im\ P_1$. Then

$$\mathcal{M}(P_1, Q_1) = \mathcal{M}(P_2, Q_2) \text{ iff } T[P_1, Q_1] = [P_2, Q_2] \text{ for some } T \in GL_{r_1}(\mathcal{M}[D]).$$

Proof: (i) Sufficiency is trivial. To prove necessity note that $\ker_{\mathcal{A}} P_1 \subset \ker_{\mathcal{A}} P_2$ yields the existence of some $T \in \mathcal{M}[D]^{r_2 \times r_1}$ such $P_2 = TP_1$, see Proposition 1.18. Now $im\ Q_1 \subset im\ P_1$ implies $Q_2(u) = P_2(z) = TP_1(z) = TQ_1(u)$ for all $u \in \mathcal{U}^m$ and thus by Lemma 2.2 (i) is proved.

(ii) Applying (i) twice yields the existence of some $T, T' \in \mathcal{M}[D]^{r_1 \times r_1}$ such that $T[P_1, Q_1] = [P_2, Q_2]$ and $P_1 = T'P_2$. Hence $P_1 = T'TP_1$ and by Lemma 1.13 $T' = T^{-1}$. □

For time-invariant systems in differential operator description *Hinrichsen and Prätzel-Wolters* (1980) have studied system equivalence via certain homomorphisms between the solution modules. The following definition extends this approach to the time-varying setting.

Definition 2.4 Suppose

$$\mathbb{P}_i = \begin{bmatrix} P_i & Q_i \\ V_i & W_i \end{bmatrix} \in \mathcal{M}[D]^{(r_i+p)\times(r_i+m)} \quad ,(i = 1, 2)$$

are system matrices.

(i) A $\mathbb{R}$−linear map $f : \mathcal{M}(P_1, Q_1) \to \mathcal{M}(P_2, Q_2)$ is called a *solution homomorphism* if

$$f(z, u) = \begin{bmatrix} T_1(D) & Y(D) \\ 0 & I_m \end{bmatrix} \begin{pmatrix} z \\ u \end{pmatrix} \tag{2.3}$$

for some $T_1 \in \mathcal{M}[D]^{r_2 \times r_1}$, $Y \in \mathcal{M}[D]^{r_2 \times m}$.

(ii) A solution homomorphism $f : \mathcal{M}(P_1, Q_1) \to \mathcal{M}(P_2, Q_2)$ is called a *system homomorphism* if

$$V_1(D)(z) + W_1(D)(u) = [V_2(D), W_2(D)](f(z, u)) \quad \text{for all } (z, u) \in \mathcal{M}(P_1, Q_1). \tag{2.4}$$

(iii) $\mathbb{P}_1, \mathbb{P}_2$ are called *system equivalent*, denoted by $\mathbb{P}_1 \overset{se}{\sim} \mathbb{P}_2$, if there exists a *system isomorphism* $f : \mathcal{M}(P_1, Q_1) \to \mathcal{M}(P_2, Q_2)$, i.e. f is a system homomorphism which is invertible as a system homomorphism.

Using the notation of the previous definition a system homomorphism f makes the following diagram commute

$$\begin{tikzcd}
 & (z,u) \arrow[ddl, mapsto] \arrow[ddr, mapsto] & \\
 & \mathcal{M}(P_1,Q_1) \arrow[ddl] \arrow[ddr] \arrow[dddd, "f"] & \\
u & & V_1(z)+W_1(u) \\
\mathcal{U}^m & & (\mathcal{C}^\infty)^p \\
u & & V_2(\tilde{z})+W_2(u) \\
 & \mathcal{M}(P_2,Q_2) \arrow[uul] \arrow[uur] & \\
 & (\tilde{z},u) \arrow[uul, mapsto] \arrow[uur, mapsto] &
\end{tikzcd}$$

In other words: f does not transform the controls u , the output of the first system corresponding to the solution pair (z,u) is the same as the output of the second system corresponding to the associated solution pair $(f(z,u),u)$, the internal variables are transformed.
Since a solution space $\mathcal{M}(P,Q)$ can directly decomposed as in Proposition 2.1 a solution homomorphism (2.3) induces the following $\mathbb{R}$ -homomorphisms

$$\begin{array}{rcl} f_0 : \ker_{\mathcal{A}} P_1 & \rightarrow & \ker_{\mathcal{A}} P_2 \\ z & \mapsto & T_1(D)(z) \end{array} \tag{2.6}$$

$$\begin{array}{rcl} f_1 : \mathcal{M}_+(P_1,Q_1) & \rightarrow & \mathcal{M}_+(P_2,Q_2) \\ (z_u,u) & \mapsto & (T_1(D)(z_u)+Y(D)(u),u) \end{array} \tag{2.7}$$

In particular, f preserves the direct decomposition, i.e.

$$\begin{array}{rcl} f(\mathcal{M}_+(P_1,Q_1)) & \subset & \mathcal{M}_+(P_2,Q_2) \\ f(\ker_{\mathcal{A}} P_1 \times \{0\}) & \subset & \ker_{\mathcal{A}} P_2 \times \{0\} \end{array}$$

Example 2.5 (i) Two state space systems of the form (2.2) associated with

$$\mathbb{P}^i_{st} = \begin{bmatrix} DI_n - A_i & -B_i \\ C_i & E_i(D) \end{bmatrix} \in \mathcal{A}[D]^{(n+p)\times(n+m)}, \quad ,i=1,2$$

are called *similar* via $T \in GL_n(\mathcal{A})$ if

$$\begin{bmatrix} T^{-1} & 0 \\ 0 & I_p \end{bmatrix} \mathbb{P}^1_{st} = \mathbb{P}^2_{st} \begin{bmatrix} T^{-1} & 0 \\ 0 & I_m \end{bmatrix}$$

This extends the concept of similarity introduced in Section 1.1, cf. also Lemma 1.1.2.
If $\mathbb{P}^1_{st}$ is similar to $\mathbb{P}^2_{st}$ it is easily verified that the map

$$f : \mathcal{M}(DI_n - A_1, B_1) \rightarrow \mathcal{M}(DI_n - A_2, B_2), \quad (z,u) \mapsto \begin{bmatrix} T^{-1} & 0 \\ 0 & I_m \end{bmatrix} \begin{pmatrix} z \\ u \end{pmatrix}$$

is a system isomorphism.

(ii) Consider the *n-th* order scalar differential equation

$$p_1(D)(z) = q_1 u$$

where $p_1(D) = a_0 + \ldots + a_{n-1}D^{n-1} + D^n \in \mathcal{A}[D]$, $q_1 \in \mathcal{A}$. It is well-known that this equation is equivalent to the first order matrix differential equation.

$$P_2(D)(z) = Q_2 u$$

where

$$P_2(D) = \left[DI_n - \begin{bmatrix} 0 & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & 0 \quad 1 & \\ -a_0 & & \cdots & & -a_{n-1} \end{bmatrix} \right], \quad Q_2 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ q_1 \end{bmatrix}$$

The solution isomorphism is given by

$$\begin{aligned} f : \mathcal{M}(p_1, q_1) &\rightarrow \mathcal{M}(P_2, Q_2) \\ (z, u) &\mapsto \begin{bmatrix} \begin{bmatrix} 1 \\ D \\ \vdots \\ D^{n-1} \end{bmatrix} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ 0 & 1 \end{bmatrix} \begin{pmatrix} z \\ u \end{pmatrix} \end{aligned}$$

Proposition 2.6 Suppose

$$\mathbb{P}^i_{st} = \begin{bmatrix} DI_{r_i} - A_i & -B_i \\ C_i & E_i(D) \end{bmatrix} \in \mathcal{A}[D]^{(r_i+p)\times(r_i+m)}$$

are associated with two state space systems of the form (2.2), $i = 1, 2$ resp. Then

$$\mathbb{P}^1_{st} \text{ is similar to } \mathbb{P}^2_{st} \quad \text{iff} \quad \mathbb{P}^1_{st} \overset{se}{\sim} \mathbb{P}^2_{st}.$$

Proof: Sufficiency is proved in Example 2.5(i). Assume $\mathbb{P}^1_{st} \overset{se}{\sim} \mathbb{P}^2_{st}$. Then $\ker_{\mathcal{A}} DI_{r_1} - A_1 \cong \ker_{\mathcal{A}} DI_{r_2} - A_2$ and thus $r_1 = r_2 =: r$. Since every state space system is similar to a state space system with constant free motion (see Remark 1.1.1) and similar systems are system equivalent we assume without restriction of generality that $A_1 = A_2 = 0$. Let

$$f : \mathcal{M}(DI_r, B_1) \rightarrow \mathcal{M}(DI_r, B_2), \quad (z, u) \mapsto \begin{bmatrix} T_1 & Y \\ 0 & I_n \end{bmatrix} \begin{pmatrix} z \\ u \end{pmatrix}$$

denote the system isomorphism, $T_1 \in \mathcal{M}[D]^{r\times r}, Y \in \mathcal{M}[D]^{r\times m}$. Choose $Q \in \mathcal{M}[D]^{r\times r}, H \in \mathcal{M}^{r\times r}$ such that $T_1 = Q\, DI_r + H$. Then

$$f(z, u) = (T_1(z) + Y(u), u) = ((Q\, B_1 + Y)(u) + H z, u).$$

Hence $f_0 : \ker_{\mathcal{A}} DI_r \rightarrow \ker_{\mathcal{A}} DI_r$ is an isomorphism described by $z \mapsto Hz$ and therefore $H \in GI_n(\mathbb{R})$. Furthermore

$$DI_r((Q\, B_1 + Y)(u) + H z) = B_2 u \qquad \text{for all } (z, u) \in \mathcal{M}(DI_r, B_1)$$

and defining $Y_1 := Q\, B_1 + Y$ yields

$$DI_r(z) = H^{-1}(B_2 - DI_r\, Y_1)(u) \quad \text{for all} \quad (z,u) \in \mathcal{M}(DI_r, B_1).$$

Thus by Lemma 2.2 $H^{-1}(B_2 - DI_r\, Y_1) = B_1$. Comparing the coefficients gives $H^{-1}\, Y_1 = 0$ and $H^{-1}\, B_2 = B_1$.
It remains to prove $C_2\, H = C_1$ and $E_1(D) = E_2(D)$. By (2.4)

$$C_1 z + E_1(D)(u) = (C_2, E_2(D))(T_1(z) + Y(u), u)^T = C_2 H z + E_2(D)(u) \text{ for all } (z,u) \in \mathcal{M}(DI_r, B_1).$$

Since $\ker_{\mathcal{A}}\, DI_r = \mathbb{R}^n$ one concludes for $u = 0$ that $C_1 = C_2 H$. Furthermore by Lemma 2.2, $E_1(D) = E_2(D)$. This proves the proposition. □

For the algebraic characterization of the injectivity and surjectivity of a solution homomorphism (2.3) a lemma is needed.

Lemma 2.7 If f is a solution homomorphism as in (2.3), then there exists a $T \in \mathcal{M}[D]^{r_2 \times r_1}$ such that

$$T[P_1, Q_1] = [P_2, Q_2] \begin{bmatrix} T_1 & Y \\ 0 & I_m \end{bmatrix}$$

Proof: Note that

$$f(\mathcal{M}(P_1, Q_1)) \subset \mathcal{M}(P_2, Q_2) \quad \Leftrightarrow \quad \mathcal{M}(P_1, Q_1) \subset \mathcal{M}(P_2 T_1, -P_2 Y + Q_2).$$

Thus the results follows from Proposition 2.3(i). □

Proposition 2.8 Suppose f, f_0 are given as in (2.3), (2.6). Then we have

(i) f is injective $\Leftrightarrow$ f_0 is injective $\Leftrightarrow$ P_1, T_1 are right coprime

(ii) If $im\; Q_1 \subset im\; P_1$ and $T \in \mathcal{M}[D]^{r_2 \times r_1}$ is given as in Lemma 2.7 then

$$f \text{ is surjective } \Leftrightarrow f_0 \text{is surjective } \Leftrightarrow T, P_2 \text{ are left coprime}$$

(iii) If f is bijective then f^{-1} is also of the form

$$\begin{bmatrix} T' & Y' \\ 0 & I_m \end{bmatrix} \quad \text{with } T' \in \mathcal{M}[D]^{r_1 \times r_2}, Y' \in \mathcal{M}[D]^{r_1 \times m}.$$

Proof: It is trivial that f is injective (surjective) iff f_0 is injective (surjective). The remainder of the proof is completely analogous to the time-invariant situation, see *Hinrichsen and Prätzel-Wolters* (1980) Lemma 5.1, and therefore omitted. □

The following algebraic characterization of system equivalence generalizes the time-invariant result of *Hinrichsen and Prätzel-Wolters* (1980).

Proposition 2.9 Suppose

$$\mathbb{P}_i = \begin{bmatrix} P_i & -Q_i \\ V_i & W_i \end{bmatrix} \in \mathcal{M}[D]^{(r_i+p)\times(r_i+m)}$$

are system matrices, $i = 1,2$. Then $\mathbb{P}_1 \overset{se}{\approx} \mathbb{P}_2$ if and only if there exist $T, T_1 \in \mathcal{M}[D]^{r_2 \times r_1}$, $X \in \mathcal{M}[D]^{p \times r_1}$, $Y \in \mathcal{M}[D]^{r_2 \times m}$ such that

(i) $\begin{bmatrix} T & 0 \\ X & I_p \end{bmatrix} \mathbb{P}_1 = \mathbb{P}_2 \begin{bmatrix} T_1 & Y \\ 0 & I_m \end{bmatrix}$

(ii) T, P_2 are left coprime and P_1, T_1 are right coprime.

Proof: Suppose the system equivalence of $\mathbb{P}_1$ and $\mathbb{P}_2$ is described by f as in (2.3). Then $\mathcal{M}(P_1, Q_1) \subset \mathcal{M}(P_2 T_1, -P_2 Y + Q_2)$ and by Proposition 2.3 there exists $T \in \mathcal{M}[D]^{r_2 \times r_1}$ such that

$$T[P_1, Q_1] = [P_2, Q_2] \begin{bmatrix} T_1 & Y \\ 0 & I_m \end{bmatrix}.$$

Thus it remains to prove the existence of some $X \in \mathcal{M}[D]^{p \times r_1}$ so that

$$X\, P_1 + V_1 = V_2\, T_1 \quad \text{and} \quad X\, Q_1 + W_1 = V_2\, Y + W_2.$$

This follows from Proposition 2.3 since by (2.4)

$$\mathcal{M}(P_1, Q_1) \subset \mathcal{M}(V_1 - V_2\, T_1, W_2 - W_1 + V_2\, Y).$$

The coprimeness conditions hold true by Proposition 2.8.
To prove sufficiency define a map f as in (2.3). Then by (ii) and Proposition 2.8 f is a system-isomorphism. (2.4) follows from (i). This completes the proof. □

By using the algebraic characterization of system equivalence it will be shown how fairly rich the equivalence class of a system matrix is.

Corollary 2.10 Suppose

$$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(r+p) \times (r+m)}$$

is a system matrix and P is *equivalent* to P_1, i.e. $UP = P_1 U_1$ for some $U, U_1 \in GL_r(\mathcal{M}[D])$. Then

$$\mathbb{P} \overset{se}{\approx} \begin{bmatrix} P_1 & -UQ \\ VU_1^{-1} & W \end{bmatrix}$$

Proof: Since

$$\begin{bmatrix} U & 0 \\ 0 & I_p \end{bmatrix} \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} = \begin{bmatrix} P_1 & -UQ \\ VU_1^{-1} & W \end{bmatrix} \begin{bmatrix} U_1 & O \\ 0 & I_m \end{bmatrix}$$

the claim follows from Proposition 2.9. □

An important result is that in every equivalence class of system matrices lies an analytic state space system. More precisely

Proposition 2.11 Every system matrix

$$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(r+p) \times (r+m)}$$

is system equivalent to some

$$\mathbb{P}_{st} = \begin{bmatrix} DI_n & -B \\ C & E(D) \end{bmatrix} \in \mathcal{A}[D]^{(n+p)\times(n+m)}, \quad n = ord\ P$$

with B, C analytic matrices. $\mathbb{P}_{st}$ is uniquely determined up to a constant similarity transformation.

Proof: By Corollary 1.15 P is equivalent to $diag(1,\dots,1,D^n)$ and to DI_n. Thus by Corollary 2.10 one may assume that $\mathbb{P}$ is of the form

$$\mathbb{P} = \begin{bmatrix} DI_n & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(n+p)\times(n+m)}.$$

Let $B \in \mathcal{M}^{n\times n}$,$C \in \mathcal{M}^{p\times n}$,$Y \in \mathcal{M}[D]^{n\times n}$,$X \in \mathcal{M}[D]^{p\times n}$ such that

$$-Q = DI_n\, Y - B \text{ and } V = X\, DI_n + C.$$

Then for $E(D) := W + X\,Q - C\,Y$ one obtains

$$\begin{bmatrix} I_n & 0 \\ -X & I_p \end{bmatrix} \mathbb{P} = \mathbb{P}_{st} \begin{bmatrix} I_n & Y \\ 0 & I_m \end{bmatrix}, \quad \text{where } \mathbb{P}_{st} = \begin{bmatrix} DI_n & -B \\ C & E(D) \end{bmatrix}$$

and by Proposition 2.9 $\mathbb{P} \overset{se}{\sim} \mathbb{P}_{st}$.
It remains to show that $\mathbb{P}_{st}$ is defined over $\mathcal{A}[D]$. Since $im\ B \subset im\ DI_n$ it follows that B cannot have poles and thus $B \in \mathcal{A}^{n\times m}$. Since $\ker_{\mathcal{A}}\ DI_n = \mathbb{R}^n$ it is allowed to insert successively $u \equiv 0$ and $z = e_j \in \mathbb{R}^n (j \in \underline{n})$ into $y = Cz + E(D)(u)$. By assumption (A3) y does not have poles and thus $C \in \mathcal{A}^{p\times m}$. Using again assumption (A3) and the same trick as in the proof of Lemma 2.2 yields $E(D) \in \mathcal{A}[D]^{p\times m}$. This completes the proof. □

Now we are in a position to give an alternative definition of a solution homomorphism (see Definition 2.4(i)).

Proposition 2.12 Suppose

$$\mathbb{P}_i = \begin{bmatrix} P_i & Q_i \\ V_i & W_i \end{bmatrix} \in \mathcal{M}[D]^{(r_i+p)\times(r_i+m)}, \qquad i = 1,2$$

are system matrices. Then a $\mathbb{R}$−linear map $f : \mathcal{M}(P_1, Q_1) \to \mathcal{M}(P_2, Q_2)$ is a solution homomorphism if and only if f satisfies for every $(z,u) \in \mathcal{M}(P_1,Q_1)$ and every compact interval J there exists an $\mathbb{R}$ – linear map $f\mid_J$ such that the diagram commutes

$$\begin{tikzcd}
\mathcal{M}(P_1,Q_1) \arrow[r, "f"] \arrow[d, "\text{restriction on } J"'] & \mathcal{M}(P_2,Q_2) \arrow[d, "\text{restriction on } J"] \\
\mathcal{M}(P_1,Q_1)\mid_J \arrow[r, "f\mid_J"] & \mathcal{M}(P_2,Q_2)\mid_J
\end{tikzcd}$$

or equivalently for every $(z,u) \in \mathcal{M}(P_1,Q_1)$ and every compact interval J we have

$$(z,u)\mid_J = 0 \;\Rightarrow\; f(z,u)\mid_J = 0$$

For the proof the following theorem of *Peetre* (1960) will be applied.

Theorem 2.13 Suppose $Y : C_c^\infty \to C_c^\infty$ is a linear map, where

$$C_c^\infty := \{f \in C^\infty \mid \text{ supp } f \text{ is compact}\}.$$

Then Y is *local*, i.e. supp $Y\,(u) \subset$ supp u for all $u \in C_c^\infty$, if and only if

$$Y(u) = \sum_{i \geq 0} \alpha_i u^{(i)}$$

where $\{\alpha_i\}_{i \in \mathbb{N}_o}$ is a unique family of distributions which is *locally finite* , i.e. for every compact interval $J \subset \mathbb{R}$ $\alpha_i\mid_J = 0$ for i sufficiently large, and *locally contained* in C_c^∞, i.e. for every $t \in \mathbb{R}$ there exists a neighbourhood J_t of t such that $\alpha_i\mid_{J_t} \in C_c^\infty$.

Proof of Proposition 2.12 If f is defined by (2.3) then it satisfies (2.8). To prove sufficiency assume f satisfies(2.7). We proceed in several steps.

(i) By Proposition 2.11 $\mathbb{P}_i$ is system equivalent to some

$$\mathbb{P}^i_{st} = \begin{bmatrix} DI_n & -B_i \\ C_i & E_i(D) \end{bmatrix} \in \mathcal{A}[D]^{(n_i+p)\times(n_i+m)}, \quad i = 1,2 \text{ resp. .}$$

The system isomorphisms describing these equivalences satisfy (2.5). Thus it is sufficient to consider the case $\mathbb{P}_i = \mathbb{P}^i_{st}$, $i = 1,2$.

(ii) f induces a $\mathbb{R}$–linear map

$$f_0 : \ker_{\mathcal{A}} DI_{n_1} \to \ker_{\mathcal{A}} DI_{n_2}, \quad z \mapsto f_0(z).$$

Since $\ker_{\mathcal{A}} DI_{n_i} (i = 1,2)$ have constant bases, f_0 can be represented by some $T \in \mathbb{R}^{n_2 \times n_1}$. Therefore f is given by

$$f(z,u) = \begin{bmatrix} T & Y \\ 0 & I_m \end{bmatrix} \begin{pmatrix} z \\ u \end{pmatrix}$$

where $Y : \mathcal{U}^m \to \mathcal{U}^{n_2}$ is an $\mathbb{R}$-linear map.

(iii) It is shown that Y is local, i.e.

$$u \mid_J = 0 \;\Rightarrow\; Y(u) \mid_J = 0 \text{ for every } u \in \mathcal{U}^m \text{ and every compact interval } J.$$

If $u \mid_J = 0$ then $(z,u) \mid_J = (z^0, 0) \mid_J$ where $z^0 \in \ker DI_{n_1}$ is some constant free motion. Since $(z - z^0, u) \in \mathcal{M}(DI_{n_1}, B_1)$ and $(z - z^0, u) \mid_J = 0$, (2.7) yields $f(z - z^0, u) \mid_J = 0$. Now

$$f(z - z^0, u) \mid_J = (T_1(z - z^0) + Y(u), u) \mid_J = (T_1(z - z^0) \mid_J + Y(u) \mid_J, 0)$$

implies $Y(u) \mid_J = 0$.

(iv) If Y is restricted to $(\mathcal{C}_c^\infty)^m$ then by (iii) $Y((\mathcal{C}_c^\infty)^m) \subset (\mathcal{C}_c^\infty)^{n_2}$. Since Y is local an application of Theorem 2.13 yields that f can be presented by

$$f(z,u) = \begin{bmatrix} T & Y \\ 0 & I_m \end{bmatrix} \begin{pmatrix} z \\ u \end{pmatrix} \quad \text{for } u \in (\mathcal{C}_c^\infty)^m$$

where $Y = \sum_{i \geq 0} Y_i D^i I_m$ and the unique families $((Y_{k\ell}^i))_{i \in \mathbb{N}_0}$ are locally finite and locally contained in $\mathcal{C}_c^\infty$ for $k \in \underline{n_2}$, $\ell \in \underline{m}$, $Y_i = ((Y_{k\ell}^i))$.

(v) It remains to show that

$$Y = \sum_{i=0}^{k} Y_i\, D^i\, I_m \in \mathcal{A}^{n_2 \times m}[D] \qquad \text{for some } k \in \mathbb{N}_0.$$

Choose $u \in \mathcal{U}^m$ and $I = (t_0, \infty)$ so that $u(\cdot) \mid_I \equiv e_j$, $j \in \underline{m}$. Then for some $z \in (\mathcal{C}^\infty)^{n_1}$ $\dot{z} = Bu$ and thus $z \mid_I \in \mathcal{A} \mid_I^{n_1}$. Since Y_0 is locally contained in $(\mathcal{C}_c^\infty)^{n_2 \times m}$ for $t \in I$ there exists an open neighbourhood $J \subset I$ of t such that $Y_0 \mid_J \in (\mathcal{C}^\infty \mid_J)^{n_2 \times m}$. From

$$f(z,u) = (Tz + Y(u), u) \in \mathcal{M}(DI_{n_2}, B_2)$$

one obtains

$$(Y(u))^\cdot \mid_J = (Y_0 e_j)^\cdot \mid_J = -DI_{n_2}(Tz) \mid_J + B_2 e_j \mid_J \;\in \mathcal{A} \mid_J^{n_2}$$

and thus $Y_0 \mid_J \in \mathcal{A} \mid_J^{n_2 \times m}$. Since $t \in \mathbb{R}$ was arbitrary one obtains $Y_0 \in \mathcal{A}^{n_2 \times m}$.

Now inserting successively $u(t) = t^i e_j$ on I gives as above $Y_i \in \mathcal{A}^{n_2 \times m}$ for $i \geq 0$. Since $\{Y_i\}_{i \in \mathbb{N}}$ is a locally finite family, the identity property of analytic functions yields the existence of some $k \geq 0$ such that

$$Y = \sum_{i=0}^{k} Y_i\, D^i\, I_m \;\in \mathcal{A}^{n_2 \times m}[D].$$

This completes the proof. □

2.3 Controllability and observability

In Chapter 1, Definition 2.1 and 3.4 controllability and observability were introduced for state space systems. Now this will be generalized for systems in differential operator representation of the form (2.1).

Definition 3.1 Suppose

$$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(r+p)\times(r+m)}$$

is a system matrix.
Then $\mathbb{P}$ is called *controllable* on $[t_0, t_1]$, $t_0 < t_1$, if for every $z^0 \in \ker_{\mathcal{A}} P(D)$ there exists a control $u \in \mathcal{U}^m$ with supp $u \subset [t_0, t_1]$ such that

$$(z^0 + z_u)(t) = \begin{cases} z^0(t) & \text{for } t \le t_0 \\ 0 & \text{for } t \ge t_1 \end{cases}$$

where z_u denotes the unique forced motion, see Proposition 2.1.
$\mathbb{P}$ is called *observable* if V acts as a monomorphism on $\ker_{\mathcal{A}} P(D)$ or, in other words, $\ker_{\mathcal{A}} P \cap \ker_{\mathcal{A}} V = \{0\}$.

Note that by Proposition 1.2.2 these are extensions of the definitions concerning state space systems. Observability is not defined on an interval since if V is a monomorphism on $\ker_{\mathcal{A}_I} P(D)$ then, because P is full wrt $\mathcal{A}$, V is injective on $\ker_{\mathcal{A}} P(D)$ as well. In Proposition 3.4 it will be seen that the analogous fact holds true for controllability.

Controllability and observability are invariant under system equivalence. More precisely we have

Proposition 3.2 Suppose

$$\mathbb{P}_i = \begin{bmatrix} P_i & -Q_i \\ V_i & W_i \end{bmatrix} \in \mathcal{M}[D]^{(r_i+p)\times(r_i+m)}, \quad i = 1, 2$$

are system equivalent and $I = [t_0, t_1]$, $t_0 < t_1$. Then
$\mathbb{P}_1$ is controllable on I (observable) iff $\mathbb{P}_2$ is controllable on I (observable).

Proof: Let the system isomorphism be given by $f : \mathcal{M}(P_1, Q_1) \to \mathcal{M}(P_2, Q_2)$ defined in (2.3).
If $\mathbb{P}_1$ is controllable on I then for $z^1 \in \ker_{\mathcal{A}} P_1$ there exists $u \in \mathcal{U}^m$ with supp $u \subset [t_0, t_1]$ such that

$$(z^1 + z_u^1)(t) = \begin{cases} z^1(t) & \text{for } t \le t_0 \\ 0 & \text{for } t \ge t_1 \end{cases}$$

where $(z_u^1, u) \in \mathcal{M}_+(P_1, Q_1)$. Since f preserves the direct decomposition of $\mathcal{M}_+(P_1, Q_1)$,

$$(z_u^2, u) = (T_1(z_u^1) + Y(u), u) \in \mathcal{M}_+(P_2, Q_2).$$

Furthermore $u \mid_{(t_1,\infty)} = 0$ yields $Y(u) \mid_{(t_1,\infty)} = 0$ and thus

$$(z^2 + z_u^2)(t) = \begin{cases} z^2(t) & \text{for } t \le t_0 \\ 0 & \text{for } t \ge t_1 \end{cases}$$

Since for every $z^2 \in \ker_{\mathcal{A}} P_2$ there exists $z^1 \in \ker_{\mathcal{A}} P_1$ such that $T_1(z^1) = z^2$ it has been shown that $\mathbb{P}_2$ is controllable on I. Using the fact that f is a system isomorphism the converse direction is proved similary.
It remains to prove the statement for observability. Assume $\mathbb{P}_1$ is observable and $x \in \ker_{\mathcal{A}} P_2 \cap \ker_{\mathcal{A}} V_2$. Then, by using the algebraic criterion in Proposition 2.9 (i). $x = T_1(x')$ for a unique $x' \in \ker_{\mathcal{A}} P_1$ and $V_2 T_1(x') = X P_1(x') + V_1(x') = 0$. Thus $x' \in \ker_{\mathcal{A}} P_1 \cap \ker_{\mathcal{A}} V_2$ and since $\mathbb{P}_2$ is observable $x' = 0$. Therefore $x = 0$ and thus $\mathbb{P}_2$ is observable.
This completes the proof. □

In order to characterize controllability by left coprimeness of P and Q, as it has been done for state space systems in Theorem 1.2.4(v), a lemma is needed.

Lemma 3.3 If

$$\mathbb{P}_i = \begin{bmatrix} P_i & -Q_i \\ V_i & W_i \end{bmatrix} \in \mathcal{M}[D]^{(r_i+p)\times(r_i+m)}, \quad i = 1,2,$$

are system equivalent, then P_1, Q_1 are left coprime if and only if P_2, Q_2 are left coprime.

Proof: It is sufficient to show one direction. Using the notation of Proposition 2.9 one obtains $TP_1 = P_2 T_1$, $Q_2 = P_2 Y + TQ_1$ and condition (ii) yields the existence of some A, B, E, F of appropriate formats so that $I_{r_2} = TA + P_2 B$, $I_{r_1} = P_1 E + Q_1 F$.
Therefore

$$T = T P_1 E + T Q_1 F = P_2 T_1 E + (Q_2 - P_2 Y) F = P_2(T_1 E - Y F) + Q_2 F$$

and the following equations are equivalent

$$\begin{aligned} T A &= P_2(T_1 E A - Y F A) + Q_2 F A \\ I_{r_2} - P_2 B &= P_2(T_1 E A + B - Y F A) + Q_2 F A \end{aligned}$$

This proves the lemma. □

Proposition 3.4 For a system matrix

$$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(r+p)\times(n+m)}$$

the following statements are equivalent

(i) $\mathbb{P}$ is controllable on $[t_0, t_1]$, $t_0 < t_1$.

(ii) $\mathbb{P}$ is controllable on every interval.

(iii) P and Q are left coprime, i.e. there exist $X \in \mathcal{M}[D]^{r\times r}$, $Y \in \mathcal{M}[D]^{m\times r}$ such that

$$P X + Q Y = I_r$$

Proof: Because of Proposition 3.2 and Lemma 3.3 it is sufficient to consider the case $\mathbb{P} = \mathbb{P}_{st}$, where $\mathbb{P}_{st}$ is associated with an analytic state space system of the form (2.2). Now the equivalence (i) $\Leftrightarrow$ (ii) follows from the analyticity of the system, see Remark 1.2.5(i). (i) $\Leftrightarrow$ (iii) is proved in Theorem 1.2.4. Thus the proof is complete. □

Proposition 3.5 A system matrix

$$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(r+p)\times(r+m)}$$

is observable if and only if P and V are right coprime, i.e. there exist $X \in \mathcal{M}[D]^{r\times r}$, $Y \in \mathcal{M}[D]^{r\times p}$ sucht that

$$X\,P + Y\,V = I_r.$$

Proof: Let G denote a *gcrd* of P and V . Then by (1.6) $\ker_{\mathcal{A}}\, G = \ker_{\mathcal{A}}\, P \cap \ker_{\mathcal{A}}\, V$. If P and V are right coprime G is necessarily invertible and thus $\ker_{\mathcal{A}}\, G = \{0\}$ whence $\mathbb{P}$ is observable. Conversely, if P and V are not right coprime then G is not invertible and thus $\ker_{\mathcal{A}}\, G \neq \{0\}$. Hence $\mathbb{P}$ is not observable. This proves the proposition. □

2.4 Input-output map and formal transfer matrix

The decomposition of the solution vector space $\mathcal{M}(P,Q)$ (see Section 2) enables us to define an input-output map of systems in differential operator description. In general, there does not exist a frequency domain analysis for time-varying systems. However, one can define a formal transfer matrix and show its close relationship to the input-output map.

Definition 4.1 Suppose

$$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(r+p)\times(r+m)}$$

is a system matrix. The *input-output map* of the system associated with $\mathbb{P}$ is defined by

$$\begin{array}{rcll} G & : & \mathcal{U}^m & \rightarrow \mathcal{U}^p \\ & & u & \mapsto V(D)(z_u) + W(D)(u) \end{array}$$

where z_u denotes the forced motion starting from zero, $(z_u, u) \in \mathcal{M}_+(P,Q)$.

Remark 4.2 The input-output maps of two system equivalent systems coincide. This is immediate from (2.3) and the fact that $f(\mathcal{M}_+(P_1,Q_1)) \subset \mathcal{M}_+(P_2,Q_2)$.

In order to define the formal transfer matrix we have to introduce the left-skew field of fractions of $\mathcal{M}[D]$

$$\mathcal{M}(D) := \{p^{-1}q \mid p \in \mathcal{M}[D]^*, q \in \mathcal{M}[D]\}$$

This field is constructed as follows (cf. *Cohn* (1971), p. 20):
For pairs $(p,q) \in \mathcal{M}[D]^* \times \mathcal{M}[D]$ define an equivalence relation between them by the condition: $(p_1,q_1) \sim (p_2,q_2)$ iff there exist $u_1, u_2 \in \mathcal{M}[D]^*$ such that

$$u_2 p_1 = u_1 p_2 \quad \text{and} \quad u_2 q_1 = u_1 q_2$$

The equivalence class containing a pair (p,q) is denoted by $p^{-1}q$. The multiplication

$$p_1^{-1}q_1 \cdot p_2^{-1}q_2 := (u_2 p_1)^{-1}(u_1 q_2) \quad \text{with } u_1, u_2 \in \mathcal{M}[D]^* \text{ such that } u_1 p_2 = u_2 q_1$$

depends only on the equivalence classes of the factors and is associative.

Definition 4.3 Suppose

$$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(r+p)\times(r+m)}$$

is a system matrix. Then

$$\hat{G} = VP^{-1}Q + W \in \mathcal{M}(D)^{p\times m}$$

is called the *formal transfer matrix* of the system associated with $\mathbb{P}$.

Proposition 4.4 If two system matrices are system equivalent then their formal transfer matrices coincide.

Proof: Suppose $\mathbb{P}_i, i = 1, 2$, satisfy condition (i) in Proposition 2.9. Then it follows that

$$\begin{aligned} V_1P_1^{-1}Q_1 + W_1 &= (V_2T_1P_1^{-1} - X)Q_1 + W_1 = V_2Y + V_2T_1P_1^{-1}Q_1 + W_1 \\ &= V_2P_2^{-1}(P_2Y + TQ_1) + W_2 = V_2P_2^{-1}Q_2 + W_2 \end{aligned}$$

and the proposition is proved. □

In the following proposition the relationship between the input output map

$$G : \mathcal{U}^m \to \mathcal{U}^p, \quad u \mapsto V(D)(z_u) + W(D)(u)$$

and the formal transfer matrix as a multiplication operator

$$\hat{G} : \mathcal{M}(D)^m \to \mathcal{M}(D)^p, \quad u(D) \mapsto V(D)P(D)^{-1}Q(D)u(D) + W(D)u(D)$$

will be clarified.

Proposition 4.5 Suppose the system matrices

$$\mathbb{P}_i = \begin{bmatrix} P_i & -Q_i \\ V_i & W_i \end{bmatrix} \in \mathcal{M}[D]^{(r_i+p)\times(r_i+m)}$$

have input-output maps G_i and formal transfer matrices $\hat{G}_i, i = 1, 2$. Then we have

$$\hat{G}_1 = \hat{G}_2 \Leftrightarrow G_1(u) = G_2(u) \qquad \text{for all } u \in \mathcal{U}^m$$

Proof: By Remark 4.2 and Proposition 4.4 it is assumed without restriction of generality that

$$\mathbb{P}_i = \mathbb{P}^i_{st} = \begin{bmatrix} DI_{n_i} & -B \\ C_i & E_i(D) \end{bmatrix} \in \mathcal{A}[D]^{(n_i+p)\times(n_i+m)}, \qquad i = 1, 2.$$

The multiplication rule (1.4) in Chapter 1 and multiplication in $\mathcal{M}(D)$ yields

$$c\,D^{-1} = (D - \frac{\dot{c}}{c})^{-1}c \quad \text{for all } c \in \mathcal{A} \setminus \{0\} \tag{4.1}$$

Denote $C_\lambda = ((c^\lambda_{ij}))$ for $\lambda = 1, 2$. Then

$$C_\lambda D^{-1} I_{n_\lambda} B_\lambda = (([D - \frac{\dot{c}^\lambda_{ij}}{c^\lambda_{ij}}]^{-1} c^\lambda_{ij}))B_\lambda, \quad \lambda = 1, 2. \tag{4.2}$$

where the entries of

$$(([D - \frac{\dot{c}_{ij}^{\lambda}}{c_{ij}^{\lambda}}]^{-1} c_{ij}^{\lambda}))$$

are defined to be zero if $c_{ij}^{\lambda} = 0$. Put

$$\begin{aligned} k(D) &= \ell c\ell m\{D - \tfrac{\dot{c}_{ij}^{\lambda}}{c_{ij}} \mid c_{ij}^{\lambda} \neq 0, i \in \underline{p}, j \in \underline{n}, \lambda = 1,2\} \\ &= s_{ij}^{\lambda}(D)(D - \tfrac{\dot{c}_{ij}^{\lambda}}{c_{ij}^{\lambda}}) \end{aligned}$$

for some $s_{ij}^{\lambda}(D) \in \mathcal{M}[D]$, $i \in \underline{p}$, $j \in \underline{n}$, $\lambda = 1,2$ resp.
Now by (4.2) the entries of

$$k(D)\, C_{\lambda}\, D^{-1} I_{n_{\lambda}} = ((s_{ij}^{\lambda}(D) c_{ij}^{\lambda})) \in \mathcal{M}[D]^{p\times n} \tag{4.3}$$

are polynomials and therefore it is allowed to write

$$\begin{aligned} (k(D)C_{\lambda}(t))(z_u^{\lambda}(t)) &= (k(D)C_{\lambda}(t)D^{-1})(D(z_u^{\lambda}(t))) \\ &= (k(D)C_{\lambda}(t)D^{-1})(B_{\lambda}(t)u(t)) \end{aligned} \tag{4.4}$$

where $(z_u^{\lambda}, u) \in \mathcal{M}_+(DI_{n_{\lambda}}, B_{\lambda})$ for $\lambda = 1,2$. This enables us to prove the proposition.

$$C_1 D^{-1} I_{n_1} B_1 + E_1(D) \;=\; \hat{G}_1 \;=\; \hat{G}_2 \;=\; C_2 D^{-1} I_{n_2} B_2 + E_2(D)$$

is valid if and only if

$$k(D)C_1D^{-1}B_1 + k(D)E_1(D) \;=\; k(D)C_2D^{-1}B_2 + k(D)E_2(D) \tag{4.5}$$

Since the entries of the matrices in (4.5) are polynomials Lemma 2.2 yields that (4.5) is equivalent to

$$(k(D)C_1D^{-1})(B_1u) + (k(D)E_1(D))(u) = (k(D)C_2D^{-1})(B_2u) + (k(D)E_2(D))(u) \text{ for all } u \in \mathcal{U}^m$$

which by (4.4) can be rewritten as

$$k(D)(C_1z_u^1) + k(D)(E_1(D)(u)) \;=\; k(D)(C_2z_u^2) + k(D)(E_2(D)u(t)) \quad \text{for all } u \in \mathcal{U}^m \tag{4.6}$$

By Proposition 1.5(iv) $k(D)$ is full wrt $\mathcal{A}$ since it is the $\ell c\ell m$ of full polynomials. For all $u \in \mathcal{U}^m$ and $\lambda = 1,2$ $\;C_{\lambda}z_u^{\lambda} + E_{\lambda}(D)(u)$ has a support bounded to the left and thus (4.6) holds true if and only if

$$G_1(u) \;=\; C_1z_u^1 + E_1(D)(u) \;=\; C_2z_u^2 + E_2(D)(u) \;=\; G_2(u) \quad \text{for all } u \in \mathcal{U}^m$$

This completes the proof. □

Corollary 4.6 Suppose two analytic state space systems associated with

$$\mathbb{P}_{st}^{i} = \begin{bmatrix} DI_{n_i} - A_i & -B_i \\ C_i & E_i(D) \end{bmatrix} \in \mathcal{A}[D]^{(n_i+p)\times(n_i+m)}$$

$(i = 1,2)$ satisfy $\hat{G}_1 = \hat{G}_2$ or $G_1(u) = G_2(u)$ for all $u \in \mathcal{U}^m$. Then $E_1(D) = E_2(D)$.

Proof: By Remark 1.1.1 $\mathbb{P}^i_{st}$ can be transformed to a system with constant free motion and $E_i(D)$ is not changed, $i = 1,2$ resp. Thus by Proposition 4.4 and Remark 4.2 it is assumed without restriction of generality that $A_i = 0$ for $i = 1,2$. Using the notation of the proof of Proposition 4.5 gives

$$k(D)[C_1D^{-1}B_1 - C_2D^{-1}B_2] = k(D)[E_2(D) - E_1(D)] \tag{4.7}$$

Let $\ell := \deg k(D)$. Then the degree of the left hand side polynomial matrix in (4.7) is smaller than ℓ and by comparing coefficients in (4.7) one obtains $D^\ell[E_2(D) - E_1(D)] = 0$. This proves the corollary. □

2.5 $\mathcal{M}[D]$-right modules and the input module

In Section 2 we analysed the solution vector space $\mathcal{M}(P,Q)$ of a system matrix

$$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(r+p)\times(r+m)}$$

and considered the *differential* equation

$$P(D)(z) = Q(D)(u) \qquad \text{for } (z,u) \in (\mathcal{C}^\infty)^r \times \mathcal{U}^m.$$

Now we will study the *algebraic* equation

$$P(D)z(D) = Q(D)u(D) \qquad \text{for } (z(D),u(D)) \in \mathcal{M}[D]^r \times \mathcal{M}[D]^m$$

It is important to distinguish between two operators induced by

$$P(D) = \sum_{i=0}^{k} P_i\, D^i \in \mathcal{M}^{r\times n}[D]$$

One is P as a *differential operator* acting on $(\mathcal{C}^\infty)^n$

$$\begin{array}{rcccl} P(D) & : & (\mathcal{C}^\infty)^n & \to & (\mathcal{C}^\infty)^r \\ & & z & \mapsto & P(D)(z) = \sum_{i=0}^k P_i z^{(i)} \end{array}$$

the other is P as the formal *multiplication operator* acting on $\mathcal{M}[D]^n$

$$\begin{array}{rcccl} P(D) & : & \mathcal{M}[D]^n & \to & \mathcal{M}[D]^r \\ & & z(D) & \mapsto & P(D)z(D) = \sum_{i=0}^k P_i D^i z(D) \end{array}$$

First, submodules of the free $\mathcal{M}[D]$-right module $\mathcal{M}[D]^r$ will be studied. If **M** is a right-(left-) $\mathcal{M}[D]$-module its *rank* is the cardinality of any maximal right-(left-) linearily independent (over $\mathcal{M}[D]$) subset of elements of **M** , see *Cohn* (1971) p.28. Since $\mathcal{M}[D]$ is a right and left Euclidean domain it follows for the free $\mathcal{M}[D]$-right module $\mathcal{M}[D]^r$ that each of its submodules is also free and of rank at most r , see *Cohn* (1971) p.46.
For a matrix $P \in \mathcal{M}[D]^{r\times k}$ the *column (row) rank* is defined as the rank of the right (left) $\mathcal{M}[D]$-submodule of $\mathcal{M}[D]^r(\mathcal{M}[D]^{1\times k})$ spanned by the columns (rows) of P . Both ranks coincide, see *Cohn* (1971) p. 195.

Definition 5.1 For $v(D) = (v^1(D), \ldots, v^r(D))^T \in \mathcal{M}[D]^r$ set

$$\deg\ v(D) = \max_{i \in \underline{r}} \deg\ v^i(D)$$

Let $V(D) = [v_1(D), \ldots, v_k(D)] \in \mathcal{M}[D]^{r \times k}$, then

$$\lambda_i := \deg\ v_i(D) \quad \text{is called the i}-\text{th index of } V,$$

$$\mathbf{W} = V(D) \cdot \mathcal{M}[D]^k$$

denotes the right $\mathcal{M}[D]$ submodule generated by V.
V is called *(ordered) minimal basis* of $\mathbf{W}$ if its columns are linearly independent over $\mathcal{M}[D]$ and the sum of its indices, $\sum_{i=1}^{k} \lambda_i$, is minimal among all bases of $\mathbf{W}$ (and $\lambda_1 \geq \ldots \geq \lambda_k$).
If $v_i(D) = \sum_{j=0}^{\lambda_i} D^j v_{ij}$ for $i \in \underline{k}$, then

$$[V(D)]_\ell := [v_{1,\lambda_1}, \ldots, v_{k,\lambda_k}] \in \mathcal{M}^{r \times k}, \ \ v_{i,\lambda_i} = 0 \text{ if } v_i(D) = 0$$

denotes the *leading (column) coefficient matrix* of V. The definition of this matrix does not depend on the side of which the coefficients of $v_i(D)$ are written.

Using these notations a minimal basis of $\mathbf{W}$ can be charactzerized as follows.

Theorem 5.2 Suppose $\mathbf{W} = V(D) \cdot \mathcal{M}[D]^k$ and $rk_{\mathcal{M}[D]} V(D) = k$. Then the following are equivalent:

(i) $V(D)$ is a minimal basis of W.

(ii) $rk_{\mathcal{M}}[V(D)]_\ell = k$

(iii) For any $x(D) = (x^1(D), \ldots, x^k(D))^T \in \mathcal{M}[D]^k \setminus \{0\}$

$$\deg\ V(D) \cdot x(D) = \max_{i \in \underline{k}} \{\lambda^i + \deg\ x^i(D) \mid x^i(D) \neq 0\}$$

(iv) For $d \in \mathbb{N}_0$ the $\mathcal{M}$−vector space

$$\mathbf{W}_d := \{v(D) \in \mathbf{W} \mid \deg\ v(D) \leq d\}$$

has dimension

$$\dim_{\mathcal{M}} \mathbf{W}_d = \sum_{i:\lambda_i \leq d} (d + 1 - \lambda_i)$$

Proof: (i) $\Rightarrow$ (ii) : Assume $(m_1, \ldots, m_k)^T \in \mathcal{M}^k \setminus \{0\}$ such that $\sum_{i=1}^{k} v_{i,\lambda_i} m_i = 0$ and λ_p is the maximal index with $\lambda_p \neq 0$. Then

$$\begin{aligned} v' &:= \sum_{i=1}^{k} v_i D^{(\lambda_p - \lambda_i)} m_i \\ &= \sum_{i=1}^{k} \Big(\sum_{j=0}^{\lambda_i - 1} D^j v_{ij} + D^{\lambda_i} v_{i,\lambda_i}\Big) D^{(\lambda_p - \lambda_i)} m_i \\ &= \sum_{i=1}^{k} \sum_{j=0}^{\lambda_i - 1} D^j v_{ij} D^{(\lambda_p - \lambda_i)} m_i + \sum_{i=1}^{k} D^{\lambda_i} (D^{(\lambda_p - \lambda_i)} v_{i,\lambda_i} + w_i) m_i \end{aligned}$$

with w_i such that deg $w_i < \lambda_p - \lambda_i$

$$= w + D^{\lambda_p} \sum_{i=1}^{k} v_{i,\lambda_i} m_i$$

with w such that deg $w < \lambda_p$

$$= w.$$

Since

$$v_p = (v' - \sum_{i=1, i \neq p}^{k} v_i D^{(\lambda_p - \lambda_i)} m_i) m_p^{-1}$$

the matrix $[v_1, \ldots, v_{p-1}, v', v_{p+1}, \ldots, v_m]$ is a basis with lower order than V. This contradicts (i).

(ii) $\Rightarrow$ (iii): Let $x = (x_1, \ldots, x_k)^T \in \mathcal{M}[D]^k \setminus \{0\}$. Then

$$\deg Vx \;=\; \deg \sum_{i=1}^{k} v_i x_i \;\leq\; \max\{\deg x_i + \lambda_i \mid x_i \neq 0\} \;=:\, a$$

Let $\ell_i := \deg\ x_i$ for $i \in \underline{k}$, and $\mathcal{N} := \{i \in \underline{k} \mid \ell_i + \lambda_i = a\}$. Then

$$\begin{aligned}
Vx &= \sum_{i=1}^{k} \sum_{j=0}^{\lambda_i} D^j v_{ij} \sum_{\mu=0}^{\ell_i} D^\mu x_{i\mu} \\
&= \sum_{i=1}^{k} \sum_{j=0}^{\lambda_i} D^j \sum_{\mu=0}^{\ell_i} (D^\mu v_{ij} + y_{\mu ij}) x_{i\mu}, \qquad \text{with } y_{\mu ij} \text{ such that deg } y_{\mu ij} < \mu \\
&= D^a \sum_{i \in \mathcal{N}} v_{i\lambda_i} x_{i\ell_i} + y, \qquad \text{with } y \text{ such that deg } y < a
\end{aligned}$$

By (ii) $\sum_{i \in \mathcal{N}} v_{i,\lambda_i} x_{i\ell_i} \neq 0$ whence (iii) follows.

(iii) $\Rightarrow$ (iv): For arbitrary $\alpha \in \mathbb{N}$ and $x = (x^1, \ldots, x^k)^T \in \mathcal{M}[D]^k$ one has

$$\deg Vx < \alpha \;\Leftrightarrow\; \deg x^i < \alpha - \lambda_i \qquad \text{for all } i \in \underline{k}, x^i \neq 0$$

Since

$$\dim_{\mathcal{M}} \mathbf{W}_d = \sum_{i=1}^{k} \dim_{\mathcal{M}} \{V e_i x^i \mid x^i \in \mathcal{M}[D], \deg V e_i x^i \leq d\},$$

where e_i denotes the i-th unit vector in $\mathbb{R}^k$, it follows that

$$\dim_{\mathcal{M}} \mathbf{W}_d = \sum_{i: \lambda_i \leq d} (d + 1 - \lambda_i)$$

(iv) $\Rightarrow$ (i): If the numbers of indices of V equal to d are denoted by $h(d) = \sum_{i:\lambda_i = d} 1,\ d \in \mathbb{N}_0$, then (iv) yields

$$\begin{aligned}
h(d) &= \sum_{i:\lambda_i \leq d} (d+1-\lambda_i) + (d-1-\lambda_i) - 2(d-\lambda_i) - \sum_{i:\lambda_i = d} (d-1-\lambda_i) \\
&= \sum_{i:\lambda_i \leq d} (d+1-\lambda_i) + \sum_{i:\lambda_i \leq d} (d-2+1-\lambda_i) - \sum_{i:\lambda_i = d} (d-1-\lambda_i) - 2 \sum_{i:\lambda_i \leq d} (d-1+1-\lambda_i) \\
&= \sum_{i:\lambda_i \leq d} (d+1-\lambda_i) + \sum_{i:\lambda_i \leq d-2} (d-2+1-\lambda_i) - 2 \sum_{i:\lambda_i \leq d-1} (d-1+1-\lambda_i) \\
&= \dim_{\mathcal{M}} \mathbf{W}_d + \dim_{\mathcal{M}} \mathbf{W}_{d-2} - 2 \dim_{\mathcal{M}} \mathbf{W}_{d-1}
\end{aligned}$$

The equation follows for $d \in \{0,1\}$ by direct calculation and defining $\dim_{\mathcal{M}} \mathbf{W}_d = 0$ for $d = 0,1$. This proves that, if (iv) is valid, then $h(d)$ is only determined by the module, not by the specific basis. If $w = [w_1, \ldots, w_k]$ is another basis of $\mathbf{W}_n$ then $\sum_{i=1}^{k} \deg\ w_i = \sum_{d \geq 1} d\, h(d)$. Since (i) $\Rightarrow$ (iv) has already been shown, every minimal basis $W = [w_1, \ldots, w_k]$ necessarily satisfies $\sum_{i=1}^{k} \deg\ w_i = \sum_{d \geq 1} d\, h(d)$. Now if (iv) is valid then for V as in the proposition

$$\sum_{i=1}^{k} \deg\ v_i = \sum_{i=1}^{k} = \sum_{d \geq 1} d\, h(d)$$

and thus V is minimal. □

The foregoing theorem is a generalization of *Forney's* Main Theorem (1975), see also *Münzner and Prätzel-Wolters* (1979) p. 293.

Remark 5.3

(i) The last part of the proof of Theorem 5.2 shows that the families of indices of different minimal bases of $\mathbf{W}$ coincide, they do only depend on the dimension of $\mathbf{W}_d$.

(ii) The direction (i) $\Rightarrow$ (ii) of the proof of Theorem 5.2 leads to an algorithm which transforms an arbitrary basis $V(D)$ of $\mathbf{W}$ in finitely many steps into a minimal basis $V'(D)$ of $\mathbf{W}$.

The set of basis transformation matrices which transform a minimal basis can be characterized as follows.

Proposition 5.4 Suppose $V(D) \in \mathcal{M}[D]^{r \times k}$ is an ordered minimal basis of submodule $\mathbf{W}$ with indices $\lambda_1, \ldots, \lambda_k$. Then $\overline{V}(D() = V(D)T(D)$ is an ordered minimal basis of $\mathbf{W}$ iff $T(D)$ is invertible over $\mathcal{M}[D]$ and satisfies

$$\begin{array}{rcll} \deg\ t_{ij}(D) & \leq & \lambda_j - \lambda_i & \text{for } \lambda_i \leq \lambda j \\ t_{ij}(D) & = & 0 & \text{for } \lambda_i > \lambda_j \end{array}$$

i.e. $T(D)$ is of the form

$$T = \begin{bmatrix} \star & & \\ & \ddots & 0 \\ \star & & \\ & & \star \end{bmatrix} \in GL_k(\mathcal{M}[D])$$

where the square diagonal blocks have meromorphic entries and the formats are corresponding to the multiple of the indices.

Proof: Put

$$V = [v_1, \ldots, v_k],\ \overline{V} = [\overline{v}_1, \ldots, \overline{v}_k] \quad \text{and} \quad T = (t_{ij}).$$

If $\overline{V}$ is also a minimal basis of $\mathbf{W}$ then $T \in GL_k(\mathcal{M}[D])$ and minimality of $\overline{V}$ yields

$$\deg \overline{v}_j = \deg \sum_{i=1}^{k} t_{ij}\, v_i = \max_{i \in \underline{k}}\{\deg\ t_{ij} + \lambda_i \mid t_{ij} \neq 0\} = \lambda_j$$

Therefore

$$\deg\ t_{ij} \leq \lambda_j - \lambda_j \ \text{ for } \ \lambda_i \leq \lambda_j \ \text{ and } \ t_{ij} = 0 \ \text{ for } \ \lambda_i > \lambda_j.$$

To prove the converse direction note that

$$\deg \overline{v}_j \;=\; \deg \sum_{i=1}^{k} t_{ij}\, v_i \;\leq\; \max_{i \in \underline{k}}\; \deg v_i\, t_{ij} \;\leq\; \max_{i \in \underline{k}}\; \lambda_i + (\lambda_j - \lambda_i) = \lambda_j$$

This proves minimality of $\overline{V}$ and the proof is complete. □

The foregoing results will now be applied to systems associated with

$$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(r+p)\times(r+m)}.$$

Consider the right $\mathcal{M}[D]$ module

$$\ker\,[P, -Q] := \{x \in \mathcal{M}[D]^{r \times m} \mid [P, -Q]x = 0\}$$

and the so called *input module*

$$\mu(P, Q) := \{u \in \mathcal{M}[D]^m \mid \exists z \in \mathcal{M}[D]^r : \begin{pmatrix} z \\ u \end{pmatrix} \in \ker\,[P, -Q]\}$$

Suppose $\begin{bmatrix} Z \\ U \end{bmatrix} \in \mathcal{M}[D]^{(r+m)\times k}$ is a $\mathcal{M}[D]$ basis of $\ker[P, -Q]$, then $U \in \mathcal{M}[D]^{m \times k}$ is a basis of $\mu(P, Q)$. This is seen as follows. If $U\,a = 0$ for some $a \in \mathcal{M}[D]^k$ then $P\,Z\,a = Q\,U\,a = 0$ and since P is invertible over $\mathcal{M}(D)$, $Z\,a = 0$.

As an immediate result we have:

Proposition 5.5 If two system matrices

$$\mathbb{P}_i = \begin{bmatrix} P_i & -Q_i \\ V_i & W_i \end{bmatrix} \in \mathcal{M}[D]^{(r_i+p)\times(r_i+m)} \qquad i = 1, 2$$

are system equialent then $\mu(P_1, Q_1) = \mu(P_2, Q_2)$.

Proof: Using the notation of Proposition 2.9(i) yields that

$$\ker[P_1, -Q_1] \to \ker[P_2, -Q_2], \quad \begin{pmatrix} z \\ u \end{pmatrix} \mapsto \begin{pmatrix} T_1(z) + Y(u) \\ u \end{pmatrix}$$

is a $\mathbb{R}$-homomorphism. Thus $\mu(P_1, Q_1) \subset \mu(P_2, Q_2)$ and since system equivalence is a symmetric relation the proof is complete. □

For state space systems of the form (2.2) the input module can be characterized in terms of $(DI_n - A)^i(B)$. For this we need the right $\mathcal{M}$ -homomorphism

$$\begin{array}{rcl} \hat{K}_{A,B} \;:\; \mathcal{M}[D]^m & \to & \mathcal{M}^n \\ \sum_{i=0}^{k} D^i u_i & \mapsto & K^k(A, B)(u_0, \ldots, u_k)^T \;=\; \sum_{i=0}^{k} (-1)^i (DI_n - A)^i(B) u_i \end{array}$$

Proposition 5.6 Let $(A, B) \in \mathcal{A}^{n\times(n+m)}$. Then

$$\mu(DI_n - A, B) = \ker\, \hat{K}_{A,B}$$

and

$$\dim_{\mathcal{M}[D]} \mu(DI_n - A, B) = m.$$

Proof: The following multiplication rules, which are easily proved by induction, will be used.

$$\begin{aligned} N\ D^i &= \sum_{\lambda=0}^{i}(-1)^\lambda \binom{i}{\lambda} D^{i-\lambda} N^{(\lambda)} \quad \text{for } i \in \mathbb{N},\ N \in \mathcal{M}^{n\times m} && (5.1)\\ N\sum_{i=0}^{k} D^i\, v_i &= \sum_{i=0}^{k} D^i \sum_{\lambda=i}^{k}(-1)^{\lambda-i}\binom{\lambda}{\lambda-i} N^{(\lambda-i)} v_\lambda \\ &\qquad \text{for } k \in \mathbb{N},\ N \in \mathcal{M}^{n\times m},\ v_1,\ldots,v_k \in \mathcal{M}^m. && (5.2)\end{aligned}$$

Using Remark 1.1.1 and equation 1(1.6) it is easily seen that it is sufficient to consider the case $A = 0$.

Suppose $u(D) = \sum_{i=0}^{k} D^i\, u_i \in \mu(DI_n, B)$. Then there exists $x(D) = \sum_{i=0}^{k-1} D^i\, x_i \in \mathcal{M}[D]^n$ so that

$$D\, x(D) = B\, u(D)$$

which by (5.2) is equivalent to

$$D\sum_{i=0}^{k-1} D^i\, x_i = \sum_{i=0}^{k} D^i \sum_{\lambda=i}^{k}(-1)^{\lambda-i}\binom{\lambda}{\lambda-i} B^{\lambda-i} u_\lambda \tag{5.3}$$

Comparing the coefficients in (5.3) yields for $i = 1,\ldots,k$

$$0 = \sum_{\lambda=0}^{k}(-1)^\lambda\, B^{(\lambda)}\, u_\lambda,\ x_{i-1} = \sum_{\lambda=i}^{k}(-1)^{\lambda-i}\binom{\lambda}{\lambda-i} B^{(\lambda-i)}\, u_\lambda \tag{5.4}$$

Thus $u(D) \in \ker \hat{K}_{0,B}$. Conversely, if $u(D) \in \ker \hat{K}_{0,B}$ then define $x(D)$ by (5.4) and (5.3) is valid. This shows $u(D) \in \mu(DI_n, B)$ and the first statement of the proof is shown. To prove the second equality use H defined in 1(4.2). Then it is obvious that for every $\ell \in \underline{m}$ there exist

$$u_i^\ell = (0,\ldots,0,\underset{\ell\text{-th element}}{*},0,\ldots,0) \in \mathcal{M}^m, \qquad i = 0,\ldots,k_\ell$$

so that

$$u_\ell(D) := \sum_{i=0}^{k_\ell} D^i\, u_i^\ell \in \ker \hat{K}_{0,B}.$$

Since the vectors $u_1(D),\ldots,u_m(D)$ are linearly independent the second equality is clear. □

2.6 Invariants of system equivalence resp. similarity

In Section 1.4 two families of invariants of state space systems were introduced: the controllability- and the geometric indices. They are both invariant with respect to similarity. In this section I shall present two other families of invariants and - using the unifying power of the input module - prove that (roughly speaking) they all four coincide.

Forney (1975) Chapter 7 conciders proper rational input-output maps $G(s) : \mathbb{R}(s)^m \to \mathbb{R}(s)^p$ and the minimal indices of the rational vector space $\{(u(s), G(s)u(s))^T \mid u(s) \in \mathbb{R}(s)^m\}$ which he calls the dynamical indices of $G(s)$. *Münzner and Prätzerl-Wolters* (1979) show that these indices coincide with those of the module $\{u(s) \in \mathbb{R}[s]^m \mid G(s)u(s) \in \mathbb{R}[s]^p\}$. By using Section 4 and 5 we are now in a position to carry over Forney's approach to time-varying polynomial systems.

Definition 6.1 Let

$$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(n+p)\times(n+m)}$$

be a system matrix with formal transfer matrix $\hat{G} = VP^{-1}Q + W \in \mathcal{M}(D)^{p\times m}$.
The indices of the minimal basis of $\mu(P,Q)$ are called the *minimal indices* of $\mathbb{P}$.
The indices of a minimal basis of the $\mathcal{M}[D]-$right module

$$\mathbf{M}_{\hat{G}} := \{u \in \mathcal{M}[D]^m \mid \hat{G}u \in \mathcal{M}[D]^p\}$$

are called the *dynamical indices* of $\hat{G}$.

By Remark 5.3(i) the minimal and dynamical indices are well defined.
Proposition 5.5 resp. Proposition 4.4 imply that the minimal resp. dynamical indices are invariant with respect to system equivalence.

Proposition 6.2 If $\mathbb{P}$ as in Definition 6.1 is observable then

$$\mathbf{M}_{\hat{G}} = \mu(P,Q)$$

and thus the families of minimal and dynamical indices of an observable system coincide.

Proof: If $z \in \mathcal{M}[D]^n$ such that $P\,z = Q\,u$ then $\hat{G}\,u = V\,z + W\,u$. Thus $u \in \mathbf{M}_{\hat{G}}$. To prove the converse note that since $\mathbb{P}$ is observable there exist $X \in \mathcal{M}[D]^{n\times n}$, $Y \in \mathcal{M}[D]^{n\times p}$ such that $I_r = X\,P + Y\,Y$. Thus, if $u \in \mathbf{M}_{\hat{G}}$,

$$z := P^{-1}Q\,u \;=\; (X\,P + Y\,V)P^{-1}Q\,u \;=\; X\,Q\,u + Y\,\hat{G}\,u \in \mathcal{M}[D]^n.$$

□

Theorem 6.3 For $(A,B) \in \mathcal{A}^{n\times(n+m)}$ the families of controllability, geometric and minimal indices coincide.

Proof: The controllability-, geometric- and minimal indices are denoted by $k_1,\ldots,k_m$; $\alpha_1(\cdot),\ldots,\alpha_m(\cdot)$ and $\lambda_1,\ldots,\lambda_m$, resp.

(i) It is shown that the family $(\alpha_i(\cdot))_{i\in\underline{m}}$ coincides with the family $(\lambda_i)_{i\in\underline{m}}$. The proof is similar to the time-invariant case, see *Münzner and Prätzel-Wolters* (1979) p. 298, and is generalized as follows. Put $\mathbf{V}_d := \{v \in \mathcal{M}[D]^m \mid \deg v \le d\}$, $d \in \mathbb{N}_0$. Then

$$\hat{K}_{A,B}(\mathbf{V}_d) \;=\; \{K^d(A,B)(u_0,\ldots,u_d)^T \mid u_0,\ldots,u_d \in \mathcal{M}^m\}$$

where the map $\hat{K}_{A,B}$ is defined in Section 5. Since the map

$$\begin{array}{rcll} f & : & \hat{K}_{A,B}(V_d) & \rightarrow \; V_d/V_d \cap \ker \hat{K}_{A,B} \\ & & K^d(u_0,\ldots,u_d)^T & \mapsto \; u_0 + \ldots + D^d\,u_d + (\mathbf{V}_d \cap \ker \hat{K}_{A,B}) \end{array}$$

is a $\mathcal{M}-$right homomorphism one obtains for $\mathbf{M}_d := \mathbf{V}_d \cap \ker \hat{K}_{A,B}$

$$\dim_{\mathcal{M}} \hat{K}_{A,B}(\mathbf{V}_d) \;=\; m(d+1) - \dim_{\mathcal{M}} \mathbf{M}_d \qquad (6.1)$$

From now on the system is considered on a non-void open interval where the $r_i(t)'s$, defined in Section 1.4, are constants. Then it remains to prove that

$$\sum_{j:\lambda_j=d} 1 =: h(d) = k(d) := \sum_{j:\alpha_j=d} 1 \tag{6.2}$$

Using the prove of (iv) $\Rightarrow$ (i) in Theorem 5.2 one has

$$h(d) \;=\; \dim_{\mathcal{M}}\ \mathbf{M}_d + \dim_{\mathcal{M}}\ \mathbf{M}_{d-2} - 2\ \dim\ \mathbf{M}_{d-1} \tag{6.3}$$

Since

$$\sum_{d:d\geq\ell} k(d) \;=\; \sum_{i:\alpha_i\geq\ell} 1 \;=\; \sum_{i:1\leq i\leq r_{\ell-1}} 1 \;=\; r_{\ell-1}$$

it follows that

$$k(d) \;=\; r_{d-1} - r_d \;=\; 2\, rk_{\mathcal{M}}\ K^{d-1}(A,B) - rk_{\mathcal{M}}\ K^{d-2}(A,B) - rk_{\mathcal{M}}\ K^{d}(A,B) \tag{6.4}$$

If $\dim_{\mathcal{M}}\ \hat{K}_{A,B}(\mathbf{V}_{-1}) := 0$ the equations $rk_{\mathcal{M}}\ K^d(A,B) = \dim_{\mathcal{M}}\ \hat{K}_{A,B}(\mathbf{V}_{d-1})$, (6.1), (6.3) and (6.4) yield for $d \in \mathbb{N}$

$$\begin{aligned} k(d) \;&=\; 2\ \dim_{\mathcal{M}}\ \hat{K}_{A,B}(\mathbf{V}_{d-1}) - \dim_{\mathcal{M}} \hat{K}_{A,B}(\mathbf{V}_{d-2}) - \dim_{\mathcal{M}} \hat{K}_{A,B}(\mathbf{V}_{d}) \\ &=\; \dim_{\mathcal{M}}\ \mathbf{M}_d + \dim_{\mathcal{M}}\ \mathbf{M}_{d-2} - 2\ \dim_{\mathcal{M}}\ \mathbf{M}_{d-1} \\ &=\; h(d). \end{aligned}$$

This proves (i).

(ii) It remains to prove that the family $(\lambda_i)_{i\in\underline{m}}$ coincides with the family $(k_i)_{i\in\underline{m}}$. Recall the Rosenbrock deleting procedure described in Section **1.4**. Put

$$w^i_\mu = (w^i_{\mu_1},\ldots,w^i_{\mu_m})^T \in \mathcal{M}^m \qquad \text{for } i\in\underline{m},\ \ \mu = 0,\ldots,k_i$$

so that

$$(DI_n - A)^{k_i}(b_i) = \sum_{\mu=0}^{k_i}(-1)^\mu(DI_n - A)^\mu(B)w^i_\mu$$

where

$$w^i_{\mu_{ij}} = 0 \quad \text{ if } (DI_n - A)^\mu(b_j) \text{ is omitted in the deleting procedure }.$$

Thus w^i_μ are uniquely defined and

$$w_i(D) := \sum_{\mu=0}^{k_i-1} D^\mu w^i_\mu + D^k(w^i_{k_i1},\ldots,w^i_{k_ii-1},-1,0,\ldots 0)^T$$

belongs to $\ker\ \hat{K}_{A,B} = \mu(DI_n - A, B)$ for every $i \in \underline{m}$. Since

$$[W]_\ell = \begin{bmatrix} -1 & & \star \\ 0 & \ddots & \\ & & -1 \end{bmatrix} \quad \text{where } W := [w_1,\ldots,w_m] \in \mathcal{M}[D]^{m\times m}$$

it follow from the the proof of (ii) $\Rightarrow$ (iii) in Theorem 5.2 that $rk_{\mathcal{M}[D]}W = m$. Thus by Theorem 5.2 W is a minimal basis of $W(D)\,\mathcal{M}[D]^m$. If V denotes a minimal basis of $\mu(DI_n - A, B)$

then there exists a non-singular $T = (t_{ij}) \in \mathcal{M}[D]^{m\times m}$ such that $W = V\,T$. By Theorem 5.2 (iii) we have

$$k_j = \deg w_j = \deg \sum_{i=1}^{m} v_i\, t_{ij} = \max_{i\in \underline{m}}\{\lambda_i + \deg\, t_{ij} \mid t_{ij} \neq 0\}$$

and therefore

$$\deg t_{ij} \leq k_j - \lambda_i \qquad \text{if } \lambda_i \leq k_j \tag{6.5}$$

$$t_{ij} = 0 \qquad \text{if } \lambda_i > k_j \tag{6.6}$$

Since T is non singular there exist m distinct numbers $\sigma(1),\ldots,\sigma(m) \in \underline{m}$ so that $t_{i,\sigma(i)} \neq 0$ for $i = 1 \in \underline{m}$. Thus (6.6) implies

$$\lambda_i \leq k_{\sigma(i)} \quad \text{for all } i \in \underline{m} \tag{6.7}$$

By (i) and formula **1**.(4.2) it follows that

$$\sum_{i=1}^{m} \lambda_i = \sum_{i=1}^{m} \alpha_i(t) = rk_{\mathbb{R}}\, K(A(t), B(t)) = \sum_{i=1}^{m} k_i \qquad \text{for all} \quad t \in I \tag{6.8}$$

where I is some suitable interval.
Finally (6.7) and (6.8) imply $(\lambda_i)_{i\in\underline{m}} = (k_i)_{i\in\underline{m}}$. □

Theorem 6.3(ii) is an improvement of Proposition 5.2 in *Ilchmann* (1985a) where controllability is assumed. This direct proof is due to *Glüsing-LüerŒen* (1987). Applying Proposition 1.4.2 to Theorem 6.3 yields

Corollary 6.4 For $(A, B) \in \mathcal{A}^{n\times(n+m)}$ and with the notation as in Theorem 6.3 the following are equivalent:

(i) (A,B) is controllable

(ii) $\sum_{i=1}^{m} k_i = n$

(iii) $\sum_{i=1}^{m} \lambda_i = n$

(iv) $\sum_{i=1}^{m} \alpha_i(t) = n$ for all $t \in I \setminus N, N$ some discrete set; I any non-void interval $I \subset \mathbb{R}$

Proposition 6.5 Let

$$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(r+p)\times(r+m)}$$

be a system matrix. Then

$$\mathbb{P} \text{ is controllable} \quad \Leftrightarrow \quad \sum_{i=1}^{m} \lambda_i = \dim\, \ker_{\mathcal{A}}\, P(D)$$

where $\lambda_1,\ldots,\lambda_m$ denote the minimal indices of $\mu(P, Q)$.

Proof: By Proposition 2.11 $\mathbb{P}$ is system equivalent to some

$$\mathbb{P}_{st} = \begin{bmatrix} DI_n & -B \\ C & E(D) \end{bmatrix} \in \mathcal{A}[D]^{(n+p)\times(n+m)}$$

where $n = \dim\, \ker_{\mathcal{A}}\, P(D)$ and $\mu(P, Q) = \mu(DI_n, B)$ (see Proposition 5.5). Since by Proposition 3.2 $\mathbb{P}$ is controllable if and only if $\mathbb{P}_{st}$ is controllable the result follows from Corollary 6.4. □

2.7 Characterizations of system equivalence

The analysis of the input module in Section 5 and of the input-output map resp. the formal transfer matrix sets us in a position to characterize system equivalence. If $f : \mathbb{P}_1 \to \mathbb{P}_2$ is a system homomorphism then the corresponding formal transfer matrices coincide. For the converse additional assumptions have to be imposed.

Proposition 7.1 Suppose

$$\mathbb{P}_i = \begin{bmatrix} P_i & -Q_i \\ V_i & W_i \end{bmatrix} \in \mathcal{M}[D]^{(r_i+p)\times(r_i+m)}$$

$i = 1,2$ are system matrices with formal transfer matrices $\hat{G}_i$, resp. Then the following statements are valid:

(i) If $\mathbb{P}_1$ and $\mathbb{P}_2$ are controllable then

$$\mathbb{P}_1 \overset{se}{\sim} \mathbb{P}_2 \quad \Leftrightarrow \quad \hat{G}_1 = \hat{G}_2 \text{ and } \mu(P_1,Q_1)) = \mu(P_2,Q_2)$$

(ii) If $\mathbb{P}_1$ and $\mathbb{P}_2$ are controllable and observable then

$$\mathbb{P}_1 \overset{se}{\sim} \mathbb{P}_2 \quad \Leftrightarrow \quad \hat{G}_1 = \hat{G}_2$$

Proof: Controlllability, the formal transfer matrix and the input module are invariant with respect to system equivalence. Thus by Proposition 2.2 without restriction of generality it is assumed that

$$\mathbb{P}_1 = \begin{bmatrix} DI_n & -B \\ C & E(D) \end{bmatrix} \in \mathcal{A}^{(n+p)\times(n+m)}, \quad \mathbb{P}_2 = \begin{bmatrix} DI_{n'} & -B' \\ C' & E'(D) \end{bmatrix} \in \mathcal{A}[D]^{(n'+p)\times(n'+m)}$$

Put $B = [b_1,\dots,b_m]$ and $B' = [b'_1,\dots,b'_m]$.
Necessity in (i) and (ii) follows from Proposition 4.4 and 5.5. I prove sufficiency in (i). Denote the controllability indices of $(0,B),(0,B')$ by k_i, k'_i resp. For arbitrary

$$u(D) = \sum_{\lambda=0}^{k_i-1} D^\lambda u_\lambda \in \mathcal{M}[D]^m \setminus \{0\}, \quad u_\lambda = (0,\dots,0,u_{i,\lambda},0,\dots,0)^T \in \mathcal{M}^m$$

it follows from the construction of the k_i's (see Section 1.4) that $u \notin \ker \hat{K}_{0,B}$. Since by Proposition 5.6 $\ker \hat{K}_{0,B} = \ker \hat{K}_{0,B'}$ one has $u \notin \ker \hat{K}_{0,B'}$ and thus $k_i \le k'_i$. On the other hand $k'_i \le k_i$ and therefore

$$k_i = k'_i \quad \text{for } i \in \underline{m}$$

Since $\mathbb{P}_1$ and $\mathbb{P}_2$ are controllable Corollary 6.4 yields $n = \sum_{i=1}^m k_i = \sum_{i=1}^m k'_i = n'$. Put

$$\begin{aligned} H &= [b_1,\dots,b_1^{(k_1-1)},b_2,\dots,b_m^{(k_m-1)}] \in GL_n(\mathcal{M}) \\ H' &= [b'_1,\dots,b_1'^{(k_1-1)},b'_2,\dots,b_m'^{(k_m-1)}] \in GL_m(\mathcal{M}) \\ T &= H'H^{-1} \in GL_n(\mathcal{M}) \end{aligned}$$

Let $u_i \in \ker \hat{K}_{0,B}$ so that $b_i^{(k_i)} = H\,u_i$. Since $\ker \hat{K}_{0,B} = \ker \hat{K}_{0,B'}$ we obtain $b_i'^{(k_i)} = H'u_i$ whence $T\,b_i^{(k_i)} = H'_i u_i = b'^{(k_i)}$. The last equation together with the definition of T yields

$$T\,b_i^{(j)} = b_i'^{(j)} \quad \text{for } i \in \underline{m}, \quad j = 0,\dots,k_i$$

Therefore $T\,H = (H')^{\cdot} = (T\,H)^{\cdot}\,H + T\,H$ whence $T = 0$.
This gives

$$T\,B = B' \qquad \text{for } T \in GL_n(\mathbb{R}) \tag{7.1}$$

Since $\hat{G}_1 = \hat{G}_2$ we have $C\,D^{-1}B + E(D) = C'D^{-1}T\,B + E'(D)$. By Theorem 3.4 the controllability of $\mathbb{P}_1$ yields the existence of $X \in \mathcal{M}[D]^{n\times n}$, $Y \in \mathcal{M}[D]^{m\times n}$ so that $DX + BY = I_n$. Multiplying the formal transfer matrix from the right by Y gives

$$C\,D^{-1}B\,Y - C'\,D^{-1}T\,B\,Y = (E'(D) - E(D))Y$$

Since $D^{-1}T = T\,D^{-1}$ this is equivalent to

$$(C - C'T)D^{-1}(I_n - D\,X)\cdot D = (E'(D) - E(D))Y\,D$$

resp.

$$(C - C'T)(I_n - X\,D) = (E' - E(D))Y\,D$$

resp.

$$(C - C'T) = ((E'(D) - E(D))Y + (C - C'T)X)D$$

By comparing the coefficients of the last equation one obtains $C = C'T$, $E(D) = E'(D)$. This together with (7.1) implies that $\mathbb{P}_1$ is similar to $\mathbb{P}_2$ and thus the proof of (i) is complete.

Sufficiency in (ii) is proved in several steps. Instead of the formal transfer matrices the input output maps are considered.

(α) Put

$$\mathcal{U}_J^m = \{u \in \mathcal{U}^m \mid \text{supp } u \supset J\} \;\text{ where }\; J = [t_0, t_1], t_0 < t_1.$$

Then $\mathbb{P}_i$ is controllable if and only if the $\mathbb{R}$-linear maps

$$\sigma_J^i : \mathcal{U}_J^m \to \ker_{\mathcal{A}} P_i,\; u(\cdot) \mapsto z^i(\cdot)$$

are surjective, where $z^i(\cdot)$ denotes the unique free motion which satisfies $z_u^i(t) = z^i(t)$ for $t \geq t_1, i = 1, 2$ resp.

(β) It is shown that there exists a unique $\mathbb{R}$-linear map f_J such that the following diagram commutes:

$$\begin{tikzcd}
 & \ker_{\mathcal{A}} P_2 \\
\mathcal{U}_J^m \arrow[ur, "\sigma_J^2"] \arrow[dr, "\sigma_J^1"'] & \\
 & \ker_{\mathcal{A}} P_1 \arrow[uu, "f_J"']
\end{tikzcd}$$

Since σ_J^1 is surjective it remains to prove that $\ker \sigma_J^1 \subset \ker \sigma_J^2$. Suppose $\sigma_J^1(u) = 0$ for $u \in \mathcal{U}_J^m$. Since $(z_u^1, u)\,|_{(t_1,\infty)} = 0$ and

$$V_1(z_u^1) + W_1(u) = G_1(u) = G_2(u) = V_2(z_u^2) + W_2(u)$$

one obtains $V_2(z_u^2)\,|_{(t_1,\infty)} = 0$. So observability of $\mathbb{P}_2$ yields $z_u^2\,|_{(t_1,\infty)} = 0$ and thus $\sigma_J^2(u) = 0$.

(γ) Applying (β) twice yields that f_J is in fact an isomorphism. The $\mathbb{R}$-linear map

$$f : \mathcal{M}(P_1, Q_1) \to \mathcal{M}(P_2, Q_2)$$

defined by

$$
\begin{aligned}
f\,|_{\ker_A P_1\times\{0\}}\,(z^1,0) &= (f_J(z^1),0)\\
f\,|_{\mathcal{M}_+(P_1,Q_1)}\,(z_u^1,u) &= (z_u^2,u)
\end{aligned}
$$

is a system isomorphism if one can show that f satisfies (2.4) and (2.7).
If $(z,u)\in\mathcal{M}(P_1,Q_1)$ is decomposed into

$$(z,u)=(z^1,0)+(z_u^1,u) \tag{7.2}$$

then

$$f(z,u)=(f_J(z^1),0)+(z_u^2,u)$$

To prove (2.4) it remains to show that

$$V_1(z^1)+V_1(z_u^1)+W_1(u) \;=\; V_2(f_J(z^1))+v_2(z_u^2)+W_2(u)$$

or equivalently, since $\sigma_1(u)=\sigma_2(u)$,

$$V_1(z^1) \;=\; V_2(f_J(z^1)) \tag{7.3}$$

Choose $u'\in\mathcal{U}_J^m$ so that $\sigma_J^1(u')=z^1$. Then

$$f_J(z^1)=\sigma_J^2(u') \quad\text{and}\quad (f_J(z^1),u')=(z_{u'}^2,0) \quad \text{for } t\ge t_1.$$

From $\sigma_1(u')=\sigma_2(u')$ it follows that

$$V_1(z^1)=V_1(z_{u'}^1)=V_2(z_{u'}^2)=V_2(f_J(z^1)) \quad \text{for } t\ge t_1$$

and (7.3) is a consequence of the identity property of analytic functions.
It remains to prove (2.7). Suppose $(z,u)\in\mathcal{M}(P_1,Q_1)$ satisfies $(z,u)\,|_J=0$ on some compact interval J. If (z,u) is decomposed as in (7.2) then by (2.4)

$$V_1(z)+W(u) \;=\; V_1(z^1+z_u^1)+W_1(u) \;=\; V_2(f_J(z^1)+z_u^2)+W_2(u)$$

Thus $V_2(f_J(z^1)+z_u^2)\,|_J=0$ and observability of $\mathbb{P}_2$ yields $f(z,u)\,|_J=0$ which proves (2.7).
This completes the proof. □

For time-invariant polynomial matrix systems Proposition 7.1 (i) is proved in *Prätzel-Wolters* (1981) Corollary 3.14. Proposition 7.1 (ii), considered in the time-invariant situation, restates the well-known uniqueness theorem of finite dimensional realization theory, cf. also *Hinrichsen and Prätzel-Wolters* (1980) Proposition 7.8.

Chapter 3

Disturbance Decoupling Problems - A Geometric Approach

3.0 Introduction

The concept of $(A, B)-$ invariance has been introduced by *Basile and Marro* (1969) and *Wonham and Morse* (1970) to solve various decoupling and pole assignment problems for linear time-invariant multivariable systems. This concept was generalized to non-linear systems (see e.g. *Hirschhorn* (1981), *Isidori, Krener, Gori-Giori and Monaco* (1981), *Isidori* (1985)) and to infinite-dimensional linear systems (see e.g. *Curtain* (1985), (1986)). In *Ilchmann* (1985b) I introduced a geometric approach for time-varying systems of the form

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) + S(t)q(t) \\ y(t) &= C(t)x(t) \end{aligned} \tag{0.1}$$

where A, B, C are piecewise analytic matrices as in 1.(1.1) and $S(\cdot) \in \mathcal{A}_p^{n\times s}$. Here $q(\cdot)$ is viewed as a disturbance entering the system via $S(\cdot)$
The main problem is as follows: When is it possible to determine a feedback matrix $F(\cdot) \in \mathcal{A}_p^{m\times n}$ such that in the closed loop system

$$\begin{aligned} \dot{x}(t) &= [A + BF](t)x(t) + S(t)q(t) \\ y(t) &= C(t)x(t) \end{aligned} \tag{0.2}$$

the disturbance $q(\cdot)$ has no influence on the output $y(\cdot)$ on a given open time interval I ?

The following example will illustrate an important difference between time-invariant and time-varying systems with respect to disturbance decoupling.

Let

$$\begin{aligned} \dot{x}(t) &= \begin{bmatrix} 0 & t & 0 \\ a_1(t) & a_2(t) & a_3(t) \\ a_4(t) & a_5(t) & a_6(t) \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ b(t) \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ 0 \\ s(t) \end{bmatrix} q(t) \\ y(t) &= [c(t), 0, 0]x(t) \end{aligned} \tag{0.3}$$

where the entries of the matrices are real analytic functions and $a_3(\cdot), b(\cdot), s(\cdot), c(\cdot)$ are not identically zero.

If the feedback matrix is denoted by

$$F(t) = [f_1(t), f_2(t), f_3(t)]$$

then

$$A + BF = \begin{bmatrix} 0 & t & 0 \\ a_1 + bf_1 & a_2 + bf_2 & a_3 + bf_3 \\ a_4 & a_5 & a_6 \end{bmatrix}.$$

By a simple calculation it is seen that $q(\cdot)$ has no influence on $y(\cdot)$ if and only if $a_3 + bf_3 \equiv 0$. If $b(t') = 0$ and $a_3(t') \neq 0$ for some $t' \in \mathbb{R}$, then $\mid f_3(t) \mid = \mid \frac{a_3(t)}{b(t)} \mid$ tends to infinity as $t \to t'$. This shows that disturbance decoupling might only be possible within certain intervals. These intervals are determined in the following time-varying geometric approach.

It is known from the time-invariant setting that the controllable subspace $im \sum_{i=0}^{n-1} A^i B$ is (A, B)–invariant. As it was shown, see equation 1.(5.4), the time-varying extension of the controllable subspace is the time-varying subspace $\Re(t) = \Phi(t, t_0)\Re(t_0)$. So it is no surprise that one has to extend the concept of (A, B)–invariance to time-varying subspaces instead of constant linear spaces. This basic tool of time-varying subspaces was studied in depth in Section 1.5.

If the entries of the matrices A, B, S and C consist of real analytic functions the present set up is a specialization of the nonlinear approach. However, there are several reasons to introduce a self-contained geometric approach for time-varying systems of the form (0.1):

- The class of *piecewise* real analytic systems is much richer than the class of time-varying systems covered by the non-linear approach.
- The mathematical approach using time-varying subspaces is a natural one for the analysis of time-varying linear disturbance decoupling problems. There is no need to use differential geometry.
- The concept of (A, B)–invariance has a nice geometric interpretation, not given in the nonlinear case (see Theorem 1.5 (iv)). It also can be dualized in a canonical way.
- The maximal intervals where disturbance decoupling is possible are determined by the zeros of certain functions of time.
- A sufficient condition when disturbance decoupling is possible on I is given. This condition can be checked on a computer if, for instance, the matrices in (0.1) are defined over $\mathbb{R}[t]$.
- If disturbance decoupling is possible a constructive algorithm is given to determine the feedback matrix F.

In this chapter we proceed as follows.
In Section 1 the concept of (A, B)–invariant time-varying subspaces is introduced and characterized.
In Section 2 an algebraic characterization of this concept is presented.
The dual relationship between (A, B)– and (C, A)–invariance is explained in Section 3. An algorithm is given which determines in a finite number of steps the smallest (C, A)–invariant family of subspaces containing a family $\mathcal{L}(t)$.
In Section 4 the disturbance decoupling problem for piecewise analytic state space systems is introduced and characterized. For analytic systems it can be checked, by means of the largest

(A, B)-invariant family of subspaces included in ker $C(t)$, if and on which intervals the disturbance decoupling problem is solvable.
In Section 5 controllability subspace families are defined and characterized.
This is used to solve the restricted decoupling problem for analytic systems in Section 6.

Sections 1 to 4 are based on *Ilchmann* (1985b), the results of Section 5 and 6 are presented in *Ilchmann* (1986).

3.1 (A, B)-invariant time-varying subspaces

Throughout this chapter piecewise analytic systems $(A, B) \in \mathcal{A}_p^{n \times (n+m)}$ are considered. The concept and notation of time-varying subspaces (see Section 1.5) will be used.

Definition 1.1 Suppose $(A, B) \in \mathcal{A}_p^{n \times (n+m)}$ and $\mathcal{V} \in \mathbf{W}_n$ is generated by $V \in \mathcal{A}_p^{n \times k}$. Then $\mathcal{V}$ is called *meromorphically* (A, B)-*invariant* if there exist $N \in \mathcal{M}_p^{k \times k}$, $M \in \mathcal{M}_p^{n \times k}$ such that

$$(DI_n - A)(V) = V\,N + B\,M \tag{1.1}$$

$\mathcal{V}$ is called (A, B)-*invariant* if (1.1) holds true for some N, M with entries in $\mathcal{A}_p$ instead of $\mathcal{M}_p$. If $B = 0$ we speak of (meromorphic) A-invariance.

This is an extension of the concept of (A, B)-invariance introduced by *Basile and Marro* (1969) for time-invariant systems $(A, B) \in \mathbb{R}^{n \times (n+m)}$, see also *Wonham* (1974). In this case a constant vector space $\mathcal{V}$ of $\mathbb{R}^n$ is called (A, B)-*invariant* if

$$A\mathcal{V} \subset \mathcal{V} + imB$$

Clearly, $\mathcal{V}$ viewed as a constant family belongs to $\mathbf{W}_n$ and $\mathcal{V}$ is (A, B)-invariant in the sense of Definition 1.1.

A simple example shall illustrate the difference between (A, B)- and meromorphic (A, B)-invariance. Put

$$\mathcal{V}(t) = v(t) \cdot \mathbb{R}, \quad v(t) = \begin{bmatrix} 0 \\ t \end{bmatrix}, \quad A = 0_{2\times 2}, \quad B(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{1.2}$$

Then

$$(DI_2 - A)(v(t)) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ t \end{bmatrix} t^{-1} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} 0$$

and thus $\mathcal{V}$ is meromorphically (A, B)-invariant.

Proposition 1.2 Suppose

(i) $(A, B) \in \mathcal{A}^{n \times (n+m)}$, $V \in \mathcal{A}^{n \times k}$ and $rk_{\mathbb{R}}[V(t), B(t)] = \text{const.}$ for all $t \in \mathbb{R}$

or

(ii) $(A, B) \in \mathcal{A}_p^{n \times (n+m)}$, $V \in \mathcal{A}_p^{n \times k}$ and $[V, B]$ has p.c. rank.

Then $\mathcal{V}$ generated by V is meromorphically (A, B)-invariant if and only if $\mathcal{V}$ is (A, B)-invariant.

Proof: (i) It has to be shown that meromorphic (A, B)–invariance implies (A, B)–invariance. Meromorphic (A, B)-invariance yields that

$$(DI_n - A(t))(V(t)) \in [V(t), B(t)] \ \mathbb{R}^{m+k} \quad \text{for almost all } t \in \mathbb{R}$$

Now the result follows from Proposition 1.5.10(i).

(ii) Since $[V, B]$ has p.c. rank it is sufficient to prove the assertion on an interval (a_v^ℓ, a_{v+1}^r) where $[V, B]_v(t)$ has constant rank. (The notation of Definition 1.5.7 is used.) Now (ii) follows from (i). □

Remark 1.3 Suppose (i) or (ii) of Proposition 1.2 is satisfied. Then it follows from the proof of Proposition 1.2 that $\mathcal{V}$ is (A, B)–invariant if and only if

$$im\,(DI_n - A(t))(V(t)) \subset \mathcal{V}(t) + im\ B(t) \qquad \text{for all } t \in N, \text{ where } N \text{ is a discrete set}$$

The following basic properties of (meromorphic) (A, B)–invariance are immediate.

Remark 1.4 For $\mathcal{V} \in \mathbf{W}_n$ with generator $V = [v_1, \dots, v_k] \in \mathcal{A}_p^{n\times k}$ the following statements hold true:

(i) Suppose $(A, B) \in \mathcal{A}_p^{n\times(n+m)}$ is similar to $(A', B') \in \mathcal{A}_p^{n\times(n+m)}$ via $T \in GL_n(\mathcal{A}_p)$. Then $\mathcal{V}$ is (meromorphically) (A, B)–invariant iff $T^{-1}\cdot\mathcal{V}$ is (meromorphically) (A', B')–invariant.

(ii) $\mathcal{V}$ is (meromorphically) (A, B)–invariant iff for every

$$v = \sum_{i=1}^{k} a_i v_i \quad , \alpha_i \in \mathcal{A}_p \ (\alpha_i \in \mathcal{M}_p)$$

there exist $r \in \mathcal{A}_p^k, s \in \mathcal{A}_p^m$ $(r \in \mathcal{M}_p^k, s \in \mathcal{M}_p^m)$ such that

$$(DI_n - A)(v) = V\,r + B\,s.$$

(iii) The sum of two (meromorphically) (A, B)–invariant families is (meromorphically) (A, B)–invariant as well.

The concept of (A, B)–invariance becomes clearer by the following theorem. Furthermore this result is important for the solvability of the disturbance decoupling problem tackled in Section 4.

Theorem 1.5 Suppose $(A, B) \in \mathcal{A}_p^{n\times(n+m)}$ and $\mathcal{V} \in \mathbf{W}_n$ is generated by $V \in \mathcal{A}_p^{n\times k}$ with $rk_{\mathbb{R}} V(t) = k$ for all $t \in \mathbb{R}$. $P(t) : \mathbb{R}^n \to \mathcal{V}(t)$ denotes the orthogonal projector on $\mathcal{V}(t)$ along $\mathcal{V}^\perp(t)$. Then the following are equivalent:

(i) $\mathcal{V}$ is (A, B)–invariant, i.e. there exist $N \in \mathcal{A}_p^{k\times k}$ and $M \in \mathcal{A}_p^{m\times k}$ such that

$$(DI_n - A(t)(V(t)) \;=\; V(t)N(t) + B(t)M(t) \qquad \text{for all } t \in \mathbb{R}$$

(ii) There exists an $F \in \mathcal{A}_p^{m\times n}$ such that $\mathcal{V}$ is $(A + B\,F)$–invariant.

(iii) There exist $\tilde{N} \in \mathcal{A}_p^{n\times n}$ and $\tilde{M} \in \mathcal{A}_p^{m\times n}$ such that

$$(DI_n - A(t))(P(t)) \;=\; P(t)\tilde{N}(t) + B(t)\tilde{M}(t) \qquad \text{for all } t \in \mathbb{R}$$

(iv) There exist $N \in \mathcal{A}_p^{k\times k}, M \in \mathcal{A}_p^{m\times k}$ such that

$$V(t)\Psi(t_0,t)^T \;=\; \Phi(t,t_0) + \int_{t_0}^{t} \Phi(t,s)B(s)M(s)\Psi(t_0,s)^T ds \qquad \text{for all } t \in \mathbb{R}$$

where Φ, Ψ denote the transition matrices of

$$\dot{x}(t) = A(t)x(t), \quad \dot{x}(t) = N(t)^T x(t)\text{ , resp.}$$

Proof: (i) $\Rightarrow$ (ii) : Define

$$F = M(V^T V)^{-1} V^T$$

Then

$$(DI_n - (A + B\,F))(V) = (DI_n - A)(V) - B\,F\,V = V\,N$$

which proves (ii).

(ii) $\Rightarrow$ (i) is trivial.

(i) $\Rightarrow$ (iii) : Put $Q = V^T(V\,V^T)^{-1}P$. Then $V\,Q = P$ and $(DI_n - A)(P) = V[N\,Q - Q] + B\,M\,Q$.

(iii) $\Rightarrow$ (i): If $Q := V^T(V\,V^T)^{-1}P$ then

$$(DI_n - A)(V)Q = (DI_n - A)(P) + V\,\dot{Q} = P\,\tilde{N} + B\,\tilde{M} + V\,\dot{Q}.$$

Since $rk_{\mathbf{R}}P(t) = k$ for all $t \in \mathbb{R}$ there exists $Q_r \in \mathcal{A}_p^{n\times k}$ so that $Q\,Q_r = I_k$. Thus

$$(DI_n - A)(V) = P\,\tilde{N}\,Q_r + B\,\tilde{M}\,Q_r + P\,Q_r\,\dot{Q}\,Q_r$$

which proves (i).

(i) $\Rightarrow$ (iv) : Multiplying the equation in (i) from the left by $T^{-1}(\cdot) = \Phi(\cdot, t_0)^{-1}$ yields

$$\dot{V}' = V'\,N + B'\,M \quad \text{where } V' = T^{-1}\,V,\;\; B' = T^{-1}\,B$$

which by Variation of Constants is equivalent to

$$V'^T(t) = \Psi(t,t_0)V'^T(t_0) + \int_{t_0}^{t} \Psi(t,s)M^T(s)B'^T(s)ds$$

resp.

$$V'(t)\Psi(t_0,t)^T = V'(t_0) + \int_{t_0}^{t} B'(s)M(s)\Psi(t_0,s)^T ds$$

Multiplying from the left by $T(t)$ gives

$$V(t)\Psi(t_0,t)^T = \Phi(t,t_0)V(t_0) + \int_{t_0}^{t} \Phi(t,s)B(s)M(s)\Psi(t_0,t)^T ds.$$

This proves (iv).
To prove (iv) $\Rightarrow$ (i) reverse the arguments in the proof of (i) $\Rightarrow$ (iv). □

As an immediate consequence of Theorem 1.5 one obtains

Corollary 1.6 Suppose $A \in \mathcal{A}_p^{n\times n}$ and $\mathcal{V} \in \mathbf{W}_n$ is generated by $V \in \mathcal{A}^{n\times k}$ with $rk_{\mathbb{R}}\, V(t) = k$ for all $t \in \mathbb{R}$. Then the following are equivalent:

(i) $\mathcal{V}$ is A–invariant.

(ii) There exists $\tilde{N} \in \mathcal{A}_p^{n\times n}$ such that

$$(DI_n - A(t))(P(t)) = P(t)\tilde{N}(t) \qquad \text{for all } t \in \mathbb{R}.$$

(iii) $\mathcal{V}(t) = \Phi(t,t_0)\mathcal{V}(t_0) \qquad$ for all $t, t_0 \in \mathbb{R}$.

Remark 1.7 If a real analytic system $(A,B) \in \mathcal{A}^{n\times(n+m)}$ is considered and V is real analytic, then in Theorem 1.5 and Corollary 1.6 all matrices are also real analytic. The proofs carry over completely.

Remark 1.8 Condition (iii) in Corollary 1.6 implies that for every $x_0 \in \mathcal{V}(t_0)$ the free trajectory $\Phi(t,t_0)x_0$ remains in $\mathcal{V}(t)$ for all $t \in \mathbb{R}$.
Condition (iv) in Theorem 1.5 says that if $x_0 \in \mathcal{V}(t_0)$ then there exists a control $u \in \mathcal{A}_p^m$ such that the forced motion

$$x_u(t;t_0,x_0) = \Phi(t,t_0)x_0 + \int_{t_0}^{t} \Phi(t,s)B(s)u(s)ds$$

can be hold in $\mathcal{V}(t)$ for every $t \in \mathbb{R}$. For time-invariant systems the latter condition is also sufficient for (A,B)–invariance. If this is also valid for time-varying systems is an open problem.

Example 1.9 For time-invariant systems $(A,B) \in \mathbb{R}^{n\times(n+m)}$ it is well-known that the controllable subspace $\sum_{i=0}^{n-1} A^i\, imB$ is the smallest A–invariant subspace which contains imB , see e.g. *Wonham* (1985) Section 1.2. This is extended to the analytic situation as follows: The controllable family $\Re$ (see Section 1.5) of an analytic system $(A,B) \in \mathcal{A}^{n\times(n+m)}$ is the smallest A–invariant family which contains $imB(\cdot)$. In fact, $\Re$ is an analytic family (see Remark 1.5.11) and thus A–invariance follows from Corollary 1.6. Use of the presentation 1.(5.5) of $\Re(t)$ yields $imB(t) \subset \Re(t)$. If $\mathcal{V} = (\mathcal{V}(t))_{t\in\mathbb{R}} \in \mathbf{W}_n$ is another A–invariant family with $imB(t) \subset \mathcal{V}(t)$ then $(DI_n - A(t))^i(B(t)) \subset \mathcal{V}(t)$ for all $i \in \mathbb{N}$. Thus $\Re(t) \subset \mathcal{V}(t)$ by 1(5.5) and therefore $\Re$ is the smallest A–invariant family which contains $imB(\cdot)$.

If we do not assume that the rank of $V(\cdot)$ is constant then the feedback constructed in Theorem 1.5 (ii) may have poles. For disturbance decoupling problems it is important to locate these poles.

Proposition 1.10 Suppose $(A,B) \in \mathcal{A}^{n\times(n+m)}$, $rk_{\mathcal{M}}B = m$ and $\mathcal{V} \in \mathbf{W}_n$ is generated by $V \in \mathcal{A}^{n\times k}$ with $rk_{\mathcal{M}}V = k$.
If $\mathcal{V}$ is meromorphically (A,B)–invariant then there exist analytic matrices U_1, U_2, T, W of formats $k\times s, m\times s, s\times s, s\times k$, resp. and $T^{-1} \in \mathcal{M}^{s\times s}$ so that

$$(DI_n - A)(V) = V\, U_1\, T^{-1}\, W + B\, U_2\, T^{-1} W \tag{1.3}$$

T^{-1} has poles at t' if and only if

$$\dim(V(t')\mathbb{R}^k + B(t')\mathbb{R}^m) \; < \; \max_{t\in\mathbb{R}} \dim(V(t)\mathbb{R}^k + B(t)\mathbb{R}^m) \tag{1.4}$$

Proof: By Lemma 1.5.1 $G = \text{gcld}\,(V, B) \in \mathcal{A}^{n\times s}$ satisfies

$$G\mathcal{A}^s = V\mathcal{A}^k + B\mathcal{A}^m \quad \text{and} \quad G = V\,U_1 + B\,U_3 \quad \text{for some} \quad U_1 \in \mathcal{A}^{k\times s},\ U_3 \in \mathcal{A}^{m\times s}.$$

Let $\hat{G} \in \mathcal{A}^{n\times s}$ with $rk_{\mathbb{R}}\,\hat{G}(t) = s$ for all $t \in \mathbb{R}$ (see Lemma 1. 5.4) such that

$$G(t)\mathbb{R}^s \overset{\subseteq}{\underset{\text{ae}}{=}} \hat{G}(t)\mathbb{R}^s$$

Then $G = \hat{G}\,T$ for some $T \in \mathcal{A}^{s\times s}$ with $T^{-1} \in \mathcal{M}^{s\times s}$. Since $\mathcal{V}$ is meromorphically (A, B)–invariant and $\hat{G}$ is left invertible over $\mathcal{A}$ there exists some $W \in \mathcal{A}^{s\times k}$ so that $(DI_n - A)(V) = \hat{G}W$. This proves (1.3). Clearly, $T(\cdot)^{-1}$ has poles at t' if and only if $rk_{\mathbb{R}}\,G(t') < rk_{\mathcal{M}}\,G(\cdot)$ which proves (1.4). □

Example 1.11 Let $(A, B) \in \mathcal{A}^{3\times(3+1)}$ be given by

$$A(t) = \begin{bmatrix} t_2 & \sin t & -t(t^2+2) \\ a_4(t) & a_5(t) & a_6(t) \\ -1 & 0 & t \end{bmatrix}, \quad B(t) = \begin{bmatrix} -t^2(t-2) \\ t^2 \\ t-2 \end{bmatrix}$$

and $\mathcal{V}(t) = V(t)\mathbb{R}$ where $V(t) = [t^2, 0, -1]^T$.
Using the notation of the proof of Proposition 1.10 one obtains by Lemma 1.5.1

$$\begin{aligned} G(t) &= \text{gcld}\,(V, B) = \begin{bmatrix} t^2 & -t^2(t-2) \\ 0 & t^2 \\ -1 & t-2 \end{bmatrix} \\ &= V\,U_1 + B\,U_3 \\ &= \begin{bmatrix} t^2 \\ 0 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -t^2(t-2) \\ t^2 \\ t-2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

and

$$G(t) = \hat{G}(t)T(t) = \begin{bmatrix} t^2 & 0 \\ 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -(t-2) \\ 0 & t^2 \end{bmatrix}.$$

$\mathcal{V}$ is meromorphically (A, B)–invariant since for

$$W(t) = \begin{bmatrix} -t(t+1) \\ -a_4(t)t^2 + a_6(t) \end{bmatrix} \quad \text{and} \quad T^{-1}(t) = \begin{bmatrix} 1 & -t^{-2}(t-2) \\ 0 & t^{-2} \end{bmatrix}$$

$$(DI_n - A(t))(V(t)) = \begin{bmatrix} -t^3(t+1) \\ -a_4(t)t^2 + a_6(t) \\ t(t+1) \end{bmatrix} = \hat{G}(t)W(t) = G(t)T(t)^{-1}W(t)$$

$$\begin{aligned} &= V(t)[t^{-2}(t-2)[-a_4(t)t^2 + a_6(t)] - t(t+1)] + B(t)[t^{-2}(-a_4(t)t^2 + a_6(t))] \\ &= V(t)N(t) + B(t)M(t) \end{aligned}$$

Since

$$rk_{\mathbb{R}}[V(t), B(t)] = \begin{cases} 1 & \text{for } t = 0 \\ 2 & \text{for } t \neq 0 \end{cases}$$

$t' = 0$ is the only pole of $T^{-1}(t')$ or equivalently

$$\dim\,(V(t')\mathbb{R} + B(t')\mathbb{R}) < 2.$$

In the proof of Theorem 1.5 an analytic feedback matrix F was determined so that $(DI_n - (A+BF))(V) \subset \mathcal{V}$. In the present example $V(\cdot)$ does not have constant rank. Therefore F is a meromorphic matrix

$$F(t) = \frac{-a_4(t)t^2 + a_6(t)}{t^2(t^4+1)}[t^2, 0, -1] = M(V^T V)^{-1} V^T$$

and

$$B(t)F(t) = \frac{-a_4(t)t^2 + a_6(t)}{t^4+1} \begin{bmatrix} -t^2(t-2) & 0 & (t-2) \\ t^2 & 0 & -1 \\ (t-2) & 0 & -\frac{t-2}{t^2} \end{bmatrix}$$

It depends on the zeros of a_4 and a_6 if F has poles.

3.2 Algebraic characterization of (A, B)–invariance

Based on *P. Fuhrmann's* realization theory (cf. *Fuhrmann*(1976)), *Emre and Hautus* (1980) and *Hautus* (1980) give a "frequency domain characterization" of (A, B)–invariant subspaces. In *Hautus* (1980) the following result can be found.

Proposition 2.1 Suppose $(A, B) \in \mathbb{R}^{n\times(n+m)}$ and $\mathcal{V}$ is a subspace of $\mathbb{R}^n$. Then $\mathcal{V}$ is (A, B)–invariant if and only if for every $v \in \mathcal{V}$ there exist strictly proper rational functions $\omega(s)$, $\xi(s)$ with $\xi(s) \in \mathcal{V}$ for all $s \in \mathbb{R}$ such that

$$v = (sI_n - A)\xi(s) - B\,\omega(s) \tag{2.1}$$

For time-varying systems a frequency domain analysis does not exist, however the skew polynomial approach developed in Chapter 2 can be used to extend the previous proposition. At first the case $B(\cdot) \equiv 0$ is studied.

Proposition 2.2 Suppose $A \in \mathcal{A}^{n\times n}$, $V \in \mathcal{A}^{n\times k}$ and $rk_{\mathbb{R}}\, V(t) =$ const. for all $t \in \mathbb{R}$.
Then $\mathcal{V}$ generated by V is A–invariant if and only if for every $v \in \mathcal{A}^n$ with $v(t) \in \mathcal{V}(t)$ for all $t \in \mathbb{R}$ there exist

$$\xi(D) = \sum_{i=0}^{\ell-1} \xi_i D^i \in \mathcal{A}^n[D], \qquad \xi_i(t) \in \mathcal{V}(t) \text{ for all } t \in \mathbb{R}$$

and

$$p(D) \in \mathcal{A}[D] \;\text{ with } \deg p > \deg \xi$$

such that

$$v \cdot p(D) = (DI_n - A) \cdot \xi(D) \tag{2.2}$$

Using Remark 1.1.1 and fomula 1.(1.6) it is easily verified that without restriction of generality one may assume $A = 0$.
Suppose $\mathcal{V}$ is A–invariant and let $v \in \mathcal{A}^n$ such that $v(t) \in \mathcal{V}(t)$ for all $t \in \mathbb{R}$. Use of the multiplication in the left-skew field $\mathcal{M}(D)$ (see Section 2.2) yields

$$D^{-1}v_i = v_i(D + \frac{\dot{v}_i}{v_i})^{-1} \quad \text{for } v_i \in \mathcal{A}^*.$$

So there exists $p \in \mathcal{A}[D]$ such that

$$D^{-1}\, v\, p(D) = \hat{\xi}(D)\, p(D) \quad \in \mathcal{A}^n[D]$$

Put

$$\sum_{i=0}^{\ell} p_i\, D^i = p(D) \quad \text{and} \quad \sum_{i=0}^{\ell-1} \xi_i\, D^i = \xi(D) := \hat{\xi}(D)\, p(D).$$

Then we get by comparing the coefficients of $v\, p(D) = D\, \xi(D)$ and using A–invariance of $\mathcal{V}$ that $\xi_i(t) \in \mathcal{V}(t)$ for all $t \in \mathbb{R}$ and $i = 0, \ldots, \ell - 1$.
To prove the converse let $p(D) = \sum_{i=0}^{\ell} D^i\, \tilde{p}_i$ and compare the coefficients of $v\, p(D) = D\, \xi(D)$. Then $v\, \tilde{p}_{\ell-1} - \dot{v}\, \tilde{p}_\ell = \xi_{\ell-2} + \dot{\xi}_{\ell-1}$. Since $v\, p_{\ell-1} = \xi_{\ell-1} + \dot{\xi}_{\ell-1}$ one obtains $(\dot{v}\, \tilde{p}_\ell)(t) \in \mathcal{V}(t)$ for all $t \in \mathbb{R}$. Thus by Proposition 1.5.10(i) $\dot{v}(t) \in \mathcal{V}(t)$ for all $t \in \mathbb{R}$ and the proof is complete. □

Proposition 2.3 Suppose $(A, B) \in \mathcal{A}^{n\times(n+m)}$, $V \in \mathcal{A}^{n\times k}$ and $rk_{\mathbb{R}}[V(t), B(t)] =$ const. for all $t \in \mathbb{R}$. Then $\mathcal{V} \in \mathbf{W}_n$ generated by V is (A, B)–invariant if and only if for every column vector v of V there exist

$$\begin{aligned}
\xi(D) &= \sum_{i=0}^{\ell-1} \xi_i\, D^i \in \mathcal{A}^n[D], \quad \xi_i(t) \in \mathcal{V}(t) \text{ for all } t \in \mathbb{R} \\
\omega(D) &= \sum_{i=0}^{\ell-1} \omega_i\, D^i \in \mathcal{A}^m[D] \\
p(D) &= \sum_{i=0}^{\ell} p_i\, D^i \in \mathcal{A}[D]
\end{aligned}$$

such that

$$v \cdot p(D) = (DI_n - A) \cdot \xi(D) - B \cdot \omega(D) \tag{2.3}$$

Proof: As in the proof of Proposition 2.2 assume $A = 0$. If $\mathcal{V}$ is (A, B)–invariant and v is a column vector of V put

$$\begin{aligned}
\hat{\xi}(D) &= (DI_n - B\, F)^{-1} v \\
\hat{\omega}(D) &= F\, \hat{\xi}(D)
\end{aligned}$$

where F is given as in Theorem 1.5 (ii).
Let $p \in \mathcal{A}[D]$ so that $\xi(D) := \hat{\xi}(D)\, p(D) \in \mathcal{A}^n[D]$ and one obtains

$$v\, p(D) = (DI_n - B\, F)\xi(D).$$

Thus for $\omega(D) := F\, \xi(D)$ (2.3) is satisfied. The multiplication rule w.r.t. $\mathcal{A}[D]$ yields that necessarily the degree conditions are valid. It remains to show that $\xi_i(t) \in \mathcal{V}(t)$ for all $t \in \mathbb{R}$, $i = 0, \ldots, \ell - 1$. Choose $T \in GL_n(\mathcal{A})$ (cf. Remark 1.1.1 and formula 1.(1.6)) so that

$$T\, v \cdot p(D) = D \cdot T\, \xi(D) \tag{2.4}$$

Since $\mathcal{V}$ is $B\, F$–invariant it follows from Remark 1.4(i) that $T\, \mathcal{V}$ is $0_{n\times n}$–invariant. Now order the coefficients of $D \cdot T\xi(D)$ to the left and compare successively the coefficients in (2.4). Then one obtains

$$T\, \xi_i(t) \in T\, \mathcal{V}(t) \text{ for all } \quad t \in \mathbb{R}, \quad i = 0, \ldots, \ell - 1.$$

To prove (A, B)-invariance use the multiplication rule to order the coefficients in (2.3) as follows

$$\sum_{i=0}^{\ell} v\, p_i\, D^i \;=\; \dot{\xi}_0 + \sum_{i=0}^{\ell-2}(\xi_{i-1} + \dot{\xi}_i)D^i + \xi_{\ell-1}\, D^{\ell} - \sum_{i=0}^{\ell-1} B\,\omega_i\, D^i$$

By comparing the coefficients at $D^{\ell-1}$ one gets

$$v\, P_{\ell-1} = \xi_{\ell-2} + \dot{\xi}_{\ell-1} - B\,\omega_{\ell-1} \in \mathcal{V} \tag{2.5}$$

and ordering the coefficients in (2.3) to the left gives (see formular 2.(5.2))

$$\begin{aligned} v \sum_{i=0}^{\ell} D^i\, P_i' = D[\xi_o \; &+ \; (D\,\xi_1 - \dot{\xi}_1) + \ldots + \sum_{\lambda=0}^{\ell-1} (-1)^{\lambda} \binom{\ell-1}{\lambda} D^{\ell-1-\lambda}\, \xi_{\ell-1}^{(\lambda)}] \\ &- \; B\,\omega_0 - \ldots - \sum_{\lambda=0}^{\ell-1}(-1)^{\lambda} \binom{\ell-1}{\lambda} D^{\ell-1-\lambda}(B\,\omega_{\ell-1})^{(\lambda)} \end{aligned}$$

Again, by comparing the coefficients at $D^{\ell-1}$ one gets

$$v\, p'_{\ell-1} - 2\,\dot{v}\, p'_{\ell} = \xi_{\ell-2} - \dot{\xi}_{\ell-1} - B\,\omega_{\ell-1}. \tag{2.6}$$

(2.5) and (2.6) imply

$$\dot{v}(t)\cdot 2\, p'_{\ell}(t) \in \mathcal{V} + im\ B(t) \qquad \text{for all } t \in \mathbb{R}$$

Thus

$$\dot{v}(t) \in [V(t), B(t)]\, \mathbb{R}^{k\times m} \quad \text{for all } t \in \mathbb{R} \setminus N$$

where N is a discrete set.

Since $rk_{\mathbb{R}}[V(t), B(t)]$ is constant in t Proposition 1.5.10(i) yields $\dot{v}(t) \in [V(t), B(t)]\ \mathbb{R}^{k\times m}$ for all $t \in \mathbb{R}$ whence $\mathcal{V}$ is $(0, B)$-invariant. □

If in Proposition 2.3 all matrices are defined over $\mathcal{A}_p$ and $[V, B]$ has p.c. rank then a polynomial characterization (2.3) is also valid with $\xi(D), \omega(D), p(D)$ defined over $\mathcal{A}_p[D]$. We omit this.

3.3 Duality between (A, B)- and (C, A)-invariance

For time-invariant systems $(A, B, C) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}$ a constant subspace $\mathcal{V}$ of $\mathbb{R}^n$ is called *(C, A)-invariant* if $A(\mathcal{V} \cap \ker C) \subset \mathcal{V}$. It is well-known (see e.g. *Schumacher* (1979)) that $\mathcal{V}$ is (A, B)-invariant if and only if $\mathcal{V}^{\perp}$ is $(B^T, -A^T)$-invariant.
For time-varying systems $(A, B, C) \in \mathcal{A}_p^{n\times n} \times \mathcal{A}_p^{n\times m} \times \mathcal{A}_p^{p\times n}$ it has already been metioned in Section 1.5 that, in general, the time-varying subspace $\ker C(\cdot)$ does not have a generator $W \in \mathcal{A}_p^{n\times k'}$. Even if $\mathcal{V} \in \mathbf{W}_n$ has a generator $V \in \mathcal{A}_p^{n\times k}$ then the orthogonal complement $\mathcal{V}^{\perp} = (\mathcal{V}(t)^{\perp})_{t\in\mathbb{R}}$ does, in general, not have a piecewise analytic generator. Therefore equivalence classes were introduced

$$\overline{\mathcal{V}} := \{W \in \mathbf{W}_n \mid W(t) \overset{ae}{=} \mathcal{V}(t)\} \quad \text{for } \mathcal{V} = (\mathcal{V}(t))_{t\in\mathbb{R}} \in \mathbf{W}_n$$

By Lemma 1.5.9 for every $\mathcal{V}$ generated by $V \in \mathcal{A}_p^{n\times k}$ and $B \in \mathcal{A}_p^{n\times m}$ one can find $W \in \mathcal{A}_p^{n\times(n-k)}$ and $W' \in \mathcal{A}_p^{n\times k'}$ with p.c. ranks so that

$$(W(t)\mathbb{R}^{n-k})_{t\in\mathbb{R}} \in \overline{\mathcal{V}^{\perp}}$$

and

$$(W'(t)\mathbb{R}^{k'})_{t\in\mathbb{R}} \in \overline{(\mathcal{V}(t)\cap\ker B^T(t))_{t\in\mathbb{R}}}$$

If

$$\overline{\mathbf{W}_n} := \{\overline{\mathcal{V}} \mid \mathcal{V}\in\mathbf{W}_n\}$$

the concept of (meromorphic) (A,B)–invariance is extended as follows

Definition 3.1 Suppose $(A,B)\in\mathcal{A}_p^{n\times(n+m)}$. Then $\overline{\mathcal{V}}\in\overline{\mathbf{W}_n}$ is called (meromorphically) (A,B)–*invariant* if there exists a $\mathcal{V}\in\overline{\mathcal{V}}$ so that $\mathcal{V}$ is (meromorphically) (A,B)–invariant.

Now (C,A)–invariance as defined above for constant systems is extended to the time-varying situation as follows.

Definition 3.2 Suppose $A\in\mathcal{A}_p^{n\times n}$, $C\in\mathcal{A}_p^{p\times n}$ and $\mathcal{V}\in\mathbf{W}_n$ is generated by $V\in\mathcal{A}_p^{n\times k}$. Choose $\hat{W}\in\mathcal{A}_p^{n\times s}$ so that

$$\hat{W}(t)\mathbb{R}^s \overset{ae}{=} \mathcal{V}(t)\cap\ker C(t)$$

Then $\overline{\mathcal{V}}$ is called (C,A)–*invariant* if

$$((DI_n - A(t))(\hat{W}(t))\cdot\mathbb{R}^s)_{t\in\mathbb{R}} \in \overline{\mathcal{V}} \tag{3.1}$$

Remark 3.3

(i) In Definition 3.2 one has some freedom in choosing $\hat{W}$. By Proposition 1.5.9 $\hat{W}$ may be chosen with p.c. rank. Also by Proposition 1.5.9 choose $\hat{V}\in\mathcal{A}_p^{n\times k}$ with p.c. rank such that $(\hat{V}(t)\mathbb{R}^k)_{t\in\mathbb{R}}\in\overline{\mathcal{V}}$. Now it follows from Proposition 1.5.10 (ii) that $\overline{\mathcal{V}}$ is (C,A)–invariant if and only if

$$(DI_n - A)(\hat{W}) = \hat{V}R \quad \text{for some } R\in\mathcal{A}_p^{k\times s} \tag{3.2}$$

(ii) Since there always exsits $\hat{V}\in\mathcal{A}_p^{n\times k}$ with p.c. rank so that $(\hat{V}(t)\mathbb{R}^k)_{t\in\mathbb{R}}\in\overline{\mathcal{V}}$, it makes no sense to introduce *meromorphic* (C,A)–invariance similar to meromorphic (A,B)–invariance.

(iii) It is easily verified that analogous statements as in Remark 1.4 hold true for (C,A)–invariance.

Proposition 3.4 Suppose $A\in\mathcal{A}_p^{n\times n}$, $C\in\mathcal{A}_p^{p\times n}$ and $\mathcal{V}\in\mathbf{W}_n$ is generated by $V\in\mathcal{A}_p^{n\times k}$. Then $\overline{\mathcal{V}}$ is (C,A)–invariant if and only if $\overline{\mathcal{V}^\perp}$ is meromorphically $(-A^T,C^T)$–invariant.

Proof: By Remark 3.3 (i) assume without restriction of generality that V has p.c. rank. Choose by Lemma 1.5.9 (ii) $\hat{U}\in\mathcal{A}_p^{n\times n}$ with p.c. rank so that

$$\mathcal{V}^\perp(t) = \hat{U}(t)\mathbb{R}^\ell$$

Since $\hat{W}\in\mathcal{A}_p^{n\times s}$ satisfies

$$\hat{W}(t)\mathbb{R}^s \overset{\subseteq}{\underset{ae}{=}} V(t)\mathbb{R}^k\cap\ker C(t)$$

one obtains for arbitrary columns $\hat{u}$ of $\hat{U}$ and $\hat{w}$ of $\hat{W}$, $<\hat{u}(t),\hat{w}(t)>=0$ for all $t\in\mathbb{R}$. Thus

$$[(DI_n - A)(\hat{w})]^T\,\hat{u} = -\hat{w}^T\dot{\hat{u}} - \hat{w}^T A^T\,\hat{u} = -\hat{w}^T[(DI_n + A^T)(\hat{u})]. \tag{3.3}$$

If $\overline{\mathcal{V}}$ is (C,A)–invariant then (3.2) yields $(DI_n - A)(\hat{w}) = V\, r$ for some $r \in \mathcal{A}_p^k$ and by (3.3) $< (DI_n - A^T)(\hat{u}), \hat{w} >= 0$. Therefore

$$(DI_n + A(t)^T)(\hat{u}(t)) \in (\mathcal{V}(t) \cap \ker C(t))^\perp + im\,(C(t)^T) = [\hat{U}(t), C(t)^T]\,\mathbb{R}^{\ell+p} \quad \text{for all } t \in \mathbb{R} \tag{3.4}$$

Consider an interval $[a_\nu, a_{\nu+1})$ so that $[\hat{U}, C^T]\,|_{[a_\nu,a_{\nu+1})}$ is real analytic and can be real analytically extended to

$$[\hat{U}, C^T]_\nu \text{ on some } I_\nu := (a_\nu^\ell, a_{\nu+1}^r), \quad a_\nu^\ell < a_\nu,\ a_{\nu+1} < a_{\nu+1}^r.$$

Then by Lemma 1.5.4 (i) there exists $\hat{G}_\nu \in \mathcal{A}\,|_{I_\nu}^{n\times(\ell+p)}$ with constant rank so that

$$[\hat{U}, C^T]_\nu(t)\mathbb{R}^{\ell+p} \overset{\subset}{\underset{=}{\scriptstyle ae}} \hat{G}_\nu(t)\mathbb{R}^{\ell+p} \qquad \text{on } I_\nu$$

Since $rk_{\mathcal{M}|_{I_\nu}}[\hat{U}, C^T] = rk_{\mathcal{M}|_{I_\nu}}\hat{G}_\nu$ there exists $T_\nu \in \mathcal{A}\,|_{I_\nu}^{(\ell+p)\times(\ell+p)}$ so that $T_\nu \in GL_{\ell+p}(\mathcal{M}\,|_{I_\nu})$ and $[\hat{U}, C^T]_\nu\, T_\nu^{-1} = \hat{G}_\nu$. Now it follows from (3.4) and Proposition 5.9 (ii) that $\overline{\mathcal{V}^\perp}$ is meromorphically $(-A^T, C^T)$–invariant.

Conversely, if $\mathcal{V}^\perp$ is meromorphically $(-A^T, C^T)$–invariant, then

$$(DI_n + A^T)(\hat{u}) = \hat{U}\, m + C^T\, n \text{ for some } \ n \in \mathcal{M}_p^\ell,\ m \in \mathcal{M}_p^p.$$

Thus (3.3) yields

$$< (DI_n - A(t))(\hat{w}(t)), \hat{u}(t) >= 0 \text{ for all } t \in \mathbb{R}$$

whence $(DI_n - A(t))(\hat{w}(t)) \in \mathcal{V}(t)$ for all $t \in \mathbb{R}$. This completes the proof. □

For time-invariant systems $(A,B) \in \mathbb{R}^{n\times(n+m)}$ it is well-known (see e.g. *Wonham* (1985) p. 91) that the *maximal (A,B)–invariant subspace* $\mathcal{V}^*$ included in a subspace $\mathcal{L} \subset \mathbb{R}^n$ can be determined as follows

$$\begin{aligned} \mathcal{V}^o &:= \mathcal{L} \\ \mathcal{V}^i &:= \mathcal{L} \cap A^{-1}(imB + \mathcal{V}^{i-1}), \quad i \in \mathbb{N} \end{aligned}$$

This sequence is decreasing, stops after at most $k = \dim \mathcal{L}$ steps and $\mathcal{V}^k = \mathcal{V}^*$. It is not clear how to generalize this algorithm to time-varying systems and subspaces. Instead one can determine the smallest (C,A)–invariant family which contains a given family $\mathcal{L}$ and use duality to obtain the largest meromorphic (A,B)–invariant family which is included in a given family $\mathcal{L}^\perp$. For $\overline{\mathcal{V}_1}, \overline{\mathcal{V}_2} \in \overline{\mathbf{W}_n}$ define

$$\overline{\mathcal{V}_1} < \overline{\mathcal{V}_2} \quad \text{if} \quad \mathcal{V}_1(t) \overset{ae}{\subset} \mathcal{V}_2(t)$$

Definition 3.5 Suppose $A \in \mathcal{A}_p^{n\times n}$, $B \in \mathcal{A}_p^{n\times m}$, $C \in \mathcal{A}_p^{p\times n}$. Then $\overline{\mathcal{V}_*}(\mathcal{L}) \in \overline{\mathbf{W}_n}$ is called the *smallest (C,A)–invariant family* which contains $\overline{\mathcal{L}} \in \overline{\mathbf{W}_n}$ if

- $\overline{\mathcal{V}_*}(\mathcal{L})$ is (C,A)–invariant
- $\overline{\mathcal{L}} < \overline{\mathcal{V}_*}(\mathcal{L})$
- if $\overline{\mathcal{W}} \in \overline{\mathbf{W}_n}$ is (C,A)– invariant and $\overline{\mathcal{L}} < \overline{\mathcal{W}}$, then $\overline{\mathcal{V}_*}(\mathcal{L}) < \overline{\mathcal{W}}$

$\overline{\mathcal{V}^*(\mathcal{L}')} \in \overline{\mathbf{W}_n}$ is called the *largest meromorphically* (A,B)*–invariant family* contained in $\overline{\mathcal{L}'} \in \overline{\mathbf{W}_n}$ if

- $\overline{\mathcal{V}^*(\mathcal{L}')}$ is meromorphically (A,B)–invariant
- $\overline{\mathcal{V}^*(\mathcal{L}')} < \overline{\mathcal{L}'}$
- if $\overline{\mathcal{W}} \in \overline{\mathbf{W}_n}$ is meromorphically (A,B)– invariant and $\overline{\mathcal{W}} < \overline{\mathcal{L}'}$, then $\overline{\mathcal{W}} < \overline{\mathcal{V}^*(\mathcal{L}')}$.

To present an algorithm which determines $\overline{\mathcal{V}_*(\mathcal{L})}$ some notations are needed. Suppose

$$\begin{aligned} \mathcal{W}(t) &\overset{\text{ae}}{\equiv} \hat{W}(t)\mathbb{R}^q && \text{for } \hat{W} \in \mathcal{A}_p^{n\times q} \text{ with p.c. ranck} \\ \ker C(t) &\overset{\text{ae}}{\equiv} \hat{C}(t)\mathbb{R}^s && \text{for } \hat{C} \in \mathcal{A}_p^{n\times s} \text{ with p.c. rank} \end{aligned}$$

Then by Lemma 1.5.3(ii) there exists $R \in \mathcal{A}_p^{q\times r_1}$ so that

$$\tilde{W} \overset{\text{ae}}{\equiv} \hat{W}\,R = \operatorname{lcrm}_{\mathcal{A}_p}(\hat{W},\hat{C}) \quad \text{for } \tilde{W} \in \mathcal{A}_p^{n\times r} \text{ with p.c. rank}$$

Now by Proposition 5.10 (ii) for $w \in \mathcal{A}_p^n$ with

$$w(t) \in \mathcal{W}(t) \cap \ker C(t) \qquad \text{for all } t \in \mathbb{R}$$

there exists $q \in \mathcal{A}_p^r$ such that

$$w(t) \overset{\text{ae}}{\equiv} \tilde{W}(t)\,q(t) \overset{\text{ae}}{\equiv} \hat{W}(t)\,R(t)\,q(t)$$

Thus it makes sense to define

$$(DI_n - A)(\mathcal{W} \cap \ker C) := (DI_n - A)(\hat{W}\,R)\mathbb{R}^r.$$

Proposition 3.6 Suppose $A \in \mathcal{A}_p^{n\times n}$, $C \in \mathcal{A}_p^{p\times n}$ and $\mathcal{L} \in \mathbf{W}_n$ is generated by $L \in \mathcal{A}_p^{n\times q}$. Then the sequence

$$\begin{aligned} \mathcal{W}_0 &:= \mathcal{L} \\ \mathcal{W}_i &:= \mathcal{W}_{i-1} + (DI_n - A)(\mathcal{W}_{i-1} \cap \ker C), \quad i \in \mathbb{N}_0 \end{aligned} \tag{3.5}$$

is increasing in the sense that

$$\overline{\mathcal{W}_i} < \overline{\mathcal{W}_{i+1}} \qquad \text{for } i \in \mathbb{N}_0$$

and there exists $k \le n$ so that

$$\overline{\mathcal{V}_*(\mathcal{L})} = \overline{\mathcal{W}_k} = \overline{\mathcal{W}_{k+\ell}} \qquad \text{for every } \ell \in \mathbb{N}$$

Proof: Let $R_1 \in \mathcal{A}_p^{q\times r_1}$ such that

$$L\,R_1 = \operatorname{lcrm}\ (L,\hat{C})$$

then

$$\begin{aligned} \mathcal{W}_1(t) &\overset{\text{ae}}{\equiv} L(t)\mathbb{R}^q + (DI_n - A(t))(L(t)R_1(t))\mathbb{R}^{r_1} \\ &= W_1(t)\mathbb{R}^{q+r_1} \end{aligned}$$

where

$$W_1 := [L, (\dot{L} - A\, L)R_1]$$

Proceeding in this way one obtains

$$\mathcal{W}_i(t) \overset{\text{ae}}{=} W_i(t)\mathbb{R}^{q+r_1+\dots+r_i}$$

where

$$W_i := [W_{i-1}, (\dot{W}_{i-1} - A\, W_{i-1})R_i], \qquad W_{i-1}R_i = \operatorname{lcrm}_{\mathcal{A}_p}(W_{i-1}, \hat{C})$$

Therefore $\overline{\mathcal{W}_i} < \overline{\mathcal{W}_{i+1}}$ for $i \in \mathbb{N}_0$ and if for some $k \in \mathbb{N}_0$, $rk_{\mathcal{M}}\, W_k = rk_{\mathcal{M}}\, W_{k+1}$ then $\overline{\mathcal{W}_k} = \overline{\mathcal{W}_{k+\ell}}$ for all $\ell \in \mathbb{N}$. By construction $\overline{\mathcal{W}_k}$ is (C, A)-invariant and $\overline{\mathcal{L}} < \overline{\mathcal{W}_k}$. So it remains to prove that if $\overline{\mathcal{V}}$ is (C, A)-invariant and $\overline{\mathcal{L}} < \overline{\mathcal{V}}$, then $\overline{\mathcal{W}_k} < \overline{\mathcal{V}}$. By assumption

$$(DI_n - A)(\mathcal{V} \cap \ker C) \overset{\text{ae}}{\subset} \mathcal{V}$$

and by induction on i one gets

$$\begin{aligned} \mathcal{W}_i &= \mathcal{W}_{i-1} + (DI_n - A)(\mathcal{W}_{i-1} \cap \ker C) \\ &< \mathcal{V} + (DI_n - A)(\mathcal{V}) \end{aligned}$$

This completes the proof. □

The duality between the smallest (C, A)– and largest meromorphically (A, B)–invariant family is given as follows

Proposition 3.7 Suppose $A \in \mathcal{A}_p^{n\times n}$, $C \in \mathcal{A}_p^{p\times n}$. If $\overline{\mathcal{L}} \in \overline{\mathcal{W}_n}$ and some $\mathcal{L} \in \overline{\mathcal{L}}$ is generated by $L \in \mathcal{A}_p^{n\times k}$ then the following are equivalent

(i) $\overline{\mathcal{V}}$ is smallest (C, A)–invariant family containing $\overline{\mathcal{L}}$.

(ii) $\overline{\mathcal{V}^\perp}$ is the largest meromorphically $(-A^T, C^T)$–invariant family included in $\overline{\mathcal{L}^\perp}$.

Proof: (i) $\Rightarrow$ (ii) : If $\mathcal{W}$ is a representative of the largest meromorphically $(-A^T, C^T)$–invariant family which is included in $\mathcal{L}^\perp$ then

$$\mathcal{V}^\perp(t) \overset{\text{ae}}{\subset} \mathcal{W}(t) \overset{\text{ae}}{\subset} \mathcal{L}^\perp(t)$$

and thus

$$\mathcal{L}(t) \overset{\text{ae}}{\subset} \mathcal{W}^\perp(t) \overset{\text{ae}}{\subset} \mathcal{V}(t)$$

Since $\overline{\mathcal{V}}$ is the smallest (C, A)–invariant family it follows that $\overline{\mathcal{W}} = \overline{\mathcal{V}}$.
The reverse direction is proved analogously. □

Remark 3.8 For a real analytic system $(A, B) \in \mathcal{A}^{n\times(n+m)}$ it is demonstrated how to determine $\overline{\mathcal{V}^*}(\ker C(t))$: By Proposition 3.7 this problem is equivalent to determine $\overline{\mathcal{V}_*}$, the smallest $(B^T, -A^T)$–invariant family containing $\overline{\ker C(t))^\perp}$.
Let

$$(\ker C(t))^\perp \overset{\subseteq}{\overset{\text{ae}}{=}} L(t)\mathbb{R}^q \quad \text{for some } L \in \mathcal{A}^{n\times q} \text{ with } rk_{\mathbb{R}}\, L(t) = q \;\; \forall\, t \in \mathbb{R} \tag{3.6}$$

and

$$\ker\; B^T(t) \overset{\supseteq}{\overset{\text{ae}}{=}} \hat{C}(t)\mathbb{R}^s \quad \text{for some } \hat{C} \in \mathcal{A}^{n\times s} \text{ with } rk_{\mathbb{R}}\, \hat{C}(t) = s \; \forall\, t \in \mathbb{R} \tag{3.7}$$

Applying algorithm (3.5) yields

$$(\ker C(t))^{\perp} \subset L(t)\mathbb{R}^{q} \subset \mathcal{W}_{i-1}(t) \subset \mathcal{W}_{i}(t) \subset \hat{V}(t)\mathbb{R}^{n-k}$$

for some

$$\hat{V} \in \mathcal{A}^{n\times(n-k)} \text{ with } rk_{\mathbb{R}}\hat{V}(t) = n-k \text{ for all } t \in \mathbb{R} \text{ and } \overline{(\hat{V}(t)\mathbb{R}^{n-k})_{t\in\mathbb{R}}} = \overline{\mathcal{W}_n}.$$

Thus

$$\ker C(t) \supset (L(t)\mathbb{R}^{q})^{\perp} \supset \mathcal{W}_{i-1}^{\perp} \supset \mathcal{W}_{i}^{\perp}(t) \supset V(t)\mathbb{R}^{k}$$

for $V \in \mathcal{A}^{n\times k}$ such that $(\hat{V}(t)\mathbb{R}^{n-k})^{\perp} = V(t)\mathbb{R}^{k}$ for all $t \in \mathbb{R}$. By Proposition 3.7 it follows that $\overline{\mathcal{V}} = \overline{\mathcal{V}^{*}}(\ker C(t))$.

Example 3.9 $\overline{\mathcal{V}^{*}}(\ker C)$ will be calculated for a system (A, B, C) where $A \in \mathcal{A}^{3\times3}$ and $B \in \mathcal{A}^{3\times1}$ are as in Example 1.11 and $C(t) := [1, 0, t^2]$. Using the notation of Remark 3.8 one has

$$L(t) = \begin{bmatrix} 1 \\ 0 \\ t^2 \end{bmatrix}, \quad \hat{C}(t) = \begin{bmatrix} 1 & 1 \\ 0 & t-2 \\ t^2 & 0 \end{bmatrix}$$

which satisfy (3.6) and (3.7). For this situation the algorithm

$$\begin{aligned} \mathcal{W}_0 &= L\,\mathbb{R}^1 \\ \mathcal{W}_i &= \mathcal{W}_{i-1} + (DI_n + A^T)(\mathcal{W}_{i-1} \cap \hat{C}\mathbb{R}^2) \end{aligned}$$

is as follows

$$\begin{aligned} \mathcal{W}_1(t) &= \begin{bmatrix} 1 \\ 0 \\ t^2 \end{bmatrix} \cdot \mathbb{R} + (D + A^T(t))\left(\begin{bmatrix} 1 \\ 0 \\ t^2 \end{bmatrix} \cdot \mathbb{R} \right) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & \sin t \\ t^2 & 0 \end{bmatrix} \cdot \mathbb{R}^2 \\ \mathcal{W}_2(t) &= \mathcal{W}_1(t) + (D + A^T(t))([1, 0, t^2]^T \cdot \mathbb{R}) \\ &= \mathcal{W}_i(t) \qquad \text{for all } i \geq 1. \end{aligned}$$

Therefore $\overline{\mathcal{V}_*}$ given by

$$\mathcal{V}_*(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ t^2 & 0 \end{bmatrix} \cdot \mathbb{R}^2$$

is the smallest $(B^T, -A^T)$–invariant family containing $\overline{(\ker C(t))^{\perp}}$ and thus $\overline{\mathcal{V}^{*}}$ is given by

$$\mathcal{V}^{*}(t) = [-t^2, 0, 1]^T \cdot \mathbb{R}$$

Now we are in a position to prove the main result of this section which is a summary of the foregoing.

Proposition 3.10 The set

$$\mathcal{S}_{C,A} \text{ of}(C,A)-\text{invariant families of } \overline{\mathbf{W}_n}$$

is a lattice with respect to

$$\begin{array}{ccccc} \overline{\mathcal{V}_1} & \wedge & \overline{\mathcal{V}_2} & = & \overline{\mathcal{V}_1 \cap \mathcal{V}_2} \\ \overline{\mathcal{V}_1} & \vee & \overline{\mathcal{V}_2} & = & \overline{\mathcal{V}_*(\mathcal{V}_1 + \mathcal{V}_2)} \end{array}$$

where $\overline{\mathcal{V}_*}(\mathcal{V}_1 + V_2)$ denotes the smallest (C,A)–invariant family which contains $\overline{\mathcal{V}_1 + \mathcal{V}_2}$.
Furthermore the set

$$\mathcal{S}_{A,B} \text{ of meromorphically } (A,B)-\text{invariant families of } \overline{\mathbf{W}_n}$$

is a lattice with respect to

$$\begin{array}{ccccc} \overline{\mathcal{V}_1} & \wedge & \overline{\mathcal{V}_2} & = & \overline{\mathcal{V}^*}(\mathcal{V}_1 \cap \mathcal{V}_2) \\ \overline{\mathcal{V}_1} & \vee & \overline{\mathcal{V}_2} & = & \overline{\mathcal{V}_1 + \mathcal{V}_2} \end{array}$$

where $\overline{\mathcal{V}^*}(\mathcal{V}_1 \cap \mathcal{V}_2)$ denotes the largest meromorphically (A,B)–invariant family which is included in $\overline{\mathcal{V}_1 \cap \mathcal{V}_2}$.
The map

$$\phi : \mathcal{S}_{C,A} \to \mathcal{S}_{-A^T,C^T}, \quad \overline{\mathcal{V}} \mapsto \overline{\mathcal{V}^\perp}$$

is a lattice anti-isomorphism, where "anti" means

$$\begin{array}{ccc} \phi(\overline{\mathcal{V}_1} \wedge \overline{\mathcal{V}_2}) & = & \phi(\overline{\mathcal{V}_1}) \vee \phi(\overline{\mathcal{V}_2}) \\ \phi(\overline{\mathcal{V}_1} \vee \overline{\mathcal{V}_2}) & = & \phi(\overline{\mathcal{V}_1}) \wedge \phi(\overline{\mathcal{V}_2}) \end{array}$$

Compare Figure 3.1.

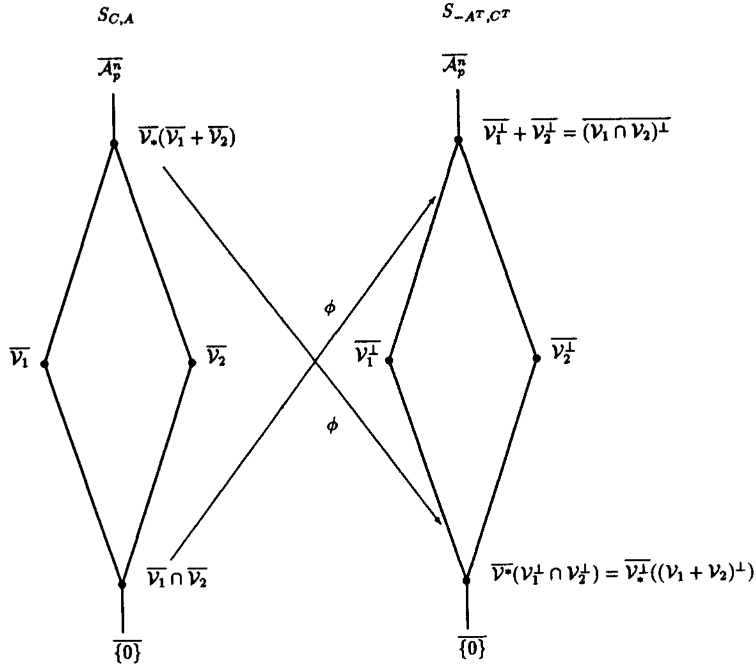

Figure 3.1.:

Proof: It is easily seen that the definition of the lattice operations does not depend on the representatives. It remains to prove that ϕ is an anti-isomorphism. Using the fact that for finite dimensional vector spaces $(\mathcal{V}_1(t)\cap\mathcal{V}_2(t))^\perp = \mathcal{V}_1(t)^\perp + \mathcal{V}_2(t)^\perp$ holds true, we obtain

$$\begin{aligned}\phi(\overline{\mathcal{V}_1}\wedge\overline{\mathcal{V}_2}) &= \phi(\overline{\mathcal{V}_1\cap\mathcal{V}_2}) = \overline{(\overline{\mathcal{V}_1}\cap\overline{\mathcal{V}_2})^\perp} = \overline{\overline{\mathcal{V}_1^\perp}+\overline{\mathcal{V}_2^\perp}}\\ &= \overline{\phi(\mathcal{V}_1)+\phi(\mathcal{V}_2)} = \phi(\overline{\mathcal{V}_1})\vee\phi(\overline{\mathcal{V}_2})\end{aligned}$$

This proves the first equation of the anti-isomorphism. To prove the second one use Proposition 3.7 to conclude

$$\begin{aligned}\phi(\overline{\mathcal{V}_1}\vee\overline{\mathcal{V}_2}) &= \phi(\overline{\mathcal{V}_*}(\mathcal{V}_1+\mathcal{V}_2)) &= \overline{\mathcal{V}_*^\perp}(\mathcal{V}_1+\mathcal{V}_2)\\ &= \overline{\mathcal{V}^*}((\mathcal{V}_1+\mathcal{V}_2))^\perp) &= \overline{\mathcal{V}^*}(\mathcal{V}_1^\perp\cap\mathcal{V}_2^\perp) = \phi(\overline{\mathcal{V}_1})\wedge\phi(\overline{\mathcal{V}_2}).\end{aligned}$$

□

3.4 Disturbance decoupling problem

In this section we consider a system $(A,B,C) \in \mathcal{A}_p^{n\times n} \times \mathcal{A}_p^{n\times m} \times \mathcal{A}_p^{p\times n}$ with an additional disturbance $q(\cdot) \in \mathcal{C}_p^s$ entering the system via $S \in \mathcal{A}_p^{n\times s}$.

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) + S(t)q(t) \\ y(t) &= C(t)x(t) \end{aligned} \tag{4.1}$$

The *disturbance decoupling problem* (**DDP**) on (t_0, t_1) is to find a state feedback matrix $F(\cdot) \in \mathcal{A}_p^{m\times n}$ such that arbitrary $q(\cdot)$ has no influence on the output

$$y(t) = C(t)[\Phi_F(t,t_0)x_0 + \int_{t_0}^{t} \Phi_F(t,s)S(s)q(s)ds] \qquad \text{for all } t \in (t_0,t_1)$$

of the closed loop system

$$\begin{aligned} \dot{x}(t) &= A_F(t)x(t) + S(t)q(t) \\ y(t) &= C(t)x(t) \end{aligned} \tag{4.2}$$

where

$$A_F(t) := A(t) + B(t)F(t)$$

and $\Phi_F(\cdot,\cdot)$ denotes the transition matrix of $\dot{x}(t) = A_F(t)x(t)$. This notation will be used throughout the remaining sections. The following definition is an extension of the time-invariant case, see *Wonham* (1985).

Definition 4.1 The system (4.1) is called *disturbance decoupled on* (t_0,t_1) if for some $F(\cdot) \in \mathcal{A}_p^{m\times n}$

$$y(t) = C(t)\int_{t_0}^{t} \Phi_F(t,s)S(s)q(s)ds = 0 \qquad \text{for all } t \in (t_0,t_1) \text{ and arbitrary } q(\cdot) \in \mathcal{C}_p^s \tag{4.3}$$

Using the controllability Gramian of the closed loop system (4.2)

$$W_F(t_0,t_1) := \int_{t_0}^{t} \Phi_F(t_0,s)S(s)S^T(s)\Phi_F^T(t_0,s)ds$$

the DDP for piecewise analytic systems can be characterized as follows

Proposition 4.2 Let $I = (t_0,t_1)$ and $F(\cdot) \in \mathcal{A}_p^{m\times n}$. $\mathbb{R} = \bigcup_{\nu\in\mathbb{Z}}[a_\nu, a_{\nu+1})$ denotes a partition so that $A_F(\cdot)$ and $S(\cdot)$ are real analytic on $(a_\nu, a_{\nu+1}) \cap I$. Then the following are equivalent

(i) (4.1) is disturbance decoupled on I by $F(\cdot)$

(ii) $\Phi_F(t,t_0)\, im\, W_F(t_0,t) \subset \ker\, C(t)$ for all $t \in I$

(iii) $\Phi_F(t,t_0)[\sum_{i\geq 0} im\,(DI_n - A_F(t_0))^i(S(t_0))$
$+ \sum_{\nu=1}^{N}\sum_{i\geq 0} im\,(DI_n - A_F(a_\nu))^i(S(a_\nu))] \subset \ker C(t)$ for all $t \in I$
where $t_0 \in [a_0,a_1)$, $t \in (a_N, a_{N+1})$.

Proof: Consider the map

$$\begin{array}{rcll} L_{t_0,t} & : & (C_p)^s & \rightarrow \mathbb{R}^n \\ & & q(\cdot) & \mapsto \int_{t_0}^{t} \Phi_F(t,s)S(s)q(s)ds \end{array}$$

By *Knobloch and Kappel* (1974) p. 103 it follows that

$$\begin{aligned} im\, L_{t_0,t} &= im \int_{t_0}^{t} \Phi_F(t,s)S(s)S^T(t)\Phi_F^T(t,s)ds \\ &= \Phi_F(t,t_0)\, im\, W_F(t_0,t) \end{aligned}$$

This proves (i) $\Leftrightarrow$ (ii). (ii) $\Leftrightarrow$ (iii) is a consequence of Proposition 1.2.6. □

If the system (4.1) is analytic then by Proposition 1.2.6 we obtain the following simple result

Corollary 4.3 An analytic system (4.1) is disturbance decoupled on I by $F(\cdot) \in \mathcal{A}\mid_I^{m\times n}$ iff

$$\sum_{i\geq 0} im\,(DI_n - A_F(t))^i(S(t)) \subset \ker C(t) \qquad \text{for all } t \in I.$$

Due to the Identity-Theorem of analytic functions, for an analytic system (4.1) condition (4.3) has to be checked only on an arbitrary small interval $(t_0, t_0+\varepsilon)$. More precisely we have

Proposition 4.4 Suppose (4.1) is analytic and $F(\cdot) \in \mathcal{A}\mid_I^{m\times m}$ where $I = (t_0, t_1)$, then the following are equivalent:

(i) (4.1) is disturbance decoupled on I by $F(\cdot)$

(ii) (4.1) is disturbance decoupled on $(t_0, t_0+\varepsilon)$ by $F(\cdot)$ for arbitrary $\varepsilon \in (0, t_1 - t_0)$

Proof: By Proposition 1.2.2, Definition 4.1 does not depend on whether we admit piecewise continuous or analytic disturbance. Since the vector function

$$t \mapsto \varphi(t,q) := \int_{t_0}^{t} \Phi_F(t,s)S(s)q(s)ds$$

is real analytic on I for every $q \in \mathcal{A}\mid_I^q$ the Identity-Theorem of analytic functions yields

$$C(t)\varphi(t,q) = 0 \qquad \text{for all } t \in I$$

if and only if

$$C(t)\varphi(t,q) = 0 \qquad \text{for all } t \in (t_0, t_0+\varepsilon),\ 0 < \varepsilon < t_1 - t_0$$

This proves the proposition. □

For an analytic system (4.1) the largest meromorphically (A,B)–invariant subspace $\overline{\mathcal{V}^*}(\ker C)$ included in $\ker C(t)$ with generator $V \in \mathcal{A}^{n\times k}$ of constant rank k was constructed in Remark 3.8. By Proposition 1.10 one obtains

$$(DI_n - A_F)(V) = VU_1T^{-1}W \qquad \text{where } F = U_2T^{-1}W(V^TV)^{-1}V^T \tag{4.4}$$

Thus the set of *critical points* for the feedback F is given by

$$\mathcal{P} = \{t' \in \mathbb{R} \mid \text{ an entry of } U_2(\cdot)T^{-1}(\cdot)W(\cdot) \text{ has a pole at } t'\}$$

Let $I \subset \mathbb{R}$ be an open interval. Then

$$F(\cdot) \text{ is analytic on } I \quad \text{if } \mathcal{P} \cap I = \emptyset$$

and furthermore by Proposition 1.10

$$\mathcal{P} \cap I = \emptyset \qquad \text{if} \quad rk_{\mathbb{R}}[V(t), B(t)] = \text{ const.} \quad \text{for all } \; t \in I$$

Now for every $I \subset \mathbb{R} \setminus \mathcal{P}$ the differential equation

$$\dot{x}(t) = A_F(t)x(t) \quad , t \in I$$

is solvable on I . This sets us in a position to state the main result of this section which is a generalization of the constant case (see *Wonham* (1985) Theorem 4.2).

Theorem 4.5 Suppose the system (4.1) is analytic and $V \in \mathcal{A}^{n \times k}$ with $rk_{\mathbb{R}} V(t) = k$ for all $t \in \mathbb{R}$ generates $\overline{\mathcal{V}^*}(\ker C)$ constructed in Remark 3.8. Then for $I = (t_0, t_1)$ we have:

(i) If the DDP is solvable on I by $F(\cdot) \in \mathcal{A}\,|_{\overline{I}}^{n \times n}$, then

$$S(t)\mathbb{R}^s \subset V(t)\mathbb{R}^k \quad \text{for all } t \in I$$

(ii) If $\overline{I} \subset \mathbb{R} \setminus \mathcal{P}$ and

$$S(t)\mathbb{R}^s \subset V(t)\mathbb{R}^k \quad \text{for all } t \in I$$

then the DDP is solvable on I by $F(\cdot) \in \mathcal{A}\,|_{\overline{I}}^{m \times n}$ given in (4.4)

Proof: (i) : By Corollary 4.3

$$im\, S(t) \subset \sum_{i \geq 0} im\,(DI_n - A_F(t))^i(S(t)) \subset \ker\, C(t) \qquad \text{for all } \; t \in I.$$

By Remark 1.5.11 there exists $\tilde{V} \in \mathcal{A}\,|_I^{n \times s}$ with constant rank on I so that

$$\tilde{\mathcal{V}}(t) := \tilde{V}(t)\mathbb{R}^s = \sum_{i \geq 0} im\,(DI_n - A_F(t))^i(S(t)) \quad \text{for all } \; t \in I$$

Thus Theorem 1.5 yields that $\tilde{\mathcal{V}}$ is (A, B)–invariant on I . This together with $\tilde{\mathcal{V}}(t) \subset \ker\, C(t)$ for all $t \in \mathbb{R}$ gives

$$\tilde{V}(t)\mathbb{R}^s \overset{ae}{\subset} V(t)\mathbb{R}^k \quad ; \quad \text{for all } t \in I$$

Since $V(t)$ has constant rank on I one gets

$$S(t)\mathbb{R}^s \subset \tilde{V}(t)\mathbb{R}^s \subset V(t)\mathbb{R}^k \qquad \text{for all } t \in I$$

which proves (i).

(ii): Since $S(t)\mathbb{R}^s \subset V(t)\mathbb{R}^k$ for all $t \in I$ and $(V(t)\mathbb{R}^k)_{t \in \mathbb{R}}$ is A_F–invariant one obtains

$$\begin{aligned} im\, S(t) \;&\subset\; \sum_{i \geq 0} im\,(DI_n - A_F(t))^i(S(t)) \\ &\subset\; \sum_{i \geq 0} im\,(DI_n - A_F(t))^i(V(t)) \\ &=\; V(t)\mathbb{R}^k \subset \ker\, C(t) \qquad\qquad \text{for all } t \in I \end{aligned}$$

Now (ii) follows from Corollary 4.3. □

Remark 4.6 If the entries of the matrices of (4.1) belong to the ring $\mathbb{R}[t]$ then it is computationally not too expensive to check the assumptions of Theorem 4.5 (ii). The main tool is to transform a matrix into an upper triangular form. This algorithm is described in detail for instance in *Wolovich* (1979). It can be implemented via the algebraic programming system *Reduce* (see *Hearn* (1985)).

Example 4.7 Consider a system of the form (4.1) specified by

$$\begin{aligned} A(t) &= \begin{bmatrix} t_2 & \sin t & -t(t^2+2) \\ a_4(t) & a_5(t) & a_6(t) \\ -1 & 0 & t \end{bmatrix}, \quad B(t) = \begin{bmatrix} -t^2(t-2) \\ t^2 \\ t-2 \end{bmatrix}, \\ C(t) &= [1, 0, t^2], \qquad\qquad\qquad\qquad S(t) = \begin{bmatrix} t^3 \\ 0 \\ -t \end{bmatrix} \end{aligned}$$

By Example 3.9 $V(t) = [-t^2, 0, 1]^T$ is a generator of $\mathcal{V}^*(\ker C)$. By Example 1.9 the set of critical points is $\mathcal{P} = \{0\}$. Since

$$im\ S(t) \subset \mathcal{V}^*(t) \quad \text{for all } t \in \mathbb{R}$$

Theorem 4.5 (ii) says that the disturbance decoupling problem is solvable on every open interval $I \subset \mathbb{R}$ with $0 \notin \bar{I}$.

3.5 Controllability subspace families

In this section we will extend the concept of controllability subspaces (see *Wonham* (1985)) to analytic time-varying system $(A, B) \in \mathcal{A}^{n\times(n+m)}$.

Definition 5.1 A family of subspaces $\mathcal{V} \in \mathbf{W}_n$ generated by $V \in \mathcal{A}^{n\times k}$ is called a *controllability subspace family* (c.s.f.) of $(A, B) \in \mathcal{A}^{n\times(n+m)}$ if

(i) $\mathcal{V}$ is (A, B)–invariant

(ii) for every $x_0 \in \mathcal{V}(t_0)$, $x_1 \in \mathcal{V}(t_1)$, $t_0 < t_1$, there exists a control $u(\cdot) \in \mathcal{C}_p^m$ such that the forced trajectory of $\dot{x}(t) = A(t)x(t) + B(t)u(t)$ satisfies

$$x(t) \in \mathcal{V}(t) \quad \text{for all } t \in (t_0, t_1) \text{ and } \quad x(t_0) = x_0, x(t_1) = x_1. \tag{5.1}$$

In case of time-invariant systems (ii) implies (i) (see *Wonham* (1985) Section 5.1).

Example 5.2 The controllable family $\Re$ (see Section 1.5) of $(A, B) \in \mathcal{A}^{n\times(n+m)}$ is a c.s.f. In Example 1.9 it is shown that $\Re$ is A–invariant whence it is (A, B)–invariant. For $x_0 \in \Re(t_0)$ and $x_1 \in \Re(t_1)$ a control $u(\cdot) \in \mathcal{C}_p^m$ satisfying (5.1) can be constructed as folllows: Set

$$\begin{aligned} \hat{x}_1 &:= \Phi(t_0, t_1)x_1 &\in \Re(t_0) \\ x_\Delta &:= x_0 - \hat{x}_1 &\in \Re(t_0) \end{aligned}$$

Choose $u(\cdot) \in \mathcal{C}_p^m$ such that

$$\hat{x} := \Phi(t, t_0)x_\Delta + \int_{t_0}^{t} \Phi(t, s)B(s)u(s)ds$$

fulfills $\hat{x}(t_1) = 0$ and $\hat{x}(t) \in \Re(t)$ for $t \in [t_0, t_1]$. Thus

$$\begin{array}{rcl} x(t) & := & \Phi(t,t_0)x_0 \ + \ \int_{t_0}^{t} \Phi(t,s)B(s)u(s)ds \\ & = & \Phi(t,t_1)x_1 \ + \ \hat{x}(t) \qquad\qquad \in \Re(t) \end{array}$$

with $x_0(t_0) = x_0$ and $x(t_1) = x_1$.

Proposition 5.3 Suppose $\mathcal{V}(t) = V(t)\mathbb{R}^k$ for some $V \in \mathcal{A}^{n\times k}$ with $rk_{\mathbb{R}}\, V(t) = k$ for all $t \in \mathbb{R}$. Then $\mathcal{V}$ is a c.s.f. of $(A, B) \in \mathcal{A}^{n\times(n+m)}$ if and only if

$$\mathcal{V}(t) = \sum_{i\geq 0} im\,(DI_n - A_F(t))^i(BG(t)) \qquad \text{for all } t \in \mathbb{R} \tag{5.2}$$

for some $F \in \mathcal{A}^{m\times n}, G \in \mathcal{A}^{m\times m}$.

Proof: Assume that (5.2) is valid. Then for given $x_0 \in \mathcal{V}(t_0)$, $x_1 \in \mathcal{V}(t_1)$ there exists by Example 5.2 $\hat{u} \in \mathcal{C}_p^m$ such that

$$\begin{array}{rcl} \dot{x}(t) & = & A_F(t)x(t) + BG(t)\hat{u} \\ x(t) & \in & \mathcal{V}(t) \qquad \text{for } t \in [t_0, t_1],\ x(t_0) = x_0,\ x(t_1) = x_1 \end{array} \tag{5.3}$$

Thus condition (ii) of Definition 5.1 is satisfied. $\mathcal{V}$ is $(A, B)-$ invariant since it is the controllable family of the system (5.3), see Remark 1.5.11.
To prove the converse let $F \in \mathcal{A}^{m\times n}$ such that for some $N \in \mathcal{A}^{m\times n}$

$$(DI_n - A_F)(V) = V\,N$$

Then $\mathcal{V}$ is also a c.s.f. of the system

$$\dot{x} = A_F(t)x(t) + B(t)u(t)$$

Choose by Lemma 1.5.1 $G \in \mathcal{A}^{m\times m}$, $L \in \mathcal{A}^{k\times m}$ such that

$$B\,G = V\,L = \ell c\ell m\,(B, V).$$

This proves ' $\subset$ ' in (5.2) . For the reverse inclusion let $\Re(t)$ denote the controllable family of (5.3). Clearly $\mathcal{V}(t) \subset \Re(t)$, and since $\Re(t)$ can be presented by the right hand side of (5.2) (see (1.(5.5)) the proof is complete. □

The foregoing and the following proposition are generalizations of the constant case, see *Wonham* (1985) p. 104.

Proposition 5.4 Suppose $\mathcal{V} \in \mathbf{W}_n$ is generated by $V \in \mathcal{A}^{n\times k}$ with constant rank k . If $\mathcal{V}$ is a c.s.f. of $(A, B) \in \mathcal{A}^{n\times(n+m)}$ and

$$imB(t) \cap \mathcal{V}(t) \ = \ imBG(t) \qquad \text{for some } G \in \mathcal{A}^{m\times n}$$

then

$$\mathcal{V}(t) = \sum_{i\geq 0} im\,(DI_n - A_F(t))^i(BG(t)) \qquad \text{for all } t \in \mathbb{R}$$

for any $F \in \mathcal{A}^{m\times n}$ which satisfies

$$(DI_n - A_F)(V) = V\,N \qquad \text{for some } N \in \mathcal{A}^{k\times k} \tag{5.4}$$

Proof: By Proposition 5.3 there exists an $F_0 \in \mathcal{A}^{m\times n}$ so that

$$\mathcal{V}(t) = \sum_{i\geq 0} im\ (DI_n - A_{F_0}(t))^i(BG(t))$$

Put

$$\mathcal{V}'(t) := \sum_{i\geq 0} im\ (DI_n - A_F(t))^i(BG(t))$$

for some F which satisfies (5.4), then $\mathcal{V}'(t) \subset \mathcal{V}(t)$. For the reverse inclusion it is sufficient to show that $\mathcal{V}'$ is $(A + BF_0)$–invariant. This is proved completely analogously to the constant case, see *Wonham* (1985) p. 105. □

3.6 Noninteracting control

Consider a state space system with several outputs

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ z_i(t) &= C_i(t)x(t), \qquad\qquad i \in \underline{k} \end{aligned} \tag{6.1}$$

where $A(\cdot)$, $B(\cdot)$, $C_i(\cdot)$ are $n\times n$, $n \times m$, $p_i \times n$ matrices, resp. defined over $\mathcal{A}$. The *restricted decoupling problem* (**RDP**) for (6.1) is to find an $F \in \mathcal{A}^{m\times n}$ and c.s.f.'s $\mathcal{V}_i \in \mathbf{W}_n$, $i \in \underline{k}$, such that the following conditions are satisfied for all $t \in \mathbb{R}$, $i \in \underline{k}$

$$\left.\begin{array}{l} \mathcal{V}_i(t) = \sum_{\lambda\geq 0} im\ (DI_n - A_F(t))^\lambda(BG_i(t)) \\ \text{where } G_i \in \mathcal{A}^{m\times m} \text{ such that } imBG_i(t) = imB(t) \cap \mathcal{V}_i(t) \end{array}\right\} \tag{6.2}$$

$$C_j(t)\mathcal{V}_i(t) = 0 \qquad \text{for } i \neq i \tag{6.3}$$

$$C_i(t)\mathcal{V}_i(t) = imC_i(t) \tag{6.4}$$

(6.3) is called the *noninteraction* condition and is equivalent to

$$\mathcal{V}_i(t) \subset \bigcap_{j\neq i} ker\ C_j(t) \tag{6.5}$$

(6.4) is called the *output controllability* condition and is equivalent to

$$\mathcal{V}_i(t) + ker\ C_i(t) = \mathbb{R}^n \tag{6.6}$$

(6.2) is referred to as the *compatibility* condition of the families $\mathcal{V}_i$.

Definition 6.1 Some families $\mathcal{V}_i \in \mathbf{W}_n$ with generator $V_i \in \mathcal{A}^{n\times r_i}$ ($i \in \underline{k}$) are called *compatible* relative to (6.1) if there exist $F \in \mathcal{A}^{m\times n}$ and $N_i \in \mathcal{A}^{r_i\times r_i}$, so that

$$(DI_n - A_F)(V_i) = V_i\, N_i \qquad \text{for } i \in \underline{k} \tag{6.7}$$

Lemma 6.2 Suppose $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_1 \cap \mathcal{V}_2 \in \mathbf{W}_n$ are generated by $V_i \in \mathcal{A}^{n\times r_i}$ with $rk_{\mathbb{R}} V_i(t) =$ const. for all $t \in \mathbb{R}$, $i \in \underline{3}$ resp. If there exist $F_i \in \mathcal{A}^{m\times n}$ and $N_i \in \mathcal{A}^{r_i\times r_i}$ so that

$$(DI_n - A_{F_i})(V_i) = V_i\, N_i \qquad\qquad \text{for } i \in \underline{3}$$

then $\mathcal{V}_1$ and $\mathcal{V}_2$ are compatible.

Proof: Let

$$\begin{array}{rcccl} P_i(t) & : & \mathbb{R} & \to & \mathcal{V}_i(t) \setminus \mathcal{V}_1(t) \cap \mathcal{V}_2(t) \qquad i = 1,2 \\ P_3(t) & : & \mathbb{R} & \to & \mathcal{V}_1(t) \cap \mathcal{V}_2(t) \end{array}$$

denote the orthogonal projection on $\mathcal{V}_i(t) \setminus \mathcal{V}_1(t) \cap \mathcal{V}_2(t)$, $\mathcal{V}_1(t) \cap \mathcal{V}_2(t)$, resp. Then by the assumptions and Proposition 1.5.6 it follows that $P_i \in \mathcal{A}^{n \times n}$ for $i \in \underline{3}$. Thus for

$$F := F_1\, P_1 + F_2\, P_2 + F_3\, P_3 \in \mathcal{A}^{m \times n}$$

(6.7) is satisfied. □

Definition 6.3 Some families $\mathcal{V}_i \in \mathbf{W}_n$, $i \in \underline{k}$, are called *independent* if

$$\mathcal{V}_i(t) \cap \sum_{j \neq i}^{k} \mathcal{V}_j = \{0\} \qquad \text{for all } i \in \underline{k}.$$

Lemma 6.4 Let $\mathcal{V}_i \in \mathbf{W}_n$ be generated by $V_i \in \mathcal{A}^{n \times r_i}$ with $rk_{\mathbb{R}} V_i(t) = \text{const.}$ for all $t \in \mathbb{R}$, $i \in \underline{k}$. If the families $\mathcal{V}_i$ are independent and

$$(DI_n - A_{F_i})(V_i) = V_i\, N_i \quad \text{for } i \in \underline{k}, \quad \text{for some} \quad F_i \in \mathcal{A}^{m \times n},\ N_i \in \mathcal{A}^{r_i \times r_i}$$

then $\mathcal{V}_i$ are compatible.

Proof: Since $\mathcal{V}_i$ are independent, there exists a $Y \in \mathbf{W}_n$ such that

$$\mathbb{R}^n = \mathcal{V}_1(t) \oplus \ldots \oplus \mathcal{V}_k(t) \oplus Y(t) \qquad \text{for all } t \in \mathbb{R}.$$

According to this decomposition we define

$$\begin{array}{rcccc} F(t) & : & \mathbb{R}^n & \to & \mathbb{R}^n \\ & & \sum_{i=1}^k v_i(t) + y(t) & \mapsto & \sum_{i=1}^k F_i(t) v_i(t) \end{array}$$

Since $\mathcal{V}_i$ have constant dimensions, $F \in \mathcal{A}^{m \times n}$. Thus F satisfies (6.7). □

Using the previous lemmata we are now in a position to prove the main result of this section, i.e. a characterization of the RDP which is a generalization of the constant case given in *Wonham* (1985) Section 9.3.

Proposition 6.5 Suppose

$$\bigcap_{i=1}^{k} ker\ C_i(t) = \{0\}$$

Then the RDP is solvable iff there exists c.s.f.s $\mathcal{V}_i$ generated by $V_i \in \mathcal{A}^{n \times r_i}$ of constant ranks, $i \in \underline{k}$ resp., such that for all $t \in \mathbb{R}$ and $i \in \underline{k}$

$$\mathcal{V}_i(t) \subset \bigcap_{j \neq i} ker\ C_j(t) \tag{6.8}$$

and

$$\mathcal{V}_i +\ ker\ C_i(t) = \mathbb{R}^n. \tag{6.9}$$

Proof: The 'only if' part is immediate from the definition. To prove the 'if' part note that by (A, B)–invariance of $\mathcal{V}_i$ there exist $F_i \in \mathcal{A}^{m\times n}$, $N_i \in \mathcal{A}^{r_i\times r_i}$ such that

$$(DI_n - A_{F_i})(V_i) = V_i\, N_i \qquad ,i \in \underline{k}$$

The families $\hat{\mathcal{K}}_i$ defined by

$$\hat{\mathcal{K}}_i(t) := \bigcap_{j\neq i} \ker\, C_j(t) \qquad ,i \in \underline{k}$$

are independent. This is proved analogously to the time-invariant case, see *Wonham* (1985) p.225. Since $\mathcal{V}_i(t) \subset \hat{\mathcal{K}}_i(t)$, it follows that the $\mathcal{V}_i$'s are also independent. By Lemma 6.4 they are compatible. Application of Proposition 5.3 yields (6.2) and the proof is complete. □

Chapter 4

Stability, Stabilizability, Robustness, and Differential Riccati Equations

4.0 Introduction

In this chapter I study various problems concerning exponential stability of linear time-varying systems of the form

$$\dot{x} = A(t)x(t) \qquad , t \geq 0 \tag{0.1}$$

It is well-known that if, for each $t \geq 0$, all eigenvalues of $A(t)$ are lying in the proper open left half complex plane, then the system (0.1) is not necessarily exponentially stable (see e.g. *Wu* (1974)). Exponential stability is secured if, additionally, the parameter variation of $A(t)$ is " slow enough" , see *Rosenbrock* (1963) and *Coppel* (1978). However, these are qualitative results. In a joint paper with *Owens and Prätzel-Wolters* (1987b) we derived *quantitative* results. This means, upper bounds for the eigenvalues and for the rate of change of $A(t)$ which ensure exponential stability of (0.1) are determined. This is presented in Section 1.

Ikeda, Maeda and Kodoma (1972) and (1975) studied the problem to stabilize a system of the form

$$\dot{x}(t) = A(t)x(t) + B(t)u(t)$$

by some state feedback. In *Ilchmann and Kern* (1987) stabilizability of (0.2a) was characterized under the additional assumption that (0.1) possesses an exponential dichotomy. Furthermore *Ikeda, Maeda and Kodoma* (1975) gave a sufficient condition so that (0.2), i.e. (0.2a) and

$$y(t) = C(t)x(t) + E(t)u(t), \tag{0.2}$$

is stabilizable by a deterministic state estimation feedback. In Section 2 these problems are studied under the assumption that (0.1) possesses an exponential dichotomy.

In the remainder of this chapter I turn to the problem of robustness of stability of (0.1). For time-invariant linear systems there exist two fundamental approaches concerning robutsness. One is the successful H^∞-approach introduced by *Zames* (1981) and *Francis and Zames* (1983) which is based on transform techniques, see *Doyle and Stein* (1981), *Postlethwaite, Edmunds and Mac Farlane* (1981), *Kwakernaak* (1984). It is not clear how to extend these techniques to the time-varying case. Recently *Hinrichsen and Pritchard* (1986 a,b) have proposed a state space approach to robustness. Their problem is closely related to a well-known problem of perturbation theory: Determine a bound for all perturbation materices $\Delta(\cdot)$ such that exponential stability of (0.1) is preserved if the generator is additively disturbed by $\Delta(\cdot)$:

$$\dot{x}(t) = [A(t) + \Delta(t)]x(t) \qquad , t \geq 0 \tag{0.3}$$

See for instance *Bohl* (1913), *Perron* (1930), *Hahn* (1967), *Daleckii and Krein* (1974), *Coppel* (1978). However these bounds are conservative. *Hinrichsen and Pritchard* (1986 a,b) - in the constant case - were interested in a ***sharp*** upper bound, that is the (complex) ***stability radius***

$$r_{\mathbb{C}}(A) \;=\; inf\{\| \Delta \|_{L_\infty} |\ (0.3) \text{ is not exponentially stable}\} \tag{0.4}$$

or, if (0.1) is subjected to structured perturbations of the form $\Delta = B\,D\,C$ (B, C are known scaling matrices), the ***structured*** (complex) ***stability radius***

$$r_{\mathbb{C}}(A; B, C) \;=\; inf\{\| D \|_{L_\infty} |\ D \in \mathbb{C}^{m\times p} \text{ and } (0.3) \text{ is not exponentially stable}\} \tag{0.5}$$

Hinrichsen and Pritchard (1986 b) proved the one-to-one correspondence between $r_{\mathbb{C}}(A; B, C)$, the norm of the *perturbation operator*

$$L : u(\cdot) \mapsto \int_0^{\cdot} C\, e^{A(\cdot - s)} B\; u(s) ds \tag{0.6}$$

and the solvability of the parameterized ***algebraic*** Riccati equation

$$A^*P + PA - \rho\, C^*C - P\,B\,B^*\,P = 0, \qquad \rho \in \mathbb{R} \tag{0.7}$$

These methods and results were partially generalized for time-varying systems by *Hinrichsen, Ilchmann and Pritchard* (1987). This is presented here in Sections 3 to 8. We proceed as follows.

In Section 3 the group of Bohl transformations, containing Lyapunov transformations as a subgroup, is introduced. A Bohl transformation applied to (0.1) as a similarity action does not change the Bohl exponent.
In Section 4 the structural stability radius for time-varying systems is defined analogously to (0.5). Its invariance properties are discussed.
A generalization of the perturbation operator (0.6)

$$L_{t_0} : u(\cdot) \mapsto \int_0^{\cdot} C(\cdot)\Phi(\cdot, s)B(s)u(s)ds \tag{0.8}$$

is studied in Section 5. Its relationship to the structured stability radius is partly clarified. However an open problem remains.
Instead of the algebraic Riccati equation (0.7), in the time-varying setting one has to study the parametrized *differential Riccati equation*

$$\dot{P}(t) + A^*(t)P(t) + P(t)A(t) - \rho\, C^*(t)C(t) - P(t)B(t)B^*(t)P(t) = 0, \;\; \rho \in \mathbb{R} \tag{0.9}$$

In Section 6 a precise relationsship between the norm of the perturbation operator (0.8) and the solvability of (0.9) is established. This result yields a complete generalization of the situation in the time-invariant case.
In Section 7 the dependence of the maximal bounded Hermitian solution of (0.9) on the parameter ρ is analysed.
In Section 8 the robustness analysis is extended to nonlinear perturbations and a common Lyapunov function for a class of perturbed systems (0.3) is determined.

4.1 Sufficient conditions for exponential stability

Consider the homogeneous linear time-varying differential equation

$$\dot{x}(t) = A(t)x(t), \qquad A(\cdot) \in PC(\mathbb{R}_+,\ \mathbb{C}^{n\times n}) \tag{1.1}$$

where $PC(\mathbb{R}_+,\ \mathbb{C}^{n\times n})$ denotes the set of piecewise continuous complex $n \times n$ matrix functions on $\mathbb{R}_+ = [0, \infty)$.

Often we will assume that $A(\cdot)$ is bounded or more generally that $A(\cdot)$ is *integrally bounded*, i.e.

$$\sup_{t\in\mathbb{R}_+} \int_t^{t+1} \| A(s) \| \, ds < \infty$$

Let $< \cdot, \cdot >$ be the usual inner product on $\mathbb{C}^n$, $n \geq 1$, $\| \cdot \|$ the associated norm and $\| B \|$ the induced operator norm for any linear operator $B \in \mathcal{L}(\ \mathbb{C}^m,\ \mathbb{C}^n)$.

Definition 1.1 The system (1.1) is said to be *exponentially stable* if there exist $M, \omega > 0$ such that

$$\| \Phi(t, t_0) \| \leq M\ e^{-\omega(t-t_0)} \qquad \text{for all } t \geq t_0 \geq 0$$

("for all $t \geq t_0 \geq 0$" means "for all $t \geq t_0$ and all $t_0 \geq 0$")

Due to the linearity of (1.1) exponential stability can be characterized as follows, cf. *Willems* (1970) p. 101.

Proposition 1.2 The system (2.1) is exponentially stable if and only if it is *uniformly asymptotically stable*, i.e. there exists k independent of t_0 such that

$$\| \Phi(t, t_0) \| \leq k \qquad \text{for all } t \geq t_0 \geq 0 \tag{1.2}$$

and

$$\lim_{t\to\infty} \| \Phi(t, t_0) \| = 0 \quad \text{uniformly in } t_0 \in \mathbb{R}_+ \tag{1.3}$$

If $A(\cdot)$ is a constant matrix it is well-known that (1.1) is exponentially stable iff the real parts of the eigenvalues of A are lying in the open left half plane. For time-varying systems, even if they are analytic and periodic, exponential stability does neither imply

$$Re\ \sigma(A(t)) \ \subset\ \mathbb{C}^- = \{s \in\ \mathbb{C} \mid Re\ s < 0\} \qquad \text{for all } t \in \mathbb{R}_+$$

nor does for some $\alpha > 0$ the condition

$$Re\ \sigma(A(t)) < -\alpha \quad \text{for all } t \in \mathbb{R}_+$$

guarantee exponential stability.

Example 1.3 (i) *Hoppenstaedt* (1966), p. 3: Let

$$A(t) = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \begin{bmatrix} -1 & -5 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$$

Then $\sigma(A(t)) = \{-1\}$ for all $t \in \mathbb{R}_+$ and it can be easily verified that a fundamental matrix is given by

$$X(t) = \begin{bmatrix} e^{t}(\cos t + \frac{1}{2}\sin t), & e^{-3t}(\cos t - \frac{1}{2}\sin t) \\ e^{t}(\sin t - \frac{1}{2}\cos t), & e^{-3t}(\sin t + \frac{1}{2}\cos t) \end{bmatrix}$$

Thus $\dot{x}(t) = A(t)x(t)$ is not exponentially stable.

(ii) *Wu* (1974) : Let

$$A(t) = \begin{bmatrix} -\frac{11}{2} + \frac{15}{2}\sin 12t & \frac{15}{2}\cos 12t \\ \frac{15}{2}\cos 12 & -\frac{11}{2} - \frac{15}{2}\sin 12t \end{bmatrix}$$

Then $\sigma(A(t)) = \{2, -13\}$ for all $t \in \mathbb{R}_+$ and a fundamental matrix is given by

$$X(t) =$$

$$\begin{bmatrix} \frac{1}{2}e^{-t}(\cos 6t + 3\sin 6t) + \frac{1}{2}e^{-10t}(\cos 6t - 3\sin 6t) & \frac{1}{6}e^{-t}(\cos 6t + 3\sin 6t) - \frac{1}{6}e^{-10t}(\cos 6t - 3\sin 6t) \\ \frac{1}{2}e^{-t}(3\cos 6t - \sin 6t) - \frac{1}{2}e^{-10t}(3\cos 6t + \sin 6t) & \frac{1}{6}e^{-t}(3\cos 6t - \sin 6t) + \frac{1}{6}e^{-10t}(3\cos 6t + \sin 6t) \end{bmatrix}$$

Since $\Phi(\cdot,\cdot)$ satisfies (1.2) and (1.3) the system (1.1) is exponentially stable by Proposition 1.2.

The system presented in Example 1.3 (i) is in some sense "too fast" in order that condition $\Re\sigma(A(t)) \leq -1$ implies exponential stability. Various assumptions on the parameter variation of $A(\cdot)$ are known, such that if $\delta > 0$ is sufficiently small then anyone of the following conditions guarantees exponential stability of (1.1):

$$\| \dot{A}(t) \| \leq \delta \quad \text{for all } t \geq 0 \quad (\textit{Rosenbrock}\ (1963)) \tag{1.4}$$

$$\| A(t_2) - A(t_1) \| \leq \delta \| t_2 - t_1 \| \qquad \text{for all } t_1, t_2 \geq 0 \quad (\textit{Coppel}\ (1978),\ \text{p.5}) \tag{1.5}$$

$$\sup_{0 \leq \tau \leq h} \| A(t+\tau) - A(t) \| \leq \delta \tag{1.6}$$

$$\left.\begin{array}{l} \dot{A}(\cdot) \text{ is continuous, } \| \dot{A}(\cdot) \| \text{ is uniformly bounded} \\ \text{and there exists } T > 0 \text{ such that} \\ \int_{t_0}^{t_0+T} \| \dot{A}(t) \| \, dt \leq \delta \cdot T \quad \text{for all } t_0 \geq 0. \end{array}\right\} \tag{1.7}$$

As a consequence of the following Proposition 1.4 (iii), (1.6) implies exponential stability if δ is small enough. (1.6) is less restrictive than a similar condition in *Kreisselmeier* (1985), Lemma 3:

$$\lim_{t \to \infty} \sup_{0 \leq \tau \leq h} \| A(t+\tau) - A(t) \| = 0 \qquad \text{for all } h > 0$$

Furthermore (1.7) is less restrictive than the criterium in *Krause and Kumar* (1986) which is

$$\left.\begin{array}{l} \text{there exists a } T_0 < \infty \text{ such that} \\ \\ \int_{t_0}^{t_0+T} \| \dot{A}(t) \| \, dt \leq \delta T \quad \text{for all } t_0 \geq 0, T \geq T_0 \end{array}\right\} \tag{1.8}$$

Krause and Kumar (1986) present a very lengthy proof to show that (1.8) implies exponential stability of (1.1) if δ is sufficiently small and moreover $A(\cdot)$ satisfies:

$$\left.\begin{array}{l} Re\ \sigma(A(t)) < -\alpha \text{ for some } \alpha > 0, \quad \text{for all } t \in \mathbb{R}_+ \\ \\ \| A(\cdot) \| \text{ and } \| \dot{A}(\cdot) \| \text{ are bounded} \\ \\ \dot{A}(\cdot) \text{ is continuous} \end{array}\right\} \tag{1.9}$$

In *Ilchmann, Owens and Prätzel-Wolters* (1987b) we proved in a short way that if the weakened Krause and Kumar condition (1.7) is assumed, then exponential stability can be derived by a slightly modification of *Rosenbrock's* (1963) proof.

The disadvantage of (1.4) - (1.7) is that they are qualitative conditions in the sense that δ must be small enough. We can improve the results and give quantitative bounds.

Proposition 1.4 Suppose $A(\cdot) \in PC(\mathbb{R}_+,\ \mathbb{C}^{n\times n})$ satisfies for some $\alpha, M > 0$ and all $t \geq 0$

$$\begin{aligned} \| A(t) \| &\leq M \\ \sigma(A(t)) &\subset \mathbb{C}^{-\alpha} = \{s \in \mathbb{C} \mid Re\ s < -\alpha\} \end{aligned}$$

Then the system

$$\dot{x}(t) = A(t)x(t), \qquad t \geq 0 \tag{1.10}$$

is exponentially stable if one of the following conditions holds true for all $t \geq 0$:

(i) $-\alpha < -4M$

(ii) $A(\cdot)$ is piecewise differentiable and

$$\| \dot{A}(t) \| \leq \delta < \frac{2}{2n-1} \cdot \frac{\alpha^{4n-2}}{2M^{4n-4}}$$

(iii) For some $k > 0, \eta \in (0,1)$ and $\alpha > 2M\eta + \frac{n-1}{k}\log\eta$

$$\sup_{0\leq\tau\leq k} \| A(t+\tau) - A(t) \| \leq \delta < \eta^{n-1}(\alpha - 2M\eta + \frac{n-1}{k}\log\eta)$$

(iv) $\alpha > n-1$ and for some $\eta \in (0.1)$

$$\sup_{h>0} \| \frac{A(t+h) - A(t)}{h} \| \leq \delta < 2\eta^{n-1}(\alpha - 2M\eta + (n-1)\log\eta)$$

Proof: We will use the following important inequality due to *Coppel* (1978):

$$\| e^{A(t)\sigma} \| \leq (\frac{2M}{\varepsilon})^{n-1} e^{(-\alpha+\varepsilon)\sigma} \quad \text{for all } \sigma, t \geq 0 \text{ and for all } \varepsilon \in (0, 2M) \tag{1.11}$$

For fixed $t_0 \in \mathbb{R}_+$ (1.10) can be rewritten in the form

$$\dot{x}(t) = A(t_0)x(t) + [A(t) - A(t_0)]x(t), \quad t \geq 0$$

and for $x(t_0) = x_0 \in \mathbb{R}^n$ its solution is given by

$$x(t) = e^{A(t_0)(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t_0)(t-s)}[A(s) - A(t_0)]x(s)ds$$

Hence by (1.11)

$$\begin{aligned} \| x(t) \| &\le \kappa_\varepsilon e^{(-\alpha+\varepsilon)(t-t_0)} \| x_0 \| \\ &+\kappa_\varepsilon \int_{t_0}^t e^{(-\alpha+\varepsilon)(t-s)} \| A(s) - A(t_0) \| \| x(s) \| ds \qquad \text{for all } t \ge t_0 \end{aligned}$$

where

$$\kappa_\varepsilon := (\frac{2M}{\varepsilon})^{n-1}$$

Multiplying this inequality by $e^{(\alpha-\varepsilon)t}$ and applying Gronwall's Lemma yields

$$\| e^{(\alpha-\varepsilon)t} x(t) \| \le \kappa_\varepsilon \, e^{(\alpha-\varepsilon)t_0} \| x_0 \| \cdot exp[\kappa_\varepsilon \int_{t_0}^t \| A(s) - A(t_0) \| ds]$$

Thus

$$\| x(t) \| \le \kappa_\varepsilon \, exp[(-\alpha+\varepsilon)(t-t_0) + \kappa_\varepsilon \int_{t_0}^t \| A(s) - A(t_0) \| ds] \| x_0 \| \quad \text{for all } t \ge t_0 \tag{1.12}$$

Now we prove the statements (i) - (iv).

(i): Since $\| A(s) - A(t_0) \| \le 2M$ for all $s, t_0 \ge 0$, (1.12) implies for $\varepsilon \in (0, 2M)$ and some $h > 0$:

$$\begin{aligned} \| x(t) \| &\le \kappa_\varepsilon e^{[-\alpha+\varepsilon+\kappa_\varepsilon 2M](t-t_0)} \| x(t_0) \| \\ &\le \kappa_\varepsilon \, e^{[\varepsilon+\kappa_\varepsilon 2M-4M-h](t-t_0)} \| x(t_0) \| . \end{aligned}$$

The function

$$\begin{aligned} f \ : \ (0, 2M] &\to R \\ \varepsilon &\mapsto \varepsilon + \kappa_\varepsilon 2M - 4M - h \end{aligned}$$

is continuous and $f(2M) = -h$. Thus there exists $\varepsilon \in (0, 2M]$ such that $f(\varepsilon) < 0$.

(ii): Consider

$$R(t) := \int_0^\infty e^{A^T(t)s} e^{A(t)s} ds \tag{1.13}$$

which solves

$$R(t)A(t) + A^T(t)R(t) = -I_n$$

and satisfies for some $c_1, c_2 > 0$

$$0 < c_1 I_n \le R(t) \le c_2 I_n \qquad \text{for all } t \ge 0.$$

The derivative of $R(\cdot)$ is given by

$$\dot{R}(t) = \int_0^\infty e^{A^T(t)s} [R(t)\dot{A}(t) + \dot{A}^T R(t)] e^{A(t)s} ds \tag{1.14}$$

(cf. *Brockett* (1970) pp. 203 and 206) . Now we show that

$$V(x, t) := x^T R(t) x$$

is a Lyapunov function of $\dot{x}(t) = A(t)x(t)$. Its time derivative along any solution is:

$$\frac{d}{dt} V(x(t), t) \ = \ x^T(t)[-I_n + \dot{R}(t)]x(t).$$

We have to show that

$$\dot{R}(t) < I_n \qquad \text{for all } t \ge 0. \tag{1.15}$$

Applying Coppel's inequality to (1.14) and (1.13) yields

$$\begin{aligned}
\| \dot{R}(t) \| &\leq \int_0^\infty (\frac{2M}{\varepsilon})^{2(n-1)} e^{2(-\alpha+\varepsilon)s} ds \cdot 2 \| R(t) \| \| \dot{A}(t) \| \\
&\leq 2[(\frac{2M}{\varepsilon})^{2(n-1)} \int_0^\infty e^{2(-\alpha+\varepsilon)s} ds]^2 \cdot \delta \\
&= 2(\frac{2M}{\varepsilon})^{4(n-1)} (\frac{1}{2(-\alpha+\varepsilon)})^2 \delta
\end{aligned}$$

and thus (1.15) holds if for some $\varepsilon \in (0,\alpha)$

$$\delta < 2 \cdot (\frac{\varepsilon}{2M})^{4(n-1)} (\alpha - \varepsilon)^2 =: g(\varepsilon).$$

It is easily verified, that $g(\cdot)$ achieves its maximum on $(0,\alpha)$ at $\varepsilon_0 = \frac{k}{k+1}\alpha$ and

$$g(\varepsilon_0) = \frac{2\alpha^{4n-2}}{(2M)^{4(n-1)} \cdot (n-1)}.$$

(iii): By (1.12) we have for every $\ell \in \mathbb{N}$, $k > 0$ and $t \in [t_0 + \ell k, t_0 + (\ell+1)k]$:

$$\begin{aligned}
\| x(t) \| &= \kappa_\varepsilon \, e^{\gamma(t-t_0-\ell k)} \| x(t_0 + \ell k) \| \\
&\leq \kappa_\varepsilon \, e^{\gamma(t-t_0-\ell k)} \cdot \kappa_\varepsilon \, e^{\gamma k} \| x(t_0 + (\ell-1)k) \| \\
&= \kappa_\varepsilon^2 \, e^{\gamma(t-t_0-(\ell-1)k)} \| x(t_0 + (\ell-1)k) \| \\
&\vdots \\
&\leq \kappa_\varepsilon^{\ell+1} \, e^{\gamma(t-t_0)} \| x(t_0) \|
\end{aligned}$$

where $\gamma := -\alpha + \varepsilon + \kappa_\varepsilon \cdot \delta$.
Thus

$$\| x(t) \| \leq \kappa_\varepsilon \, e^{\log \kappa_\varepsilon + \gamma(t-t_0)} \| x(t_0) \| .$$

and since $(t - t_0) \geq \ell \cdot k$

$$\| x(t) \| \leq \kappa_\varepsilon \, e^{(\frac{\log \kappa_\varepsilon}{k} + \gamma)(t-t_0)} \| x(t_0) \| .$$

It remains to determine $\varepsilon < 2M$ and $k > 0$ such that

$$\frac{\log \kappa_\varepsilon}{k} + \gamma < 0$$

which is equivalent to

$$0 < \delta < \frac{1}{\kappa_\varepsilon}(\alpha - \varepsilon - \frac{\log \kappa_\varepsilon}{k})$$

However, for every $\varepsilon \in (0,\alpha)$ there exists $k > 0$ such that

$$\alpha - \varepsilon - \frac{\log \kappa_\varepsilon}{k} > 0$$

and thus (1.10) is exponentially stable for every $A(\cdot)$ which satisfies

$$\sup_{0 \leq \tau \leq k^*} \| A(t+\tau) - A(t) \| \leq \delta < \frac{1}{\kappa_\varepsilon}(\alpha - \varepsilon - \frac{\log \kappa_\varepsilon}{k})$$

Now (iii) follows with $\eta = \frac{\varepsilon}{2M}$.

(iv): Assume

$$\| \frac{A(t+h) - A(t)}{h} \| \leq \delta$$

for every $h > 0$. Then by (1.12) we have

$$\begin{aligned} \| x(t) \| &\leq \kappa_\varepsilon \, e^{(-\alpha+\varepsilon)(t-t_0)+\kappa_\varepsilon \int_0^{t-t_0} h \cdot \delta dh} \| x(t_0) \| \\ &\leq \kappa_\varepsilon \, e^{(\varepsilon-\alpha+\kappa_\varepsilon \frac{t-t_0}{2} \cdot \delta)(t-t_0)} \| x(t_0) \| \,. \end{aligned}$$

and

$$\| x(t) \| \leq \kappa_\varepsilon \, e^{\gamma(t-t_0)} \| x(t_0) \| \qquad \text{for } t \in [t_0, t_0+1]$$

where $\gamma = \varepsilon - \alpha + \frac{1}{2}\,\kappa_\varepsilon \delta$.
If $t \in [t_0 + \ell, t_0 + \ell + 1]$, $\ell \in \mathbb{N}$ we conclude as in the proof of (iii)

$$\| x(t) \| \leq \kappa_\varepsilon \, e^{(\log \kappa_\varepsilon + \gamma)(t-t_0)} \| x(t_0) \|$$

Now $\log \kappa_\varepsilon + \gamma < 0$ if

$$\delta < 2(\frac{\varepsilon}{2M})^{n-1}(\alpha - \varepsilon - \log(\frac{2M}{\varepsilon})^{n-1})$$

and (iv) follows with $\eta = \frac{\varepsilon}{2M}$. □

Note that the proof of (iii) presents a short proof of Lemma 3 in *Kreisselmeier* (1985).

If additional information on the exponential decay of $e^{A(t)\tau}$ is known the bounds in Theorem 1.4 can be simplified as follows.

Proposition 1.5 Suppose $A(\cdot) \in PC(\mathbb{R}_+, \mathbb{C}^{n\times n})$ satisfies for some $M, K, \omega > 0$ and all $t \geq 0$

$$\begin{aligned} \| A(t) \| &\leq M \\ \| e^{A(t)\sigma} \| &\leq K \, e^{-\omega\sigma} \end{aligned}$$

Then the system

$$\dot{x}(t) = A(t)x(t), \qquad t \geq 0$$

is exponentially stable if one of the following conditions holds true for all $t \geq 0$:

(i) $M < \frac{1}{2}\frac{\omega}{K}$.

(ii) $A(\cdot)$ is piecewise differentiable and $\| \dot{A}(t) \| \leq \delta < 2(\frac{\omega}{K})^2$.

(iii) There exists $h > 0$ such that

$$\sup_{0 \leq \tau \leq h} \| A(t+\tau) - A(t) \| \leq \delta < \frac{\omega}{K} - \frac{1}{h}\frac{\log K}{K}.$$

(iv) $\sup_{h>0} \| \frac{A(t+h)-A(t)}{h} \| \leq \delta < 2\frac{\omega}{K} - 2\frac{\log K}{K}$.

Proof: (i): Since $\| A(s) - A(t_0) \| \leq 2M$, (1.12) implies

$$\| x(t) \| \leq K \, e^{-\omega(t-t_0)+K\,2M(t-t_0)} \| x_0 \|$$

which proves the statement.

(ii): Similarly to the proof of Proposition 1.4 (ii) one obtains

$$\begin{aligned}
\| \dot{R}(t) \| &\leq \int_0^\infty K^2 e^{-2\omega s} ds \cdot 2 \| R(t) \| \| \dot{A}(t) \| \\
&\leq 2[K^2 \int_0^\infty e^{-2\omega s} ds]^2 \cdot \delta \\
&= \frac{K^4}{2\omega^2}\delta
\end{aligned}$$

Since by assumption $\varepsilon := \frac{K^2}{2\omega^2}\delta < 1$ one concludes

$$\dot{V}(x) = x^*[A\,R + R\,A]x + x^*\dot{R}x = x^*[-I_n + \dot{R}]x < (\varepsilon - 1) \| x \|^2$$

This completes the proof.

(iii): For $t \in [t_0, t_0 + h]$, (1.12) implies

$$\| x(t) \| \leq K\, e^{(-\omega + K\delta)(t - t_0)} \| x(t_0) \|$$

Hence for $\ell \in \mathbb{N}$ and $t \in [t_0 + \ell h, t_0 + (\ell + 1)h]$

$$\begin{aligned}
\| x(t) \| &\leq K\, e^{\gamma(t - t_0 - \ell h)} \| x(t_0 + \ell h) \| \\
&\leq K\, e^{\gamma(t - t_0 - \ell h)} \cdot K\, e^{\gamma h} \| x(t_0 + (\ell - 1)h) \| \\
&= K^2\, e^{\gamma(t - t_0 - (\ell - 1)h)} \| x(t_0 + (\ell - 1)h) \| \\
&\vdots \\
&\leq K^{\ell + 1}\, e^{\gamma(t - t_0)} \| x(t_0) \|
\end{aligned}$$

where $\gamma = -\omega + K\delta$. Since $(t - t_0) \geq \ell \cdot h$

$$\begin{aligned}
\| x(t) \| &\leq K\, e^{\ell \cdot \log K + \gamma(t - t_0)} \| (t_0) \| \\
&\leq K\, e^{(\frac{\log K}{h} + \gamma)(t - t_0)} \| x(t_0) \|
\end{aligned}$$

This proves (iii).

(iv): If $t \in [t_0, t_0 + 1]$ then (1.12) yields

$$\begin{aligned}
\| x(t) \| &\leq K\, exp[-\omega(t - t_0) + K \int_0^{t - t_0} \delta h\, dh] \| x(t_0) \| \\
&\leq K\, e^{(-\omega + K\frac{\delta}{2})(t - t_0)} \| x(t_0) \|
\end{aligned}$$

For $t \in [t_0 + \ell, t_0 + (\ell + 1)]$, $\ell \in \mathbb{N}$ we conclude as in (iii)

$$\begin{aligned}
\| x(t) \| &\leq K^{\ell + 1}\, e^{(-\omega + K\frac{\delta}{2})(t - t_0)} \| x(t_0) \| \\
&\leq K\, e^{(\log K - \omega + K\frac{\delta}{2})(t - t_0)} \| x(t_0) \|
\end{aligned}$$

which proves (iv). □

The previous propositions show that there is an interplay between the bound of the real parts of the eigenvalues of $A(t)$ and the parameter variation of $A(\cdot)$. However, for these sufficient condition the assumption that $A(\cdot)$ is bounded is essential. Only few results are known to have sufficient conditions for exponential stability if $A(\cdot)$ is unbounded. In the remainder of this section we will present some results where $A(\cdot)$ is not bounded but of the special structure $A(t) = A - k(t)D$ where $k(\cdot)$ is a scalar function.

Proposition 1.6 Suppose $A, D \in \mathbb{C}^{n\times n}$ and $k(\cdot) \in PC(\mathbb{R}_+, \mathbb{R}_+)$ is monotonically nondecreasing with $\lim_{t\to\infty} k(t) = \infty$. Then the following statements are equivalent:

(i) $\sigma(D) \subset \mathbb{C}_+ = \{s \in \mathbb{C} \mid Re\, s > 0\}$

(ii) $\lim_{t\to\infty} Re\, \lambda_i(t) = -\infty$ for $i \in \underline{n}$
where $\lambda_i(t)$, $i \in \underline{n}$, denote the eigenvalues of $A - k(t)D$.

(iii) the system

$$\dot{x}(t) = (A - k(t)D)x(t), \quad t \geq 0 \tag{1.16}$$

is *arbitrary fast exponentially stable,* i.e. for some $\omega(\cdot) \in PC(\mathbb{R}_+, \mathbb{R}_+)$ monotonically non decreasing with $\lim_{t\to\infty} \omega(t) = \infty$ and some $M > 0$ the transition matrix of (1.16) satisfies

$$\| \Phi(t, t_0) \| \leq M\, e^{-\omega(t_0)(t-t_0)} \quad \text{for all } t \geq t_0 \geq 0.$$

Proof: (i) $\Rightarrow$ (ii): Select $k_1 > 0$ and $T \in GL_n(\mathbb{C})$ such that

$$T(k_1 D)T^{-1} = \begin{bmatrix} \lambda_1 & \delta_1 & & & \\ & & & \ddots & 0 \\ & & \ddots & & \\ 0 & & & & \delta_{n-1} \\ & & & & \lambda_n \end{bmatrix}$$

and $Re\, \lambda_i > 1, \delta_i \in \{0, 1\}$ for $i \in \underline{n}$. The spectrum of $A - k(t)D$ and

$$\hat{A} + \hat{D}(t) := T\,A\,T^{-1} + (-\frac{k(t)}{k_1})T\,k_1 D\,T^{-1}$$

coincide and by Gerschgorin's Circle Theorem (see e.g. *Noble and Daniel* (1977)) we have

$$\sigma(\hat{A} + \hat{D}(t)) \subset \bigcup_{i=1}^{n} C_i(t)$$

where

$$C_i(t) := \{\mu(t) \in \mathbb{C} : |-\frac{k(t)}{k_1}\lambda_i + \hat{a}_{ii} - \mu(t)| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |\hat{a}_{ij}| + \frac{k(t)}{k_1}\delta_i\}, \quad i \in \underline{n}.$$

Since $Re\, \lambda_i > 1$ we conclude

$$\lim_{t\to\infty} Re\, \mu(t) = -\infty$$

for every $\mu(t) \in C_i(t)$, whence (ii) follows.

(ii) $\Rightarrow$ (i): Assume there is an eigenvalue $\tilde{\lambda}_i$ of D with $Re\, \tilde{\lambda}_i \leq 0$. Then the real part of the corresponding eigenvalue $k(t)\tilde{\lambda}_i$ of $k(t)D$ either remains 0 or tends to $-\infty$ as $t \to \infty$. Thus by Gerschgorin's Theorem there exists $t^* > 0$ such that at least for one $\lambda_i(k(t))$ we have $Re\, \lambda_i(k(t)) \geq N > -\infty$ for all $t > t^*$. This contradicts (ii).

(i) $\Rightarrow$ (iii): Let $P = P^T > 0$ and $Q = Q^T > 0$ such that $D^T P + P\,D = 0$. It is proved that $V(x(t)) := x(t)^T P x(t)$ is a Lyapunov function for (1.16). Differentiation yields

$$\begin{aligned} \dot{V}(x(t)) &= x(t)^T[A' - (k(t) - k(t'))D]^T P x(t) + x(t)^T P[A' - (k(t) - k(t'))D]x(t) \\ &= x(t)^T[A'^T P + P\,A']x(t) - (k(t) - k(t'))x(t)^T Q x(t) \end{aligned}$$

where

$$A' := A - k(t')D.$$

Choose t' sufficiently big such that for some $p_1 > 0$

$$x^T[A'^T P + P A']x \leq -p_1 \| x \|^2 .$$

Since $q_1 \| x \|^2 \leq x^T Q x$ for some $q_1 > 0$ one obtains

$$\dot{V}(x(t)) \leq -[P_1 + (k(t) - k(t'))q_1] \| x(t) \|^2 \quad \text{for all } t \geq t'$$

This proves (iii).

(iii) $\Rightarrow$ (i): Since $\lim_{t_0 \to \infty} \omega(t_0) = \infty$ one may assume that $A(\cdot) \equiv 0$, see a disturbance result in *Coppel* (1978) p.2.
Now (i) follows since if some eigenvalue $\tilde{\lambda}$ of D satisfies $Re\ \tilde{\lambda} \leq 0$ then $\dot{x}(t) = -k(t)Dx(t)$ is unstable. □

If the conditions on A and D are relaxed the system (1.16) is, in general, not arbitrary fast exponentially stable, however it is exponentially stable. To prove this result the following "interconnections lemma" is needed.

Lemma 1.7 (Interconnections lemma)
Consider the interconnected system

$$\dot{x}(t) = \begin{bmatrix} A_1(t) & A_2(t) \\ A_3(t) & A_4(t) \end{bmatrix} x(t), \qquad t \geq 0 \tag{1.17}$$

where A_1, A_2, A_3, A_4 are piecewise continuous complex $\ell \times \ell$, $\ell \times (n-\ell)$,$(n-\ell) \times \ell$, $(n-\ell) \times (n-\ell)$, resp. matrix functions and

$$\| A_2(t) \| \leq a_2, \quad \| A_3(t) \| \leq a_3 \quad \text{for all } t \in \mathbb{R}_+ \text{ and some } a_2, a_3 > 0.$$

Moreover we assume for the transition matrices generated by A_1 and A_4

$$\begin{aligned} \| \Phi_1(t,s) \| &\leq M_1 \, e^{-\lambda(t-s)} \\ \| \Phi_4(t,s) \| &\leq M_4 \, e^{-\varepsilon(t-s)} \quad \text{for } t \geq s \geq 0 \end{aligned}$$

resp., for some $M_1, M_4, \lambda, \varepsilon > 0$. Then if $\lambda > \varepsilon + \varepsilon^{-1}(M_1 M_4 a_2 a_3)$ the interconnected system is exponentially stable with decay $-h = -\varepsilon - (\varepsilon - \lambda)^{-1}(M_1 M_4 a_2 a_3)$.

Proof: Let $x_1(t) \in \mathbb{C}^\ell, x_2(t) \in \mathbb{C}^{n-\ell}$, then the following integral equations are equivalent to (1.17)

$$\begin{aligned} x_1(t) &= \Phi_1(t,t_0)x_1(t_0) + \textstyle\int_{t_0}^t \Phi_1(t,s)A_2(s)x_2(s)ds \\ x_2(t) &= \Phi_4(t,t_0)x_2(t_0) + \textstyle\int_{t_0}^t \Phi_4(t,s)A_3(s)x_1(s)ds \end{aligned}$$

Now

$$\begin{aligned} \| x_2(t) \| \leq \; & M_4 \, e^{-\varepsilon(t-t_0)} \| x_2(t_0) \| + \int_{t_0}^t M_4 \, e^{-\varepsilon(t-s)} a_3 [M_1 \, e^{-\lambda(s-t_0)} \\ & \| x_1(t_0) \| + \int_{t_0}^s M_1 \, e^{-\lambda(s-\tau)} a_2 \| x_2(\tau) \| d\tau] ds \end{aligned}$$

and

$$\| e^{\varepsilon t} x_2(t) \| \leq M_4\, e^{\varepsilon t_0} \| x_2(t_0) \| \; + \; \frac{M_4 a_3 M_1}{\delta}[e^{\delta t + \lambda t_0} - e^{\varepsilon t_0}] \| x_1(t_0) \|$$
$$+ \; K \int_{t_0}^{t} e^{\delta s} \int_{t_0}^{s} e^{\lambda \tau} \| x_2(\tau) \| \, d\tau \, ds$$

where $\delta := \varepsilon - \lambda < 0$ by assumption, $K := M_1 M_4 a_2 a_3$.
Integration by parts gives

$$\int_{t_0}^{t} e^{\delta s} \int_{t_0}^{s} e^{\lambda \tau} \| x_2(\tau) \| \, d\tau \, ds$$
$$= \delta^{-1} e^{\delta s} \textstyle\int_{t_0}^{s} e^{\lambda \tau} \| x_2(\tau) \| \, d\tau \, |_{t_0}^{t} \; - \; \int_{t_0}^{t} \delta^{-1} e^{\delta s} \cdot e^{\lambda s} \| x_2(s) \| \, ds$$
$$= \delta^{-1} e^{\delta t} \textstyle\int_{t_0}^{t} e^{\lambda \tau} \| x_2(\tau) \| \, d\tau \; - \; \int_{t_0}^{t} \delta^{-1} e^{\varepsilon s} \| x_2(s) \| \, ds$$

Applying this to the above inequality yields

$$\| e^{\varepsilon t} x_2(t) \| \leq e^{\varepsilon t_0}[M_4 \| x_2(t_0) \| - \frac{M_1 M_4 a_3}{\delta} \| x_1(t_0) \|] - \frac{K}{\delta} \int_{t_0}^{t} e^{\varepsilon s} \| x_2(s) \| \, ds$$

and by Gronwall's inequality one obtains

$$\| x_2(t) \| \leq [M_2 \| x_2(t_0) \| - \delta^{-1} M_1 M_4 a_3 \| x_1(t_0) \|] e^{-\varepsilon(t-t_0)} \, e^{-\frac{K}{\delta}(t-t_0)}$$

Setting $\hat{K} = max\ \{M_2, -\delta^{-1} M_1 M_4 a_3\}$ and choosing the maximum norm we have

$$\| x_2(t) \| \leq \hat{K}\, e^{-h(t-t_0)} \| x(t_0) \| \qquad \text{for } t \geq t_0$$

Consider now $x_1(t)$, the solution of

$$\dot{x}_1(t) = A_1(t) x_1(t) + A_2(t) x_2(t)$$

which satisfies

$$\| x_1(t) \| \leq M_1\, e^{-\lambda(t-t_0)} \| x_1(t_0) \| + \int_{t_0}^{t} M_1\, e^{-\lambda(t-s)} a_2\, \hat{K}\, e^{-h(s-t_0)} \| x(t_0) \| \, ds$$

Thus

$$\| e^{\lambda t} x_1(t) \| \leq M_1\, e^{\lambda t_0} \| x_1(t_0) \| + \frac{M_1 a_2 \hat{K}}{\lambda - h}[e^{(\lambda - h)t + h t_0} - e^{\lambda t_0}] \| x(t_0) \|$$

and

$$\| x_1(t) \| \leq [M_1 \| x_1(t_0) \| - \frac{M_1 a_2 \hat{K}}{\lambda - h} \| x(t_0) \|] e^{-\lambda(t-t_0)} + \frac{M_1 a_2 \hat{K}}{\lambda - h} \| x(t_0) \| \, e^{-h(t-t_0)}$$

Finally $\lambda > h$ gives

$$\| x_1(t) \| < \tilde{K}\, e^{-h(t-t_0)} \| x(t_0) \|$$

for some $\tilde{K} > 0$ and this completes the proof. □

Proposition 1.8 (High gain feedback)
Suppose $A, D \in \mathbb{C}^{n \times n}$ satisfy

(i) there exist $k^*, \varepsilon > 0$ such that

$$\sigma(A - kD) \subset \mathbb{C}^{-\varepsilon} \text{ for all } k \geq k^*$$

(ii) if $0 \in \sigma(D)$ then 0 is *semisimple*, i.e. all corresponding blocks in the Jordan canonical form are of size 1×1.

Then the system

$$\dot{x}(t) = [A - k(t)D]x(t), \qquad t \geq 0$$

is exponentially stable for every piecewise continuous $k(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ with $\lim_{t\to\infty} k(t) = \infty$.

Proof: The invariance of exponential stability with respect to constant coordinate transformations together with (ii) implies that we can assume D to be of the form

$$D = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix}, \text{ where } \Lambda = \begin{bmatrix} \lambda_1 & \delta_1 & & \\ & & \ddots & \\ & & \ddots & \delta_{\ell-1} \\ & & & \lambda_\ell \end{bmatrix}, \lambda_i \neq 0, \delta_i \in \{0,1\}$$

Then

$$A - k(t)D = \begin{bmatrix} A_1 - k(t)\Lambda & A_2 \\ A_3 & A_4 \end{bmatrix}$$

and choose $S \in Gl_{n-\ell}(\mathbb{C})$ such that

$$A_4^J := S\,A_4 S^{-1} = \begin{bmatrix} \lambda_{\ell+1} & \delta_{\ell+1} & & \\ & \ddots & & \\ & & \ddots & \\ & & & \delta_{n-1} \\ & & & \lambda_n \end{bmatrix}, \quad \delta_j \in \{0,1\}.$$

For $r > 0$ and $T = diag(\alpha_{\ell+1}, \ldots, \alpha_n) \in GL_{n-\ell}(\mathbb{R})$ one obtains

$$\begin{bmatrix} I_s & 0 \\ 0 & rST \end{bmatrix} [A - k(t)D] \begin{bmatrix} I_s & 0 \\ 0 & (rST)^{-1} \end{bmatrix} = \begin{bmatrix} A_1 - k(t)\Lambda & A_2(rST)^{-1} \\ rSTA_3 & TA_4^J T^{-1} \end{bmatrix}$$

where

$$TA_4^J T^{-1} = \begin{bmatrix} \lambda_{\ell+1} & \frac{\alpha_{\ell+1}}{\alpha_{\ell+2}} \cdot \delta_{\ell+1} & & \\ & & \ddots & \\ & & & \frac{\alpha_{n-1}}{\alpha_n} \cdot \delta_{n-1} \\ & \ddots & & \\ & & & \lambda_n \end{bmatrix}.$$

For r sufficiently small and α_i suitable chosen Gerschgorin's Theorem together with (i) implies

$$Re\ \lambda_i(k(t)) \to -\infty \text{ as } t \to \infty, \;\; i \in \underline{\ell}$$

and

$$Re\ \lambda_i < -\frac{3}{4}\varepsilon \qquad \text{for } i = \ell+1, \ldots, n$$

Now the result follows by Proposition 1.6 and Lemma 1.7. □

The previous proposition was also proved by *Mårtensson* (1986). However the proof given here, resp. in *Ilchmann, Owens and Prätzel-Wolters* (1987a), was found independently and the presentation is completely different.

4.2 Stabilizability of systems with exponential dichotomy

In this section it is asumed that the linear differential equation

$$\dot{x}(t) = A(t)x(t) \qquad , A(t) \in PC(\mathbb{R}_+, \mathbb{R}^{n\times n}) \tag{2.1}$$

possesses an *exponential dichotomy* , i.e. for some fundamental matrix $X(\cdot)$ of (2.1) there exist $K, L, \alpha, \beta > 0$ such that

$$\left.\begin{array}{lll} \| X(t)P_1x_0 \| & \leq \; K\, e^{-\alpha(t-s)} \, \| X(s)P_1x_0 \| & \text{for } t \geq s \\ \| X(t)P_2x_0 \| & \leq \; L\, e^{-\beta(s-t)} \, \| X(s)P_2x_0 \| & \text{for } s \geq t \\ \| X(t)P_1X^{-1}(t) \| & \leq \; M & \text{for } t \geq 0 \end{array}\right\} \tag{2.2}$$

where $P_1 \in \mathbb{C}^{n\times n}$ is a projection, i.e. $P_1^2 = P_1$, and $P_2 := I_n - P_1$. The first and the second inequality in (2.2) say that at each time the state space splits into the direct sum of two subspaces such that the free trajectories starting in one subspace are exponentially decaying whereas the trajectories starting in the other subspace are exponentially increasing. The third condition in (2.2) means that the angular distance between the subspaces $\mathcal{V}_1(t) = X(t)P_1\mathbb{R}^n$ and $\mathcal{V}_2(t) = X(t)P_2\mathbb{R}^n$ cannot become arbitrary small under a variation of t. More precisely, there exists some $c > 0$ such that

$$inf\,\{\| v_1(t) - v_2(t) \| \mid v_i(t) \in \mathcal{V}_i(t), \| v_i(t) \| = 1, i = 1,2\} \geq c \qquad \text{for all } t \geq 0$$

(This is proved in *Daleckii and Krein* (1974) p. 163)

Remark 2.1 The conditions (2.2) are equivalent to

$$\left.\begin{array}{lll} \| X(t)P_1X^{-1}(s) \| & \leq \; K'\, e^{-\alpha(t-s)} & \text{for } t \geq s \\ \| X(t)P_2X^{-1}(s) \| & \leq \; L'\, e^{-\beta(s-t)} & \text{for } s \geq t \end{array}\right\} \tag{2.3}$$

for some $K', L' > 0$ (see *Coppel* (1978) p.11).
Using this fact it is immediate from Remark 1.3.2 that the adjoint system of (2.1) possesses an exponential dichotomy of the form

$$\left.\begin{array}{lll} \| X^a(t)P_2^Tx_0 \| & \leq \; \tilde{L}\, e^{-\alpha(t-s)} \, \| X^a(s)P_2^Tx_0 \| & \text{for } t \geq s \\ \| X^a(t)P_1^Tx_0 \| & \leq \; \tilde{K}\, e^{-\beta(t-s)} \, \| X^a(s)P_1^Tx_0 \| & \text{for } s \geq t \\ \| X^a(t)P_1^TX^{a^{-1}}(t) \| & \leq \; M & \text{for } t \geq 0 \end{array}\right\} \tag{2.4}$$

for some $\tilde{L}, \tilde{K} > 0$.

In the following we will use the notation of Section 1.2 and split the vector space $X(\cdot)\mathbb{R}^n$ of free motions into

$$X(\cdot)\mathbb{R}^n = \mathcal{V}_1(\cdot) \oplus \mathcal{V}_2(\cdot)$$

and the projections associated with each subspace are given by

$$P_i(t) = X(t)P_iX^{-1}(t) \qquad , i = 1, 2.$$

With respect to the control system

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

where $A(\cdot) \in PC(\mathbb{R}_+, \mathbb{R}^{n \times n})$ and $B(\cdot) \in PC(\mathbb{R}_+, \mathbb{R}^{n \times m})$ we will analyse the following

Stabilization Problem : (2.6)

Under which (necessary and) sufficient conditions does exist a state feedback $F(\cdot) \in PC(\mathbb{R}_+, \mathbb{R}^{m \times n})$ such that the closed loop system of (2.5)

$$\dot{x}(t) = (A + BF)(t)x(t), \qquad t \geq 0 \tag{2.7}$$

is exponentially stable. If a feedback exists, how can it be constructed?

Two ideas are essential to answer these questions: There is a result by *Coppel* (1978) which says that every system (2.1) of exponential dichotomy can be transformed by a Lyapunov transformation into a disconnected system

$$\dot{x}(t) = \begin{bmatrix} A_1(t) & 0 \\ 0 & A_2(t) \end{bmatrix} x(t), \quad t \geq 0$$

The second idea is the concept of controllability into resp. reconstructibility wrt subspaces introduced in Sections **1.2** and **1.3**.

Suppose (2.5) is uniformly completely controllable into $\mathcal{V}_1$, i.e. there exist $\sigma, a, b > 0$ such that

$$a\, I_n \leq W_2(t, t+\sigma) \leq b\, I_n \quad \text{ on } P_2^T(t)\mathbb{R}^n \quad \text{ for all } t \in \mathbb{R}_+ \tag{2.8}$$

In order to solve the problem we have to introduce

$$\tilde{W}_2(t, t+\sigma) \; := \; X(t) P_2 \int_t^{t+\sigma} X^{-1}(s) B(s) B^T(s) X^{-1^T}(s) e^{-2\alpha'(t-s)} ds \; P_2^T X^T(t)$$

for some $\alpha' > \alpha$.
Thus

$$a\, e^{-2\sigma\alpha'} I_n \; \leq \; \tilde{W}_2(t, t+\sigma) \; \leq \; b\, e^{2\sigma\alpha'} I_n \qquad \text{on } P_2^T(t)\mathbb{R}^n \quad \text{for all } t \in \mathbb{R}_+$$

and (see *Coppel* (1971) p. 41)

$$b^{-1} e^{-2\sigma\alpha'} I_n \; \leq \; \tilde{W}_2^{-1}(t, t+\sigma) \; \leq \; a^{-1} e^{2\sigma\alpha'} I_n \qquad \text{on } P_2(t)\mathbb{R}^n \quad \text{for all} \quad t \in \mathbb{R}_+ \tag{2.9}$$

Now we can state the main result

Proposition 2.2 Suppose (2.8) and for some $c > 0$

$$\| P_1(t) B(t) (P_2(t) B(t))^T \| \leq c \qquad \text{for all } t \in \mathbb{R}_+ \tag{2.10}$$

Then the transition matrix $\Phi_F(\cdot, \cdot)$ of the closed loop system (2.7) with respect to the feedback

$$F(t) = -\frac{1}{2}\, B^T(t) \tilde{W}_2^{-1}(t, t+\sigma) P_2(t)$$

satisfies for some $K' > 0$

$$\| \Phi_F(t, s) \| \leq K'\, e^{-\alpha(t-s)} \qquad \text{for all } t \geq s \geq 0.$$

Proof: Use a Lyapunov transformation, see *Coppel* (1978) Lecture 5, to transform the system (2.5) by a similarity action into the form

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} A_1(t) & 0 \\ 0 & A_2(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} B_1(t) \\ B_2(t) \end{bmatrix} u(t) \tag{2.11}$$

with fundamental matrix

$$X(t) = \begin{bmatrix} X_1(t) & 0 \\ 0 & X_2(t) \end{bmatrix}.$$

Then

$$\tilde{W}_2(t, t+\sigma) = \begin{bmatrix} 0 & 0 \\ 0 & V_2(t, t+\sigma) \end{bmatrix},$$

where

$$V_2(t, t+\sigma) = \int_t^{t+\sigma} \Phi_2(t,s) B_2(s) B_2^T \Phi_2^T(t,s) e^{-2\alpha'(t-s)} ds.$$

The feedback law becomes

$$u(t) = F(t)x(t) = -\frac{1}{2} B_2^T(t) V_2^{-1}(t, t+\sigma) x_2(t)$$

and the closed loop system is of the form

$$\begin{aligned} \dot{x}\,(t) &= A_1(t)x_1(t) - \frac{1}{2} B_1(t)B_2^T(t)V_2^{-1}(t,t+\sigma)x_2(t) \\ \dot{x}_2(t) &= [A_2(t) - \frac{1}{2} B_2(t)B_2^T(t)V_2^{-1}(t,t+\sigma)]x_2(t). \end{aligned} \tag{2.12}$$

Ikeda, Maeda and Kodoma (1975), Theorem 3.1, proved that the free motions of (2.12) are uniformly asymptotically bounded, more precisely:

$$\| x_2(t) \| \leq c_1 \, e^{-\alpha'(t-s)} \| x_2(s) \| \quad \text{for } t \geq s \geq 0 \tag{2.13}$$

for some $c_1 > 0$.
(2.12) is equivalent to

$$x_1(t) = \Phi_1(t,s)x_1(s) - \frac{1}{2} \int_s^t \Phi_1(t,\tau)B_1(\tau)B_2^T(\tau)V_2^{-1}(\tau,\tau+\sigma)x_2(\tau)d\tau \tag{2.14}$$

Let $c_2 := \frac{1}{2}\, K\, c\, a^{-1} c_1\, e^{2\sigma\alpha'}$ and apply (2.2), (2.10), (2.9) and (2.13) to (2.14). Then

$$\| x_1(t) \| \leq K\, e^{-\alpha(t-s)} \| x_1(s) \| + c_2 \int_s^t K\, e^{-\alpha(t-\tau)} e^{-\alpha'(\tau-s)} \| x_2(s) \| \, d\tau. \tag{2.15}$$

Because $\alpha' > \alpha$ we obtain

$$\begin{aligned} \| x_1(t) \| &\leq K\, e^{-\alpha(t-s)} \| x_1(s) \| + \frac{c_2}{\alpha' - \alpha} [e^{-\alpha(t-s)} - e^{-\alpha'(t-s)}] \| x_2(s) \| \\ &\leq c_3\, e^{-\alpha(t-s)} \; [\| x_1(s) \| + (1 - e^{(\alpha-\alpha'(t-s)}) \| x_2(s) \|] \\ &\leq c_3\, e^{-\alpha(t-s)} \; [\| x_1(s) \| + \| x_2(s) \|] \end{aligned} \tag{2.16}$$

where $c_3 := max \; \{K, \frac{c_2}{\alpha'-\alpha}\}$. Finally, the result follows by (2.13) and (2.16). □

For a bounded system (2.5) it can be shown that uniform complete controllability into $\mathcal{V}_1$ is also *necessary* for stabilizability.

Proposition 2.3 Suppose (2.5) is bounded, i.e. $A(\cdot)$ and $B(\cdot)$ are uniformly bounded in t. Then there exists a bounded $F(\cdot)$ such that the closed loop system (2.7) is exponentially stable if and only if (2.5) is uniformly completely controllable into $\mathcal{V}_1$.

Proof: Clearly the feedback given in Proposition 2.2 is bounded if (2.5) is bounded and uniformly completely controllable into $\mathcal{V}_1$. To prove the converse note that the boundedness of (2.5) implies the upper bounded in (2.8), see Remark 1.2.9(i). The lower bound is proved in the same way as in *Ikeda, Maeda and Kodoma* (1972), the necessity part of Theorem 3. □

The opposite problem of (2.6) is treated in the following

$$\textbf{Anti-stabilization problem:} \tag{2.17}$$

Under which (necessary and) sufficient conditions does exists a state feedback $F(\cdot) \in PC(\mathbb{R}_+, \mathbb{R}^{m\times n})$ such that (2.7) is *completely exponentially unstable, i.e.*

$$\| \Phi_F(t,s) \| \le L' \, e^{-\beta(s-t)} \quad \text{for all } s \ge t \text{ and some } \; L', \beta > 0,$$

where $\Phi_F(\cdot,\cdot)$ denotes the transition matrix of (2.7). This means every solution of the closed loop system is bounded from below by an exponentially growing function.

Definition 2.4 The system (2.5) is said to be *uniformly completely reachable from* $\mathcal{V}_2$ if there exist $\sigma, a', b' > 0$ such that

$$a' I_n \le Y_1(t-\sigma, t) \le b' I_n \qquad \text{on } P_1^T(t)\mathbb{R}^n, \quad \text{for all} \quad t \ge \sigma \tag{2.18}$$

Using the matrix

$$\tilde{Y}_1(t-\sigma,t) := \; X(t)P_1 \int_{t-\sigma}^{t} X^{-1}(s)B(s)B^T(s)e^{-2\beta'(t-s)}X^{-1^T}(s)ds \; P_1^T X^T(t) \quad \text{for } t \ge \sigma$$

and some $\beta' > \beta$ one obtains the following result.

Proposition 2.5 Suppose (2.18) and for some $c > 0$

$$\| P_2(t)B(t)(P_1(t)B(t))^T \| \le c \qquad \text{for all } t \ge \sigma \tag{2.19}$$

Then every fundamental matrix $X_F(\cdot)$ of the closed loop system (2.7) with respect to the feedback

$$F(t) = \frac{1}{2} B^T(t)\tilde{Y}_1^{-1}(t-\sigma,t)P_1(t)$$

satisfies for some $L' > 0$

$$\| X_F(t)x_0 \| \le L' \, e^{-\beta(s-t)} \, \| X_F(s)x_0 \| \qquad \text{for all} \quad s \ge t \ge \sigma, \; x_0 \in \mathbb{R}^n \tag{2.20}$$

Proof: Suppose (2.5) is of the form (2.11). Then

$$\tilde{Y}_1(t-\sigma,t) = \begin{bmatrix} \tilde{Y}_1(t) & 0 \\ 0 & 0 \end{bmatrix}$$

where

$$\tilde{Y}_1(t) = \int_{t-\sigma}^{t} \Phi_1(t,s)B_1(s)B_1^T(s)\Phi_1^T(t,s)e^{-2\beta'(t-s)}ds$$

Application of the feedback

$$F(t) = \frac{1}{2}\begin{bmatrix} B_1(t) \\ B_2(t) \end{bmatrix}^T \begin{bmatrix} \tilde{Y}_1^{-1}(t) & 0 \\ 0 & 0 \end{bmatrix} = \frac{1}{2}[B_1^T(t)\tilde{Y}_1^{-1}(t),\ 0]$$

yields

$$\begin{aligned} \dot{x}_1(t) &= [A_1(t) + \frac{1}{2}\, B_1(t)B_1^T(t)\tilde{Y}_1^{-1}(t)]x_1(t) \\ \dot{x}_2(t) &= A_2(t)x_2(t) + \frac{1}{2}\, B_2(t)B_1^T(t)\tilde{Y}_1^{-1}(t)x_1(t) \end{aligned} \tag{2.21}$$

Ikeda, Maeda and Kodoma (1975) Theorem 3.3 have proved that

$$\| x_1(t) \| \leq L_1\, e^{-\beta'(s-t)} \| x_1(s) \| \qquad \text{for all } s \geq t \geq \sigma \tag{2.22}$$

and some $L_1 > 0$. (2.21) yields

$$x_2(t) = \Phi_2(t,s)x_2(s) + \int_s^t \Phi_2(t,\tau)\frac{1}{2}\, B_2(\tau)B_1^T(\tau)\tilde{Y}_1^{-1}(\tau)x_1(\tau)d\tau$$

Taking norms and applying (2.2b), (2.19) and (2.22) gives for some $L_2 > 0$

$$\| x_2(t) \| \leq L_2\, e^{-\beta(s-t)} \| x_2(s) \| + \int_s^t L_2\, e^{-\beta(\tau-t)}\frac{1}{2}c \,\| \tilde{Y}_1^{-1}(\tau) \|\, L_1\, e^{-\beta'(s-\tau)}d\tau\, \| x_1(s) \|$$

(2.18) yields $\tilde{Y}_1^{-1}(t) \leq \frac{1}{a'}\, e^{2\sigma\beta'}$. Since $\beta' - \beta > 0$, we conclude

$$\begin{aligned} \| x_2(t) \| &\leq L_2\, e^{-\beta(s-t)} \| x_2(s) \| + \tilde{L} \int_s^t e^{(-\beta+\beta')\tau+\beta t-\beta' s} d\tau\ \| x_1(s) \| \\ &= L_2\, e^{-\beta(s-t)} \| x_2(s) \| + \frac{\tilde{L}}{\beta' - \beta}[e^{-\beta'(s-t)} - e^{-\beta(s-t)}] \| x_1(s) \| \\ &\leq L_2\, e^{-\beta(s-t)} \| x_2(s) \| \end{aligned} \tag{2.23}$$

for all $s \geq t \geq \sigma$. Now by (2.22) and (2.23) the result follows for $L' := max\ \{L_1, L_2\}$. □

Similar as in Proposition 2.3, for bounded systems one can prove that uniform complete reachability from $\mathcal{V}_2$ is necessary and sufficient for anti-stabilization by a bounded feedback.

Proposition 2.6 Suppose the system (2.5) is bounded. Then there exists a bounded $F(\cdot)$ such that every fundamental solution of the closed loop system (2.7) satisfies (2.20) iff (2.5) is uniformly completely reachable from $\mathcal{V}_2$.

Proof: Sufficiency follows from Proposition 2.5. To prove the necessity part note that boundedness of $A(\cdot)$ and $B(\cdot)$ implies a uniform upper bound for $Y_1(t-\sigma, t)$. This can be shown similar to Remark 1.2.9(i). The lower bound for $Y_1(t-\sigma, t)$ is proved analogously to *Ikeda, Maeda and Kodoma* (1972) Theorem 3.4. □

As an application of the previous results, in the remainder of this section we will treat the problem of stabilization by feedback of determinsitic state estimation. This problem was analysed for time-varying systems by *Ikeda, Maeda and Kodoma* (1975).
Here we will show how controllability and observability assumptions can be relaxed if systems of exponential dichotomy are considered.

Consider the system

$$\left.\begin{array}{rcl} \dot{x}(t) & = & A(t)x(t) + B(t)u(t) \\ \dot{y}(t) & = & C(t)x(t) + D(t)u(t) \end{array}\right\} \tag{2.24}$$

and assume that $\dot{x}(t) = A(t)x(t)$ possesses an exponential dichotomy as in (2.2) and $B(\cdot) \in PC(\mathbb{R}_+, \mathbb{R}^{n \times m})$, $C(\cdot) \in PC(\mathbb{R}_+, \mathbb{R}^{p \times n})$, $D(\cdot) \in PC(\mathbb{R}_+, \mathbb{R}^{p \times m})$.
Following *Johnson* (1969) the n-dimensional deterministic state estimation should be of the form

$$\dot{z}(t) = (A - HC)(t)z(t) + (B - HD)(t)u(t) + H(t)y(t) \tag{2.25}$$

where the design parameter $H(\cdot) \in PC(\mathbb{R}_+, \mathbb{R}^{n \times p})$ has to stabilize the homogeneous equation

$$\dot{e}(t) = (A - HC)(t)e(t) \quad , \quad e(t) := x(t) - z(t) \tag{2.26}$$

Applying the statements of Remark 2.1 and Proposition 1.3.7 concerning the adjoint system to Proposition 2.5 yields the following corollary.

Corollary 2.7 Suppose the system (2.24) is uniformly completely reconstructible wrt $\mathcal{V}_1^a$ and

$$\| P_1^T(t)\, C^T(t)(P_2^T(t)C^T(t))^T \| \leq c \qquad \text{for all } t \geq \sigma$$

and some $c > 0$. Then there exists a compensator $H(\cdot) \in PC([\sigma, \infty), \mathbb{R}^{n \times p})$ such that the transition matrix $\Phi_H(\cdot,\cdot)$ of (2.26) satisfies

$$\| \Phi_H(t,s) \| \leq \tilde{L}\, e^{-\beta(t-s)} \qquad \text{for all} \quad t \geq s \geq \sigma \quad \text{and some} \quad \tilde{L} > 0.$$

The state estimate $z(\cdot)$ given by (2.25) will now be substituted for the real state $x(\cdot)$ into the control law

$$u(t) = F(t)x(t) + v(t)$$

After a straightforward calculation one gets for the closed loop system

$$\begin{array}{rcl} \begin{bmatrix} \dot{x}(t) \\ \dot{z}(t) \end{bmatrix} & = & \begin{bmatrix} A(t) & B(t)F(t) \\ H(t)C(t) & A(t) - H(t)C(t) + B(t)F(t) \end{bmatrix} \begin{bmatrix} x(t) \\ z(z) \end{bmatrix} + \begin{bmatrix} B(t) \\ B(t) \end{bmatrix} v(t) \\ y(t) & = & [C(t), D(t)F(t)] \begin{bmatrix} x(t) \\ z(t) \end{bmatrix} + D(t)v(t) \end{array} \tag{2.27}$$

Using the previous results, sufficient conditions can be derived to ensure exponential stability of the homogeneous part of (2.27).

Proposition 2.8 Suppose the system (2.24) is uniformly completely controllable into $\mathcal{V}_1$ and uniformly completely reconstructible wrt $\mathcal{V}_1^a$ with controllability and reconstructibility intervals of length $\sigma > 0$. Moreover the following inequalities

$$\begin{array}{rcll} \| P_1(t)B(t)\, (P_2(t)B(t))^T \| & \leq & c & \qquad \text{for all } t \geq 0 \\ \| P_1^T(t)C^T(t)(P_2^T(t)C^T(t))^T \| & \leq & c & \qquad \text{for all } t \geq 0 \end{array}$$

hold for some $c > 0$.

Then there exist a feedback $F(\cdot) \in PC(\mathbb{R}_+, \mathbb{R}^{m \times n})$ and an estimator gain $H(\cdot) \in PC(\mathbb{R}_+, \mathbb{R}^{n \times p})$ such that for some $M > 0$

$$\begin{matrix} x(t) \\ z(t) \end{matrix} \leq M \begin{matrix} x(t_0) \\ z(t_0) \end{matrix} e^{-\omega(t-t_0)} \qquad \text{for } t \geq t_0 \geq 0$$

where $\omega := min\ \{\alpha, \beta\}$.

Proof: By Proposition 2.2 and Corollary 2.7 there exist $F(\cdot)$ and $H(\cdot)$ such that

$$\begin{aligned} \dot{x}(t) &= (A - BF)(t)x(t) \\ \dot{e}(t) &= (A - HC)(t)e(t) \end{aligned}$$

are exponentially stable with decay rate α resp. β. Now we are in a position to mirror completely the proof of *Ikeda, Maeda and Kodoma* (1975) pp. 323-325 for our situation. This goes through without difficulties, we therefore omit it. □

4.3 Bohl exponent and Bohl transformations

Consider a differential equation of the form

$$\dot{x}(t) = A(t)x(t), \quad t \geq 0 \tag{3.1}$$

where $A(\cdot) \in PC(\mathbb{R}_+, \mathbb{C}^{n\times n})$. For a characterization of the stability behaviour of (3.1) the following definition due to *Bohl* (1913) is useful.

Definition 3.1 (Bohl exponent)
The (upper) *Bohl exponent* $k_B(A)$ of the system (3.1) is given by

$$k_B(A) = \inf\{-\omega \in \mathbb{R} \mid \exists\, M_\omega > 0 : t \geq t_0 \geq 0 \Rightarrow \| \Phi(t,t_0) \| \leq M_\omega\, e^{-\omega(t-t_0)}\}$$

It is possible that $k_B(A) = \pm\infty$. If (3.1) is time-invariant, i.e. $A(\cdot) \equiv A \in \mathbb{C}^{n\times n}$, then

$$k_B(A) = \max_{i \in \underline{n}} Re\, \lambda_i(A)$$

where $\lambda_i(A), i \in \underline{n}$, are the eigenvalues of A.

The following properties of the Bohl exponent can be found in *Daleckii and Krein* (1974) pp. 119 - 121.

Proposition 3.2 (i) The Bohl exponent of the system (3.1) is finite if and only if

$$\sup_{0 \leq |t-s| \leq 1} \| \Phi(t,s) \| < \infty \tag{3.2}$$

In particular $k_B(A)$ is finite if $A(\cdot)$ is integrally bounded (cf. Section 1).

(ii) If $k_B(A) < \infty$ it can be determined via

$$k_B(A) = \limsup_{s,t-s\to\infty} \frac{\log \| \Phi(t,s) \|}{t-s}. \tag{3.3}$$

For later use we need the following more restrictive definition

Definition 3.3 (Strict Bohl exponent)
The Bohl exponent of the system (3.1) is said to be *strict* if it is finite and

$$k_B(A) = \lim_{s,t-s\to\infty} \frac{\log \| \Phi(t,s) \|}{t-s}.$$

Lemma 3.4 Suppose $a(\cdot) \in PC(\mathbb{R}_+, \mathbb{C})$ has a strict finite Bohl exponent and $A(\cdot) \in PC(\mathbb{R}_+, \mathbb{C}^{n\times n})$, then

(i) $k_B(-a) = -k_B(a)$

(ii) $k_B(aI_n + A) = k_B(a) + k_B(A)$ (shift property).

Proof: (i) follows from

$$\begin{aligned} k_B(a) &= \lim_{s,t-s\to\infty} \frac{\log|e^{\int_s^t a(\tau)d\tau}|}{t-s} &= \lim_{s,t-s\to\infty} \frac{\int_s^t a(\tau)d\tau}{t-s} \\ &= -\lim_{s,t-s\to\infty} \frac{\int_s^t -a(\tau)s\tau}{t-s} &= -k_B(-a). \end{aligned}$$

In order to prove (ii) note that the transition matrix of $\dot{z}(t) = [A(t) + a(t)I_n]z(t)$ is given by

$$\psi(t,s) = \Phi(t,s)\exp\left(\int_s^t a(\tau)d\tau\right).$$

Thus by Definition 3.1

$$k_B(A) \leq k_B(A + aI_n) - k_B(a), \tag{3.4}$$

which proves (ii). □

In the literatur the *(upper) Lyapunov exponent* is better known

$$k_L(A) = \inf\{-\omega \in \mathbb{R} \mid \exists M_\omega > 0 : t \geq 0 \Rightarrow \| \Phi(t,0) \| \leq M_\omega\, e^{-\omega t}\}.$$

For time-invariant systems the Bohl and Lyapunov exponents coincide whereas in general

$$k_L(A) \leq k_B(A).$$

The following example due to *Perron* (1930) illustrates that these exponents may be different even for scalar systems.

Example 3.5 (Perron equation)
Consider the scalar system

$$\dot{x}(t) = a(t)x(t), \qquad t \geq 0, \tag{3.5}$$

where $a(\cdot) \in PC(\mathbb{R}_+, \mathbb{C})$ is given by

$$a(t) = \sin\log t + \cos\log t$$

The transition matrix is

$$\Phi(t,s) = e^{t\sin\log t - s\sin\log s}$$

and since $a(\cdot)$ is integrally bounded, we have

$$k_L(a) = \lim_{t\to\infty}\sup\ \sin\log t = 1.$$

The Bohl exponent, however, can be shown to be $\sqrt{2}$ (see *Daleckii and Krein* (1974) p. 123). □

Remark 3.6 For the system (3.1) one has $k_L(A) < 0$ if and only if (3.1) is *asymptotically stable*, i.e. (1.2) holds true but k may depend on t_0 and the convergence in (1.3) need not be uniform.

The following characterizations of exponential stability are proved in *Daleckii and Krein* (1974) p. 129 and p. 130.

Theorem 3.7 Suppose $A(\cdot)$ is integrally bounded and $p \in (0,\infty)$, then the following statements are equivalent:

(i) (3.1) is exponentially stable

(ii) $k_B(A) < 0$

(iii) there exists a constant c_p, such that

$$\int_{t_0}^{\infty} \| \Phi(t,t_0) \|^p \, dt \le c_p \qquad \text{for all } t_0 \ge 0.$$

(iv) For every bounded $f(\cdot) \in PC(\mathbb{R}_+, \mathbb{C}^n)$, the solution of the initial value problem

$$\dot{x}(t) = A(t)x(t) + f(t), \qquad t \ge 0, \;\; x(0) = 0$$

is bounded.

Under the weaker assumption $k_B(A) < \infty$, conditions (i) - (iii) are equivalent.

We now analyse the effect of time-varying linear coordinate transformations $z(t) = T(t)^{-1}x(t)$ on the system (3.1); where $T(\cdot) \in PC^1(\mathbb{R}_+, GL_n(\mathbb{C}))$ (the piecewise continuously differentiable $n \times n$ functions on $\mathbb{R}_+$ which have nonsingular values), cf. Section **1.1**.
The group of Lyapunov transformations preserves the properties of stability, instability and asymptotic stability. The property of exponential stability is invariant with respect to a larger group of transformations.

Definition 3.8 (Bohl transformation)
$T(\cdot) \in PC^1(\mathbb{R}_+, GL_n(\mathbb{C}))$ is said to be a *Bohl transformation* if

$$\inf\{\varepsilon \in \mathbb{R} \mid \exists M_\varepsilon > 0 : \| T(t)^{-1} \| \cdot \| T(s) \| \le M_\varepsilon \, e^{\varepsilon|t-s|} \; \forall \, t,s \ge 0\} = 0.$$

In the following example scalar Bohl transformations are characterized.

Example 3.9 Suppose $\theta(\cdot) \in PC^1(\mathbb{R}_+, \mathbb{C}^*)$, and let $a(\cdot) = \dot{\theta}(\cdot)\theta(\cdot)^{-1}$ so that

$$\dot{\theta}(t) = a(t)\theta(t) \quad \text{and} \quad (\theta(t)^{-1})^{\cdot} = -a(t)\theta(t)^{-1}.$$

The fundamental solutions of these differential equations are

$$\varphi(t,t_0) = \theta(t)\theta(t_0)^{-1} \quad \text{and} \quad \tilde{\varphi}(t,t_0) = \theta(t)^{-1}\theta(t_0).$$

By Definition 3.8 $\theta(\cdot)$ is a Bohl transformation if and only if for every $\varepsilon > 0$ there exists $M_\varepsilon > 0$ such that

$$M_\varepsilon^{-1} \, e^{-\varepsilon(t-s)} \;\le\; \varphi(t,s) \;\le\; M_\varepsilon \, e^{\varepsilon(t-s)} \qquad \text{for all} \quad t \ge s \ge 0$$

and this condition holds if and only if $a(\cdot)$ has strict Bohl exponent 0 .

The following proposition implies, in particular, that Bohl transformations preserve exponential stability (but not necessarily stability and asymptotic stability).

Proposition 3.10

(i) The set of Bohl transformations forms a group with respect to (pointwise) multiplication.

(ii) The Bohl exponent is invariant with respect to Bohl transformations.

Proof: (i) is an immediate consequence of Definition 3.8. To prove (ii), let $\dot{x}(t) = \hat{A}(t)x(t)$ be similar to (3.1) via the Bohl transformation $T(\cdot)$. Since the transition matrix of $\dot{x}(t) = \hat{A}(t)x(t)$ is given by $\hat{\Phi}(t,s) = T(t)^{-1}\Phi(t,s)T(s)$ by Definition 3.8 one obtains

$$k_B(A) \leq k_B(\hat{A}).$$

By (i), it follows that $k_B(A) = k_B(\hat{A})$. □

Example 3.11 Consider a *periodic* scalar system

$$\dot{x}(t) = a(t)x(t), \quad t \geq 0, \tag{3.6}$$

where $a(\cdot) \in PC(\mathbb{R}_+, \mathbb{C})$ is of period $\mu > 0$. Set

$$a_0 := \frac{1}{\mu}\int_0^\mu a(\tau)d\tau \text{ and } a(t) = a_0 + d(t).$$

Then $| \int_0^t d(\tau)d\tau |$ is bounded in $t \geq 0$ and so

$$\theta(t) := e^{\int_0^t d(\tau)d\tau}$$

defines a Bohl transformation. Hence by Proposition 3.10

$$k_B(a) = k_B(a - \theta^{-1}\dot{\theta}) = a_0.$$

Thus (3.6) is exponentially stable if and only if

$$\frac{1}{\mu}\int_0^\mu a(\tau)d\tau < 0.$$

Proposition 3.12 Every scalar system

$$\dot{x}(t) = a(t)x(t), \qquad t \geq 0$$

for which $a(\cdot)$ has strict, finite Bohl exponent $a(\cdot) \in PC(\mathbb{R}_+, \mathbb{C})$ can be transformed via the Bohl transformation

$$\theta(t) = \exp(\int_0^t (a(\tau) - k_B(a))d\tau)$$

into the time-invariant linear system

$$\dot{z}(t) = k_B(a)z(t), \quad t \geq 0.$$

Proof: Lemma 3.4 yields

$$k_B(a - k_B(a)) = k_B(a) - k_B(a) = 0.$$

Thus by Example 3.11 the $\theta(\cdot)$ which solves

$$\dot{\theta}(t) = (a(t) - k_B(a))\theta(t), \quad \theta(0) = 1$$

defines a Bohl-transformation. Setting $z(t) = \theta^{-1}(t)x(t)$ yields

$$\dot{z}(t) = [a(t) - \theta^{-1}(t)\dot{\theta}(t)]z(t) = k_B(a) \cdot z(t).$$

□

Remark 3.13 The Perron equation in Example 3.5 together with the previous proposition implies that, in general, a Bohl transformation does not preserve the Lyapunov exponent.

This section is concluded by stating some known perturbation results concerning the Bohl exponent for the system

$$\dot{x}(t) = [A(t) + \Delta(t)]x(t), \ \ t \geq 0 \tag{3.7}$$

where $\Delta(\cdot) \in PC(\mathbb{R}_+, \ \mathbb{C}^{n \times n})$.

Proposition 3.14 For any $\varepsilon > 0$ there exists $\delta > 0$, such that

$$\limsup_{s,t-s\to\infty} \frac{1}{t-s} \int_s^t \| \Delta(\tau) \| \, d\tau < \delta$$

implies

$$k_B(A + \Delta) \leq k_B(A) + \varepsilon.$$

The proof is straightforward and can be found in *Daleckii and Krein* (1974) p. 125.

Corollary 3.15

(i) Suppose $\Delta \in PC(\mathbb{R}_+, \ \mathbb{C}^{n \times n})$ satisfies

$$\limsup_{s,t-s\to\infty} \frac{1}{t-s} \int_s^t \| \Delta(\tau) \| \, d\tau = 0.$$

Then

$$k_B(A + \Delta) = k_B(A).$$

(ii) If systems of the form (3.1) are identified with the corresponding matrix functions $A(\cdot)$, then the set of exponentially stable systems (3.1) is open in $PC(\mathbb{R}_+, \ \mathbb{C}^{n \times n})$ with respect to the L_∞-topology.

Systems (3.1), (3.7) are called *asymptotically equivalent* (resp. *integrally equivalent*) if

$$\lim_{t\to\infty} \| \Delta(t) \| = 0 \quad (\text{resp. } \int_0^\infty \| \Delta(t) \| \, dt < \infty).$$

The above corollary shows that asymptotically or integrally equivalent systems have the same Bohl exponent.

4.4 The structured stability radius

In this section it is assumed that the nominal system (3.1) is subjected to perturbations of the form $\Delta(t) = B(t)D(t)C(t)$, so that the perturbed system is

$$\dot{x}(t) = [A(t) + B(t)D(t)C(t)]x(t), \quad t \geq 0 \tag{4.1}$$

where $D(\cdot)$ is an unknown, bounded, time-varying disturbance matrix ($D(\cdot) \in PC_b(\mathbb{R}_+, \ \mathbb{C}^{m \times p})$) and $B(\cdot)$ and $C(\cdot)$ are known time-varying *scaling matrices* defining the structure of the perturbation. Throughout this section we assume the tripel $\sum = (A, B, C)$ consists of matrix functions

$$A(\cdot) \in PC(\mathbb{R}_+, \ \mathbb{C}^{n \times n}), \ B(\cdot) \in PC(\mathbb{R}_+, \ \mathbb{C}^{n \times m}), \ C(\cdot) \in PC(\mathbb{R}_+, \ \mathbb{C}^{p \times n}). \tag{4.2}$$

Consider for instance the time-varying oscillator

$$y(t) + a_1(t)\dot{y}(t) + a_2(t)y(t) = 0.$$

This can be written in the form of (3.1) with

$$A(t) = \begin{bmatrix} 0 & 1 \\ -a_2(t) & -a_1(t) \end{bmatrix}.$$

If the parameters a_1, a_2 are uncertain we can model this by setting the scaling matrices $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $C = I_2$, whereas if only a_2 is uncertain we set $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $C = [1, 0]$.

By Corollary 3.15 (ii) the set of exponentially stable systems is open in $PC(\mathbb{R}_+, \mathbb{C}^{n\times n})$ with respect to the L_∞–norm. [1] Its complement which is closed will be denotes by $\mathcal{U}_n(\mathbb{R}_+, \mathbb{C})$. We will call the elements of $\mathcal{U}_n(\mathbb{R}_+, \mathbb{C})$ *unstable* (not exponentially stable). Note, however, that with respect to this shorthand terminology an unstable system may in fact be asymptotically stable. The following definition extends the concept of stability radius introduced in *Hinrichsen and Pritchard* (1986a,b) to time-varying systems.

Definition 4.1 (Stability radius)
Given $\sum = (A, B, C)$, the *(complex) stability radius* $r_{\mathbb{C}}(A; B, C)$ is defined by

$$r_{\mathbb{C}}(A; B, C) = \inf\{\| D \|_\infty;\ D \in PC(\mathbb{R}_+, \mathbb{C}^{m\times p}), A + B\,D\,C \in \mathcal{U}_n(\mathbb{R}_+, \mathbb{C})\}. \tag{4.3}$$

The *unstructured stability radius* of (3.1) is defined by

$$r_{\mathbb{C}}(A) = r_{\mathbb{C}}(A; I_n, I_n).$$

Note that $r_{\mathbb{C}}(A; B, C) = \inf \emptyset = \infty$ if there does not exist a perturbation matrix $D \in PC_b(\mathbb{R}_+, \mathbb{C}^{m\times p})$ such that $A + B\,D\,C \in \mathcal{U}_n(\mathbb{R}_+, \mathbb{C})$.

Remark 4.2

(i) The unstructed stability radius $r_{\mathbb{C}}(A)$ measures the distance of $A(\cdot)$ from the set $\mathcal{U}_n(\mathbb{R}_+, \mathbb{C})$ of unstable matrices with respect to the L_∞–norm.

(ii) If $\sum = (A, B, C)$ consits of *real* matrix functions the real stability radius $r_{\mathbb{R}}(A; B, C)$ is defined in an analogous fashion. Unfortunately, this stability radius - although more important for applications - is much more difficult to analyse and so we concentrate on the complex stability radius.

(iii) In the time-invariant case it is shown in *Hinrichsen and Pritchard* (1986 b) that

$$r_{\mathbb{C}}(A; B, C) = \frac{1}{\max_{\omega\in\mathbb{R}} \| G(i\omega) \|}$$

where $G(i\omega) = C(i\omega - A)^{-1}B$ (in particular $r_{\mathbb{C}}(A; B, C) = \infty$ if $G \equiv 0$).

The unstructed stability radius has the following properties

[1] This expression is used although $\| \cdot \|_{L_\infty}$ is only a pseudo-norm on $PC(\mathbb{R}_+, \mathbb{C}^{n\times n})$.

Lemma 4.3

(i) $r_{\mathbb{C}}(A) = 0 \;\Leftrightarrow\; A \in \mathcal{U}_n(\mathbb{R}_+,\ \mathbb{C})$

(ii) $r_{\mathbb{C}}(\alpha A) = \alpha\, r_{\mathbb{C}}(A)$ for all $\alpha \geq 0$

(iii) $A \mapsto r_{\mathbb{C}}(A)$ is continuous on $PC(\mathbb{R}_+,\ \mathbb{C}^{n\times n})$

(iv) $r_{\mathbb{C}}(A+\Delta) \geq r_{\mathbb{C}}(A) - \| \Delta(\cdot) \|_{L_\infty(0,\infty;\ \mathbb{C}^{n\times n})}$ for any $\Delta \in PC_b(\mathbb{R}_+,\ \mathbb{C}^{n\times n})$

(v) $0 < r_{\mathbb{C}}(A) \leq -k_B(A)$ if $A(\cdot)$ is exponentially stable.

Proof: (i) - (iv) follow directly from the definition. (i) yields the first inequality in (v) and the second is a consequence of $A - k_B(A)I_n \in \mathcal{U}_n(\mathbb{R}_+,\ \mathbb{C})$ (since $k_B(A - k_B(A)) = 0$ by Lemma 3.4 (ii)). □

Remark 4.4 Suppose $\sum = (A, B, C)$ and $k_B(A) < 0$, then it is easily verified that

$$r_{\mathbb{C}}(A) \leq \| B(\cdot)_{|[t_o,\infty)} \|_{L_\infty} \cdot \| C(\cdot)_{|[t_0,\infty)} \|_{L_\infty} \cdot r_{\mathbb{C}}(A; B, C) \quad \text{for all } t_0 \geq 0.$$

(where we define $0 \cdot \infty = \infty$).

Now several important *invariance properties* of $r_{\mathbb{C}}(A; B, C)$ are derived

Proposition 4.5 (Asymptotical or integral equivalence)
Suppose the system (4.1) is asymptotically or integrally equivalent to $\dot{x}(t) = \hat{A}(t)x(t)$, then

$$r_{\mathbb{C}}(A; B, C) = r_{\mathbb{C}}(\hat{A}; B, C).$$

Proof: By Corollary 3.15(i)

$$k_B(A + B\,D\,C) = k_B(\hat{A} + B\,D\,C)$$

for arbitrary $D \in PC_b(\mathbb{R}_+,\ \mathbb{C}^{m\times p})$. Thus the class of destabilizing D's is the same for $A+BDC$ and $\hat{A} + B\,D\,C$. □

Corollary 4.6 Suppose B, C are constant and $\dot{x}(t) = A(t)x(t)$ is asymptotically or integrally equivalent to a constant exponentially stable system $\dot{x}(t) = A_0 x(t)$, $A_0 \in \mathbb{C}^{n\times n}$, then

$$r_{\mathbb{C}}(A; B, C) = \frac{1}{\max_{\omega\in\mathbb{R}} \| C(i\omega I - A_0)^{-1} B \|}.$$

In contrast with the Bohl exponent the unstructed stability radius is not invariant with respect to Bohl transformations. In fact any exponentially stable time-invariant system $\dot{x}(t) = Ax(t)$ can be brought arbitrary close to an unstable system by constant similarity transformations. The following example illustrates that there exists sequences of time-invariant systems such that $k_B(A_k) \to -\infty$, $r_{\mathbb{C}}(A_k) \to 0$ as $k \to \infty$.

Example 4.7 Set

$$A_k = -\begin{bmatrix} k & k^3 \\ 0 & k \end{bmatrix}, D_k = \begin{bmatrix} 0 & 0 \\ -k^{-1} & 0 \end{bmatrix}$$

then $\sigma(A_k) = \{-k\}$, $\sigma(A_k + D_k) = \{0, -2k\}$. Since $\| D_k \| \to 0$ as $k \to \infty$, $r_{\mathbb{C}}(A_k) \to 0$ as $k \to \infty$, but

$$k_B(A_k) = \max_{i=1,2} Re\, \lambda_i(A_k) = -k \to -\infty \quad \text{as } k \to \infty.$$

For *scalar* Bohl transformations one obtains

Proposition 4.8 Suppose $\sum = (A, B, C)$ and $\theta(\cdot) \in PC^1(\mathbb{R}_+,\ \mathbb{C})$ is a Bohl tranformation, then

$$r_{\mathbb{C}}(A - \theta^{-1}\dot{\theta} I_n; B, C) \ = \ r_{\mathbb{C}}(A; B, C).$$

Proof: By Proposition 3.10 (ii)

$$k_B(A - \theta^{-1}\dot{\theta} I_n + B\, D\, C) = k_B(A + B\, D\, C)$$

for every $D(\cdot) \in PC_b(\mathbb{R}_+,\ \mathbb{C}^{m \times p})$. The result follows just as in the proof of Proposition 4.5. □

The stability radius $r_{\mathbb{C}}(A; B, C)$ is invariant with respect to general Bohl transformations if the scaling matrices $B(\cdot)$, $C(\cdot)$ are transformed as well as the nominal system matrix $A(\cdot)$.

Proposition 4.9 Suppose $\sum = (A, B, C)$ and $T(\cdot) \in PC^1(\mathbb{R}_+, GL_n(\ \mathbb{C}))$ defines a Bohl transformation, then

$$r_{\mathbb{C}}(T^{-1} A\, T - T^{-1}\dot{T}; T^{-1}B, C\, T) = r_{\mathbb{C}}(A; B, C).$$

Proof: By Proposition 3.10 (ii)

$$k_B(T^{-1} A\, T - T^{-1}\dot{T} + T^{-1} B\, D\, C\, T) = k_B(A + B\, D\, C)$$

for every $D(\cdot) \in PC_b(\mathbb{R}_+,\ \mathbb{C}^{m \times p})$. □

For exponentially stable scalar systems we have the nice result that the unstructured stability radius coincides with the negative of the Bohl exponent. This is a direct consequence of the previous proposition and Proposition 3.12 for the case where the scalar system has a strict finite Bohl exponent. However the same result holds without this assumption.

Proposition 4.10 Suppose $a(\cdot) \in PC(\mathbb{R}_+,\ \mathbb{C})$ and the scalar system $\dot{x}(t) = a(t)x(t)$, $t \geq 0$ is exponentially stable, then

$$r_{\mathbb{C}}(a) = -k_B(a). \tag{4.4}$$

We omit the proof which is straightforward.

Note that the proof of Lemma 4.3 shows that for time-varying scalar systems the *constant* disturbance $d(\cdot) \equiv r_{\mathbb{C}}(A)$ destabilizes the nominal system.

4.5 The perturbation operator

In the time-invariant case (see *Hinrichsen and Pritchard* (1986b)), the stability radius can be characterized via the convolution operator

$$\begin{array}{rrcl} L_0 \ : & L_2(0,\infty;\ \mathbb{C}^m) & \to & L_2(0,\infty;\ \mathbb{C}^p) \\ & u(\cdot) & \mapsto & (t \mapsto \int_0^t C e^{A(t-s)} B u(s) ds). \end{array} \tag{5.1}$$

where $L_q(t_0,\infty;\ \mathbb{C}^m)$ denotes the set of functions $h:\ [t_0,\infty)\to\ \mathbb{C}^m$ such that $\int_{t_0}^{\infty}\ \| h(s)\|^q\ ds<\infty$, $t_0\in\mathbb{R},\ \ q>0$.

Proposition 5.1 Let $\sum=(A,B,C)$ be a tripel of constant matrices and $\dot{x}(t)=Ax(t)$ be exponentially stable, then

$$r_{\mathbb{C}}(A;B,C)=\frac{1}{\| L_0 \|}$$

where $\| L_0 \|$ is the induced norm of the operator L_0 defined in (5.1).

In order to explore the possibility of obtaining similar results for time-varying systems we assume, throughout this section,

$$\left.\begin{array}{l} A(\cdot)\in PC(\mathbb{R}_+,\ \mathbb{C}^{n\times n}),\quad B(\cdot)\in PC_b(\mathbb{R}_+,\ \mathbb{C}^{n\times m}) \\ C(\cdot)\in PC_b(\mathbb{R}_+,\ \mathbb{C}^{p\times n}),\quad k_B(A)<0 \end{array}\right\} \tag{5.2}$$

With any such tripel $\sum=(A,B,C)$ we associate a parametrized family of *perturbation operators* $(L_{t_0}^{\Sigma})_{t_0\in\mathbb{R}_+}$ defined by

$$\begin{array}{rcll} L_{t_0}^{\Sigma}\ :\ L_2(t_0,\infty;\ \mathbb{C}^m) & \to & L_2(t_0,\infty;\ \mathbb{C}^p) & ,t_0\geq 0 \\ u(\cdot) & \mapsto & (t\mapsto \int_{t_0}^{t}\ C(t)\Phi(t,s)B(s)u(s)ds) & \end{array} \tag{5.3}$$

In the following proposition we will show that these maps are well defined. Note that in the time-invariant case $\| L_{t_0}^{\Sigma} \|=\| L_0 \|$ for all $t_0\geq 0$.

Proposition 5.2 Suppose (5.2) and let $\sum=(A,B,C)$. Then

(i) $L_{t_0}^{\Sigma}$ is a bounded operator.

(ii) $t_0\mapsto\| L_{t_0}^{\Sigma} \|$ is monotonically decreasing on $\mathbb{R}_+$.

(iii) $\| L_{t_0}^{\Sigma} \|^{-1}\leq r_{\mathbb{C}}(A;B,C)$ for all $t_0\geq 0$.

(iv) If A,B,C are periodic with some common positive period, then

$$\| L_{t_0}^{\Sigma} \|=\| L_{t_1}^{\Sigma} \| \qquad \text{for all } t_0,t_1\in\mathbb{R}_+.$$

(v) In the unstructured case, i.e. $B(\cdot)\equiv C(\cdot)\equiv I_n$, if

$$\| \Phi(t,s) \|\leq M\ e^{-\omega(t-s)}\quad \text{for all}\quad t\geq s\geq 0 \text{ and some}\quad M,\omega>0$$

then

$$\frac{\omega}{M}\ \leq \| L_{t_0}^{\Sigma} \|^{-1}\leq \lim_{t_n\to\infty}\ \| L_{t_n}^{\Sigma} \|^{-1}\leq\ r_{\mathbb{C}}(A)\ \leq\ -k_B(A). \tag{5.4}$$

Proof: We write as short hand notations L_{t_0} instead of $L_{t_0}^{\Sigma}$ and $L_q(t_0,r)$ instead of $L_q(t_0,\infty;\ \mathbb{C}^r)$, $q,r\geq 1$.

(i) Let $u(\cdot)\in L_2(t_0,m)$ then by changing variables and using the inequality

$$\begin{array}{rcl} \| f\star v \|_{L_2} & \leq & \| f \|_{L_1}\cdot\| v \|_{L_2}\quad \text{for } f\in L_1,\ \ v\in L_2 \\ \| L_{t_0}\, u \|^2_{L_2(t_0,p)} & = & \displaystyle\int_{t_0}^{\infty}\ \| \int_{t_0}^{t}\ C(t)\Phi(t,s)B(s)u(s)ds \|^2\ dt \end{array}$$

$$
\begin{aligned}
&\leq (\| C \|_{L_\infty} \| B \|_{L_\infty} \ M)^2 \int_{t_0}^{\infty} [\int_{t_0}^{t} e^{-\omega(t-s)} \| u(s) \| ds]^2 dt \\
&= (\| C \|_{L_\infty} \| B \|_{L_\infty} \ M)^2 \int_{0}^{\infty} | \int_{0}^{\tau} e^{-\omega(\tau-\sigma)} \| u(\sigma + t_0) \| d\sigma |^2 d\tau \\
&\leq (\| C \|_{L_\infty} \| B \|_{L_\infty} \ M)^2 \| e^{-\omega\cdot} \|^2_{L_1(0,1)} \cdot \| u(\cdot + t_0) \|^2_{L_2(0,m)} \\
&\leq [(\| C \|_{L_\infty} \| B \|_\infty \ M)^2/\omega^2] \| u \|^2_{L_2(t_0,m)} .
\end{aligned}
$$

This shows that L_{t_0} is bounded and the first inequality holds in (5.4).
(ii) Suppose $0 \leq t_0 < t_1$ and $u(\cdot) \in L_2(t_1, m)$, $\| u(\cdot) \| = 1$. Extending $u(\cdot)$ to $\overline{u}(\cdot)$ by $u(t) = 0$ for $t \in [t_0, t_1)$ yields $\overline{u}(\cdot) \in L_2(t_0, m)$ with $\| \overline{u}(\cdot) \| = 1$. Now

$$
\begin{aligned}
\| L_{t_1} u \|^2_{L_2(t_1,p)} &= \int_{t_1}^{\infty} \| \int_{t_1}^{t} C(t)\Phi(t,s)B(s)u(s)ds \|^2 dt \\
&= \int_{t_0}^{\infty} \| \int_{t_0}^{t} C(t)\Phi(t,s)B(s)\overline{u}(s)ds \|^2 dt \\
&= \| L_{t_0} \overline{u} \|^2_{L_2(t_0,p)} .
\end{aligned}
$$

From which (ii) follows.
(iii) Let $D(\cdot) \in PC_b(t_0, \infty; \ \mathbb{C}^{m\times p})$ be such that

$$
\| D \|_{L_\infty} < \ \| L_{t_0} \|^{-1} \tag{5.5}
$$

then we have to show that the perturbed system

$$
\dot{x}(t) = [A(t) + B(t)D(t)C(t)]x(t), \quad t \geq t_0, \tag{5.6}
$$

is exponentially stable. By Theorem 3.7 and Proposition 3.2 it is sufficient to prove that the solutions $x(\cdot) = x(\cdot; t_0', x_0)$ of (5.6) (with $t_0' \geq t_0$) satisfy for some $k > 0$

$$
\sup_{t_0' \geq t_0} \| x(\cdot; t_0', x_0) \|_{L_2(t_0',n)} \leq \ k \| x_0 \| \qquad \text{for all } x_0 \in \ \mathbb{C}^n \tag{5.7}
$$

$$
\sup_{0 \leq |t-t_0'| \leq 1} \| x(t; t_0', x_0) \| \leq \ k \| x_0 \| \qquad \text{for all } x_0 \in \ \mathbb{C}^n. \tag{5.8}
$$

Now by variations of constants, for $t \geq t_0'$

$$
x(t; t_0', x_0) = \Phi(t, t_0')x_0 + \int_{t_0'}^{t} \Phi(t,s)B(s)D(s)C(s)x(s; t_0', x_0)ds \tag{5.9}
$$

and hence for $y(t) := C(t)x(t)$, $y_0(t) := C(t)\Phi(t, t_0')x_0 \in L_2(t_0', p)$

$$
y(t) = y_0(t) + (L_{t_0'} Dy)(t).
$$

By the contradiction principle and (5.5) equation (5.9) has a unique solution in $L_2(t_0', p)$ and

$$
\begin{aligned}
\| y \|_{L_2(t_0',p)} &\leq \| (I - L_{t_0'} D)^{-1} \| \cdot \| y_0 \|_{L_2(t_0',p)} \\
&\leq (1 - \| L_{t_0'} D \|)^{-1} \| y_0 \|_{L_2(t_0',p)} \\
&\leq (1 - \| L_{t_0'} \| \| D \|)^{-1} \| y_0 \|_{L_2(t_0',p)} .
\end{aligned}
$$

So the norm is uniformly bounded in $t_0' \geq t_0$.
Replacing $C(s)x(s; t_0', x_0)$ by $y(s)$ in (5.9) yields

$$
x(t; t_0', x_0) = \Phi(t, t_0')x_0 + \int_{t_0'}^{t} \Phi(t,s)B(s)D(s)y(s)ds.
$$

Similar estimates to these used in (i) show that the input to state map

$$\begin{array}{rccl} M_{t_0'} & : \; L_2(t_0', m) & \to & L_2(t_0', n) \\ & u(\cdot) & \mapsto & (t \mapsto \int_{t_0'}^{t} \Phi(t,s)B(s)u(s)ds) \end{array} \tag{5.10}$$

is uniformly bounded in $t_0' \geq t_0$. Hence (5.7) is satisfied and a similar estimate as in (i) applied to (5.9) yields (5.8).

(iv) Let $\mu > 0$ be the common period of A, B, C. The right shift S_μ

$$\begin{array}{rccl} S_\mu & : \; L_2(t_0, r) & \to & L_2(t_0 + \mu, r) \\ & v(t) & \mapsto & v(t - \mu) \end{array}$$

is an isometry. Now $\Phi(t + \mu, s + \mu) = \Phi(t, s)$, hence

$$\begin{aligned} (S_\mu \circ L_{t_0} u)(t) &= \int_{t_0}^{t-\mu} C(t-\mu)\Phi(t-\mu, s)B(s)u(s)ds \\ &= \int_{t_0}^{t-\mu} C(t)\Phi(t, s+\mu)B(s)u(s)ds \\ &= \int_{t_0+\mu}^{t} C(t)\Phi(t,\tau)B(\tau-\mu)u(\tau-\mu)d\tau \\ &= \int_{t_0+\mu}^{t} C(t)\Phi(t,\tau)B(\tau)S_\mu u(\tau)d\tau \\ &= (L_{t_0+\mu} \circ S_\mu u)(t). \end{aligned}$$

Hence the following diagram

$$\begin{tikzcd} L_2(t_0, m) \arrow[r, "L_{t_0}"] \arrow[d, "S_\mu"'] & L_2(t_0, p) \arrow[d, "S_\mu"] \\ L_2(t_0 + \mu, m) \arrow[r, "L_{t_0+\mu}"'] & L_2(t_0 + \mu, p) \end{tikzcd}$$

Figure 5.1.

commutes. This proves $\| L_{t_0} \| = \| L_{t_0+\mu} \|$ and the result follows since $t_0 \mapsto \| L_{t_0} \|$ is decreasing.

(v) The second and third inequalities in (5.4) follow from (ii) and (iii) and the last is a consequence of Lemma 4.3 (v). □

Throughout the remainder of this paper we use the notation

$$\ell(A;B,C) := \lim_{t_0\to\infty} \| L_{t_0}^{\Sigma} \|^{-1} . \tag{5.12}$$

As a consequence of (iii) we obtain the following roboustness result:

Corollary 5.3 Suppose $\sum = (A, B, C)$ and (5.2). If $D(\cdot) \in P\,C_b(0,\infty;\ \mathbb{C}^{m\times p})$ satisfies

$$\lim_{t_0\to\infty} \| D(\cdot)\,|_{[t_0,\infty)}\|_{L_\infty} < \ \ell(A;B,C) \tag{5.13}$$

then the perturbed system (5.6) is exponentially stable.

In the unstructed case it is known that perturbations $D \in PC_b(\mathbb{R}_+,\ \mathbb{C}^{m\times p})$ of norm $\| D(\cdot) \|_{L_\infty} < \frac{\omega}{M}$ (ω, M as in Proposition 5.2 (v)) do not destroy the exponential stability of the system (see *Coppel* (1978) Proposition 1.1). In view of (5.4), condition (5.13) is less conservative.

In contrast with time-invariant systems the following example shows that the inequality inequality

$$\ell(A;B,C) \ \leq \ r_{\mathbb{C}}(A;B,C) \tag{5.14}$$

is in general strict.

Example 5.4 Consider the scalar system

$$\dot{x}(t) = a(t)x(t), \quad t \geq 0$$

where $a(t) = -1 + k\alpha(t)$, $k \in \mathbb{R}$, $\alpha(\cdot) \in PC(\mathbb{R}_+,\ \mathbb{C})$ is periodic with period $3T, T = \ell n2$, given by

$$\alpha(t) = \begin{cases} 0 & t \in [3iT, (3i+1)T) \\ 1 & t \in [(3i+1)T, (3i+2)T) \ , \qquad i \in \mathbb{N}_0 \\ -1 & t \in [(3i+2)T, 3(i+1)T) \end{cases}$$

Let $\sum = (a, 1, 1)$ then in view of Example 3.11, Proposition 4.10 and Proposition 5.2 (iv), we have

$$-k_B(a) = r_{\mathbb{C}}(a) = 1 \quad \text{and} \quad \ell(A;1,1) = \| L_0^{\Sigma} \|^{-1} .$$

We will show that $\| L_0^{\Sigma} \|^{-1} < 1$.
Let $\beta(t) := k \int_0^t \alpha(\tau)d\tau$ and $u(t) = e^{\beta(t)-2t}$. A straightforward calculation shows

$$\begin{aligned} \| L_0^{\Sigma} u \|^2 - \| u \|^2 &= \int_0^\infty \left(\left| \int_0^t e^{-(t-s)+\beta(t)-\beta(s)} e^{\beta(s)-2s} ds \right|^2 - e^{2\beta(t)-4t}\right) dt \\ &= \int_0^\infty (e^{2\beta(t)}[e^{-2t} - e^{-t}]^2 - e^{2\beta(t)-4t}) dt \\ &= \int_0^T e^{2t}(1 - 2e^{-t})dt + \int_T^\infty e^{2\beta(t)-2t}(1 - 2e^{-t})dt. \end{aligned}$$

Since $1 - 2e^{-t} > 0$ for $t > T$ one can choose k so that the right hand side is positive.

Equality holds in (5.14) if the system $\sum$ is asymptotically or integrally equivalent to a time-invariant system. To prove this we need the following proposition

Proposition 5.5 Suppose that $\sum = (A, B, C)$ satisfies (5.2) and let $\dot{x}(t) = \hat{A}(t)x(t)$ be asymptotically or integrally equivalent to $\dot{x}(t) = A(t)x(t)$. Then for $\hat{\sum} = (\hat{A}, B, C)$

$$\lim_{t_0\to\infty} \| L^{\Sigma}_{t_0} - L^{\hat{\Sigma}}_{t_0} \| = 0. \tag{5.15}$$

In particular

$$\ell(A; B, C) = \ell(\hat{A}; B, C). \tag{5.16}$$

Proof: If $A(\cdot)$ generates $\Phi(\cdot,\cdot)$ and $\hat{A}(\cdot)$ generates $\hat{\Phi}(\cdot,\cdot)$, there exist $M, \omega > 0$ such that

$$\| \Phi(t,s) \| \leq Me^{-\omega(t-s)}, \ \| \hat{\Phi}(t,s) \| \leq Me^{-\omega(t-s)}, \quad \text{for all} \quad t \geq s \geq 0$$

(since $k_B(A) = k_B(\hat{A})$) . Set $\Delta(t) = \hat{A}(t) - A(t)$, then

$$\hat{\Phi}(t,s) = \Phi(t,s) + \int_s^t \Phi(t,\tau)\Delta(\tau)\hat{\Phi}(\tau,s)d\tau, \quad \text{for all} \quad t \geq s \geq 0$$

and so

$$\begin{aligned} \| \hat{\Phi}(t,s) - \Phi(t,s) \| &\leq \int_s^t Me^{-\omega(t-\tau)} \| \Delta(\tau) \| Me^{-\omega(\tau-s)}d\tau \\ &= M^2e^{-\omega(t-s)} \int_s^t \| \Delta(\tau) \| d\tau. \end{aligned}$$

Now let $u(\cdot) \in L_2(t_0, m)$, $\| u \| = 1$ then

$$\begin{aligned} \| (L^{\Sigma}_{t_0} - L^{\hat{\Sigma}}_{t_0})u \|^2 &= \int_{t_0}^{\infty} \| \int_{t_0}^t C(t)[\Phi(t,s) - \hat{\Phi}(t,s)]B(s)u(s)ds \|^2 \, dt \\ &\leq \| C \|^2_{L_\infty} \| B \|^2_{L_\infty} \int_{t_0}^{\infty} [\int_{t_0}^t M^2e^{-\omega(t-s)} \int_s^t \| \Delta(\tau) \| d\tau \, \| u(s) \| \, ds]^2 dt. \end{aligned}$$

In the case of integral equivalence this yields for $f(t) := e^{-\omega t}$

$$\begin{aligned} \| (L^{\Sigma}_{t_0} - L^{\hat{\Sigma}}_{t_0})u \|^2 &\leq K(\int_{t_0}^{\infty} \| \Delta(\tau) \| d\tau)^2 \int_0^{\infty} [\int_0^t e^{-\omega(t-s)} \| u(t_0+s) \| ds]^2 dt \\ &\leq K(\int_{t_0}^{\infty} \| \Delta(\tau) \| d\tau)^2 \| f \|^2_{L_1(0,1)} \| u \|^2_{L_2(t_0,m)} \\ &\leq \frac{K}{\omega^2}(\int_{t_0}^{\infty} \| \Delta(\tau) \| d\tau)^2 \| u \|^2_{L_2(t_0,m)} \end{aligned}$$

In the case of asymptotic equivalence, we have for $g(t) := te^{-\omega t}$

$$\begin{aligned} \| (L^{\Sigma}_{t_0} - L^{\hat{\Sigma}}_{t_0})u \|^2 &\leq K \sup_{\tau\geq t_0} \| \Delta(\tau) \|^2 \int_0^{\infty} [\int_0^t e^{-\omega(t-s)}(t-s) \| u(t_0+s) \| ds]^2 dt \\ &\leq K \sup_{\tau\geq t_0} \| \Delta(\tau) \|^2 \cdot \| g \|^2_{L_1(0,1)} \cdot \| u \|^2_{L_2(t_0,m)} \\ &\leq \frac{K}{\omega^4} \sup_{\tau\geq t_0} \| \Delta(\tau) \|^2 \cdot \| u \|^2_{L_2(t_0,m)} \end{aligned}$$

and this time the result follows since $\sup_{\tau\geq t_0} \| \Delta(\tau) \|^2 \to 0$ as $t_0 \to \infty$. □

Proposition 4.5 and 5.5 show that asymptotically or integrally equivalent systems have the same stability radius and the same limit $\ell(A; B, C)$. Hence Propositions 5.1 and 5.5 imply

Corollary 5.6 Suppose $\sum = (A, B, C)$ satisfies (5.2) and B, C are constant matrices. If $\dot{x}(t) = A(t)x(t)$ is asymptotically or integrally equivalent to a time-invariant $\dot{x}(t) = A_0 x(t)$, then

$$r_{\mathbb{C}}(A; B, C) = \ell(A; B, C) = r_{\mathbb{C}}(A_0; B, C) = [\max_{\omega \in \mathbb{R}} \| C(i\omega I - A_0)^{-1} B \|]^{-1}.$$

It is clear from the definition of $L_{t_0}^{\Sigma}$ that this operator is invariant with respect to Bohl transformations if the transformation is applied not only to $A(\cdot)$ but also to $B(\cdot)$ and $C(\cdot)$:

$$L_{t_0}^{\Sigma} = L_{t_0}^{\Sigma_T}, \quad t_0 \geq 0 \quad \text{for } \Sigma_T = (TAT^{-1} - \dot{T}T^{-1}, TB, CT^{-1}).$$

However contrary to the Bohl exponent and the stability radius, $\ell(A; B, C)$ is not invariant when scalar Bohl transformations are applied to $A(\cdot)$ alone. In fact applying Proposition 3.12 this is demonstrated by Example 5.4.

In order to fill the gap between $\ell(A; B, C)$ and $r_{\mathbb{C}}(A; B, C)$ one might try to use scalar Bohl transformation θ and consider $\sum_\theta = (A - \theta^{-1}\dot{\theta} I_n, B, C)$. Then $r_{\mathbb{C}}(A; B, C) = r_{\mathbb{C}}(A - \theta^{-1}\dot{\theta} I_n; B, C)$ and it is easy to see that $L_{t_0}^{\Sigma_\theta} = \theta^{-1} L_{t_0}^{\Sigma} \theta$. Unfortunately we have not been able to prove or disprove the following

Conjecture 5.7 Suppose (5.2) and $\sum = (A, B, C)$, then

$$r_{\mathbb{C}}(A; B, C) \;=\; \sup\{\ell(A - \theta^{-1}\dot{\theta} I_n; B, C); \theta \quad \text{a scalar Bohl transformation}\}$$

By Proposition 3.12 the conjecture holds true for scalar systems.

This section is concluded with an alternative interpretation of the perturbation operator. From a control theoretic viewpoint $L_{t_0}^{\Sigma}$ may be thought of as the *input-output operator* of the system

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \quad , \quad x(t_0) = 0 \\ y(t) &= C(t)x(t) \quad , \quad t \geq t_0 \end{aligned} \tag{5.17}$$

since $y(t) = (L_{t_0}^{\Sigma} u)(t)$. Formally (4.1) may be interpreted as a closed loop system obtained by applying the time-varying feedback $u(t) = D(t)y(t)$ to (5.17).

If the triple $\sum = (A, B, C)$ is such that $k_B(A) < 0$ (*internal stability*) then by Proposition 5.2(i) the input-output operator $L_{t_0}^{\Sigma}$ is bounded (*external stability*).
Several authors, for example *Anderson* (1972), *Anderson and Moore* (1969), *Brockett* (1970), *Megan* (1976a), *Silverman and Anderson* (1968), have investigated the relationship between internal and external stability. Clearly one cannot expect external stability to imply internal stability without additional assumptions concerning the interaction between the input $u(\cdot)$, output $y(\cdot)$ and state $x(\cdot)$ in (5.17). One has to impose uniform controllability (see Definition 1.2.8) and the following definition.

Definition 5.8 The system (5.17) is said to be *uniformly observable* if for some $\beta_0, \beta_1, \sigma > 0$

$$\beta_0 I_n \leq \int_{t-\sigma}^{t} \Phi^*(s, t-\sigma) C^*(s) C(s) \Phi(s, t-\sigma) ds \leq \beta_1 I_n \qquad \text{for all } t \geq \sigma.$$

(here $\leq$ denotes the order relation between Hermitian matrices.)

The following proposition can be deduced from a result of *Anderson* (1972)

Proposition 5.9 Suppose the system $\sum = (A, B, C)$ with B, C bounded is uniformly controllable and uniformly observable. Then the following are equivalent

(i) $\dot{x} = A(t)x(t)$ is exponentially stable

(ii) $L_{t_0}^{\Sigma} : L_2(t_0, m) \to L_2(t_0, p)$ is bounded independent of $t_0 \geq 0$.

Using a result of *Megan* (1976a) the characterization in Proposition 5.9 can be extended.

Proposition 5.10 Let $q \in [2, \infty]$ and $\sum = (A, B, C)$ satisfy the conditions of Proposition 5.9. Then $k_B(A) < 0$ if and only if the operator

$$\begin{array}{rcccl} L_{t_0,q}^{\Sigma} & : & L_q(t_0, m) & \to & L_q(t_0, p) \\ & & u(\cdot) & \mapsto & (t \mapsto \int_{t_0}^{t} C(t)\Phi(t, s)B(s)u(s)ds) \end{array}$$

is bounded uniformly in $t_0 \geq 0$.

Proof: Necessity follows as in Proposition 5.2(i). To prove sufficiency note that in *Megan* (1976a) Corollary 5.3 it is shown that boundedness of the *input-state operator*

$$\begin{array}{rcccl} \tilde{M}_{t_0} & : & L_2(t_0, m) & \to & L_\infty(t_0, m) \\ & & u(\cdot) & \mapsto & (t \mapsto \int_{t_0}^{t} \Phi(t, s)B(s)u(s)ds) \end{array}$$

implies $k_B(A) < 0$. Therefore it remains to prove

$$L_{t_0}^{\Sigma} \text{ is bounded } \quad \Rightarrow \quad \tilde{M}_{t_0} \text{ is bounded .}$$

Suppose $\tilde{M}_{t_0}$ is not bounded. Then for $N > 0$ there exist $u(\cdot) \in L_2(t_0, m)$ with $\| u(\cdot) \|_{L_2} = 1$ and $t - \sigma \geq t_0$ such that $\| (\tilde{M}_{t_0} u)(t - \sigma) \| = \| x(t - \sigma) \| > N$, where $x(\cdot)$ solves

$$\dot{x}(t) = A(t)x(t) + B(t)\tilde{u}(t), \quad x(t_0) = 0, t \geq 0$$

$$\text{and } \tilde{u}(s) = \begin{cases} u(s) & , \quad s \in [t_0, t - \sigma] \\ 0 & , \quad s > t - \sigma \end{cases}$$

Uniform observability yields

$$\begin{aligned} N^2 \beta_0 &\leq \| x(t-\sigma)^2 \| \beta_0 \leq \int_{t-\sigma}^{t} \| C(s)\Phi(s, t-\sigma)x(t-\sigma) \|^2 ds \\ &= \int_{t-\sigma}^{t} \| y(s) \|^2 ds \leq \| y(\cdot) \|_{L_2(t_0,p)}^2 = \| (L_{t_0}^{\Sigma} \tilde{u})(\cdot) \|_{L_2(t_0,p)} \\ &\leq \| L_{t_0}^{\Sigma} \| . \end{aligned}$$

Thus $L_{t_0}^{\Sigma}$ is not bounded and the proof is complete. □

4.6 The associated parametrized differential Riccati equation

In this section we examine the parametrized *differential Riccati equation* $(\mathbf{DRE})_\rho$

$$\dot{P}(t) + A^*(t)P(t) + P(t)A(t) - \rho C^*(t)C(t) - P(t)B(t)B^*(t)P(t) = 0, \quad t \geq t_0, \quad \rho \in \mathbb{R}$$

associated with the system

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t), \quad x(t_0) = x_0 \in \mathbb{C}^n \\ y(t) &= C(t)x(t) \end{aligned} \tag{6.1}$$

Throughout this section we assume (5.2).
For time-invariant $\sum = (A, B, C)$ it has been shown in *Hinrichsen and Pritchard* (1986b) that the *algebraic Riccati equation* $(\mathbf{ARE})_\rho$

$$A^*P + PA - \rho C^*C - PBB^*P = 0$$

admits a Hermitian solution P if and only if $\rho \leq r_{\mathbb{C}}(A; B, C)$. Guided by this result we wish to determine the maximal ρ for which there exist bounded Hermitian solutions of $(\mathrm{DRE})_\rho$ on $[t_0, \infty)$. *Kalman* (1960) and *Reghis and Megan* (1977), among others, have studied differential Riccati equations, however their results cannot be applied to $(\mathrm{DRE})_\rho$ if $\rho > 0$.

We will proceed via the following *optimal control problem*:

$(\mathbf{OCP})_\rho$ Minimize the cost functional

$$J_\rho(x_0, [t_0, t_1),\ u(\cdot)) := \int_{t_0}^{t_1} [\| u(s) \|^2 - \rho \| y(s) \|^2] ds$$

for $u(\cdot) \in L_2(t_0, t_1;\ \mathbb{C}^m)$ subject to (6.1)

where $0 \leq t_0 \leq t_1 \leq \infty, x_0 \in \mathbb{C}^n$ and $\rho \in \mathbb{R}$. We begin by examining the *finite time* problem where $t_1 < \infty$. Since the optimal control is expected to be feedback we start with some lemmata on the cost of feedback controls $u(t) = -F(t)x(t)$. To describe these costs we need the following well-known lemma about differential Lyapunov equations.

Lemma 6.1 Let $\tilde{A}(\cdot), R(\cdot) \in PC([t_0, \infty);\ \mathbb{C}^{n\times n})$, $\tilde{\Phi}(\cdot, \cdot)$ be the transition matrix of $\dot{x}(t) = \tilde{A}(t)x(t)$.

(i) The unique solution of the differential Lyapunov equation

$$\dot{P}(t) + \tilde{A}^*(t)P(t) + P(t)\tilde{A}(t) + R(t) = 0, \qquad t \in [t_0, t_1] \tag{6.2}$$

with final value $P(t_1) = 0$, is given by

$$P(t) = \int_t^{t_1} \tilde{\Phi}^*(s, t)R(s)\tilde{\Phi}(s, t)ds, \qquad t \in [t_0, t_1].$$

(ii) If $\dot{x}(t) = \tilde{A}(t)x(t)$ is exponentially stable and $R(\cdot)$ is bounded, then

$$P(t) = \int_t^{\infty} \tilde{\Phi}^*(s, t)R(s)\tilde{\Phi}(s, t)ds$$

is the unique bounded solution of (6.2) on $[t_0, \infty)$.

Lemma 6.2 Suppose $F(\cdot) \in PC([t_0, t_1];\ \mathbb{C}^{m\times n})$, $t_1 < \infty$, $A_F(t) = A(t) - B(t)F(t)$ with transition matrix $\Phi_F(\cdot, \cdot)$ and let

$$u_F(t) = -F(t)x(t), \qquad t \in [t_0, t_1]$$

where $x(\cdot)$ satisfies

$$\dot{x}(t) = A_F(t)x(t), \qquad t \in [t_0, t_1], x(t_0) = x_0.$$

Then

$$J_\rho(x_0, [t_0, t_1), u_F(\cdot)) = < x_0, P_F(t_0)x_0 > \tag{6.3}$$

where

$$P_F(t) = \int_t^{t_1} \Phi_F^*(s,t)[F^*(s)F(s) - \rho C^*(s)C(s)]\Phi_F(s,t)ds, \quad t \in [t_0, t_1] \tag{6.4}$$

is the solution of the *differential Lyapunov equation*, $(\mathbf{DLE})_\rho$

$$\dot{P}(t) + A_F^*(t)P(t) + P(t)A_F(t) - \rho C^*(t)C(t) + F^*(t)F(t) = 0, \qquad t \in [t_0, t_1]$$

with final value $P(t_1) = 0$.

Proof: By (6.4) and the definition of J_ρ we obtain

$$\begin{aligned} < x_0, P_F(t_0)x_0 > &= \int_{t_0}^{t_1} [\| F(s)\Phi_F(s,t)x_0 \|^2 - \rho \| C(s)\Phi_F(s,t_0)x_0 \|^2] ds \\ &= \int_{t_0}^{t_1} [\| u_F(s) \|^2 - \rho \| y_F(s) \|^2] ds \\ &= J_\rho(x_0, [t_0, t_1), u_F(\cdot)). \end{aligned}$$

That P_F solves $(\mathrm{DLE})_\rho$ follows from Lemma 6.1(i) by setting $\tilde{A}(t) = A_F(t)$ and $R(t) = -\rho C^*(t)C(t) + F^*(t)F(t)$. □

Note the following relationsship between the differential Riccati equation $(\mathrm{DRE})_\rho$ and the differential Lyapunov equation $(\mathrm{DLE})_\rho$.

Remark 6.3 $P(\cdot)$ is a solution of $(\mathrm{DRE})_\rho$ on $[t_0, t_1]$ if and only if $P(\cdot)$ is a solution of $(\mathrm{DLE})_\rho$ on $[t_0, t_1]$ with $F(t) = B^*(t)P(t)$.

Our construction procedure for solutions of $(\mathrm{DRE})_\rho$ (cf. proof of Theorem 6.7) is based on this simple observation.

Lemma 6.4 Let $F(\cdot) \in PC([t_0, t_1], \mathbb{C}^{m \times n})$, $\overline{u}(\cdot) \in L_2(t_0, t_1; \mathbb{C}^m)$, $u_F(t) = -F(t)x(t)$, $t \in [t_0, t_1]$, where now

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)(u_F(t) + \overline{u}(t)), & t \in [t_0, t_1] \\ &= A_F(t)x(t) + B(t)\overline{u}(t), & x(t_0) = x_0. \end{aligned} \tag{6.5}$$

If $u(t) = u_F(t) + \overline{u}(t)$, $t \in [t_0, t_1]$, then

$$\begin{aligned} J_\rho(x_0, [t_0, t_1), u(\cdot)) = &< x_0, P_F(t_0)x_0 > \\ &+ \textstyle\int_{t_0}^{t_1} \| u(s) + B^*(s)P_F(s)x(s) \|^2 \, ds \\ &- \textstyle\int_{t_0}^{t_1} \| [F(s) - B^*(s)P_F(s)]x(s) \|^2 \, ds \end{aligned}$$

where $P_F(\cdot)$ is defined by (6.4).

Proof: Differentiation of $V(t) := < x(t), P_F(t)x(t) >$, $t \in [t_0, t_1]$, along the solution $x(\cdot)$ of (6.5) gives (we leave out the argument t)

$$\begin{aligned}
\dot{V} &= < A_F x + B\overline{u}, P_F x > + < x, \dot{P}_F x > + < x, P_F(A_F x + B\overline{u}) > \\
&= < B\overline{u}, P_F x > + < x, P_F B\overline{u} > + < x, (\rho C^* C - F^* F)x > \\
&= - \| u_F \|^2 + \rho \| Cx \|^2 + 2\, Re < \overline{u}, B^* P_F x > \\
&= - \| \overline{u} \|^2 - \| u \|^2 - 2\, Re < \overline{u}, u > + \rho \| Cx \|^2 + 2\, Re < \overline{u}, (B^* P_F - F)x > \\
&= - \| u \|^2 + \rho \| Cx \|^2 + \| u + B^* P_F x \|^2 - \| (B^* P_F - F)x \|^2
\end{aligned}$$

Integrating on $[t_0, t_1]$ and using $P_F(t_1) = 0$ yields

$$\begin{aligned}
- < x_0, P_F(t_0 x_0 > = &- J_\rho(x_0, [t_0, t_1), u(\cdot)) + \int_{t_0}^{t_1} \| u(s) + B^*(s)P_F(s)x(s) \|^2\, ds \\
&- \int_{t_0}^{t} \| [B^*(s)P_F(s) - F(s)]x(s) \|^2\, ds
\end{aligned}$$

from which the result follows. □

If $\rho \geq 0$ and $0 \leq t_0 < t_1 < t_2 \leq \infty$, then

$$\begin{aligned}
0 &\geq \inf_{u \in L_2(t_0, t_1;\, \mathbb{C}^m)} J_\rho(x_0, [t_0, t_1), u(\cdot)) \\
&\geq \inf_{u \in L_2(t_0, t_2;\, \mathbb{C}^m)} J_\rho(x_0, [t_0, t_2), u(\cdot))
\end{aligned} \tag{6.6}$$

whereas the converse inequalities hold if $\rho \leq 0$. These inequalities show that the minimal costs are finite over an arbitrary interval if they are finite over $[0, \infty)$.

Lemma 6.5

(i) $\inf_{u \in L_2(t_0, m)} J_\rho(0, [t_0, \infty), u(\cdot)) = 0 \quad \Leftrightarrow \quad \rho \leq \| L_{t_0}^\Sigma \|^{-2}$

(here by definition $\| L_{t_0}^\Sigma \|^{-2} = \infty$ if $\| L_{t_0}^\Sigma \| = 0$).

(ii) For every $\rho \in (-\infty, \| L_{t_0}^\Sigma \|^{-2})$ there exists a constant $c_\rho > 0$ such that

$$\inf_{u \in L_2(t, m)} J_\rho(x_0, [t, \infty), u(\cdot)) \geq -c_\rho \| x_0 \|^2 \quad \text{for all } t \geq t_0, x_0 \in \mathbb{R}^n \tag{6.7}$$

Proof: (i)

$$\inf_{u \in L_2(t_0, m)} J_\rho(0, [t_0, \infty), u(\cdot)) = 0 \quad \Leftrightarrow \quad [\| u \|^2 - \rho \| L_{t_0}^\Sigma u \|^2] \geq 0 \quad \text{for all } u \in L_2(t_0, m)$$

which proves the equivalence in (i).
(ii) We need only consider the case $\rho \in (0, \| L_{t_0}^\Sigma \|^{-2})$. Since

$$2\, Re < a, b > \leq \alpha \| a \|^2 + \alpha^{-1} \| b \|^2 \quad \text{for all } \alpha > 0, a, b \in L_2(t_0, p),$$

we have

$$\begin{aligned}
J_\rho(x_0, [t_0, \infty), u(\cdot)) &= \| u(\cdot) \|^2 - \rho \| (L_{t_0}^\Sigma u)(\cdot) + C(\cdot)\Phi(\cdot, t_0)x_0 \|^2 \\
&= \| u(\cdot) \|^2 - \rho \| C(\cdot)\Phi(\cdot, t_0)x_0 \|^2 - 2\rho\, Re < (L_{t_0}^\Sigma u)(\cdot), C(\cdot)\Phi(\cdot, t_0)x_0 > \\
&\geq \| u(\cdot) \|^2 - \rho(1 + \alpha) \| (L_{t_0}^\Sigma u)(\cdot) \|^2 - \rho(1 + \alpha^{-1}) \| C(\cdot)\Phi(\cdot, t_0)x_0 \|^2 .
\end{aligned}$$

For sufficciently small α

$$J_\rho(x_0,[t_0,\infty),u(\cdot)) \geq -\rho(1+\alpha^{-1}) \parallel C(\cdot)\Phi(\cdot,t_0)x_0 \parallel^2 .$$

Since $\dot{x}(t) = A(t)x(t)$ is exponentially stable, there exists $c > 0$ such that

$$\parallel C(\cdot)\Phi(\cdot,t_0)x_0 \parallel^2 \leq c \parallel x_0 \parallel^2 \quad \text{for all } t_0 \geq 0.$$

So we may take $c_\rho = \rho(1+\alpha^{-1})c$ to ensure (6.7) for t_0. The result for any $t \geq t_0$ follows since $\parallel L_t^\Sigma \parallel \leq \parallel L_{t_0}^\Sigma \parallel$. □

Lemma 6.6 Suppose $A_k(\cdot) \in PC(t_0,t_1;\ \mathbb{C}^{n\times n})$, $k \in \mathbb{N}$, $t_1 < \infty$ converges pointwise to $\tilde{A}(\cdot) \in PC(t_0,t_1;\ \mathbb{C}^{n\times n})$ on $[t_0,t_1]$, i.e.

$$\lim_{k\to\infty} \parallel A_k(t) - \tilde{A}(t) \parallel = 0 \qquad \text{for all } t \in [t_0,t_1]$$

and $\parallel A_k(t) \parallel < c$ for all $t \in [t_0,t_1]$, $k \in \mathbb{N}$. If $A_k(\cdot)$ generates $\Phi_k(\cdot,\cdot)$ and $\tilde{A}(\cdot)$ generates $\tilde{\Phi}(\cdot,\cdot)$, then for every $\varepsilon > 0$, there exists $k_0 \in \mathbb{N}$, such that

$$\parallel \Phi_k(t,s) - \tilde{\Phi}(t,s) \parallel < \varepsilon \qquad \text{for all } k \geq k_0,\ t_0 \leq s \leq t \leq t_1. \tag{6.8}$$

Proof: The proof is similar to the one of Lemma 2.2 in *Reghis and Megan* (1977). Put $\Delta_k(t,s) = \parallel \Phi_k(t,s) - \tilde{\Phi}(t,s) \parallel$. Since

$$\Phi_k(t,s) = I_n + \int_s^t A_k(\tau)\Phi_k(\tau,s)d\tau \quad \text{and} \quad \tilde{\Phi}(t,s) = I_n + \int_s^t \tilde{A}(\tau)\tilde{\Phi}(\tau,s)d\tau$$

one obtains for $\alpha := \max\{\parallel \tilde{\Phi}(\tau,s) \parallel \ | \ t_0 \leq s \leq \tau \leq t_1\}$

$$\begin{aligned}
\Delta_k(t,s) &= \parallel \int_s^t A_k(\tau)\Phi_k(\tau,s) - \tilde{A}(\tau)\tilde{\Phi}(\tau,s)d\tau \parallel \\
&\leq \parallel \int_s^t (A_k(\tau) - \tilde{A}(\tau))\tilde{\Phi}(\tau,s)d\tau \parallel + \int_s^t \parallel A_k(\tau) \parallel \Delta_k(\tau,s)d\tau \\
&\leq \alpha \int_s^t \parallel A_k(\tau) - \tilde{A}(\tau) \parallel d\tau - c\int_s^t \Delta_k(\tau,s)d\tau.
\end{aligned}$$

Since $\parallel A_k(t) - \tilde{A}(t) \parallel \to 0$ for all $t \in [t_0,t_1]$ by Lebesgue's dominated convergence theorem for every $\varepsilon > 0$ there exists $k_0 \in \mathbb{N}$ such that

$$\int_{t_0}^{t_1} \parallel A_k(\tau) - \tilde{A}(\tau) \parallel d\tau \ \leq \ \frac{\varepsilon}{\alpha}e^{-c(t_1-t_0)} \qquad \text{for all } k \geq k_0.$$

Hence

$$\Delta_k(t,s) \ \leq \ e^{-c(t_1,t_0)} \cdot \varepsilon + c\int_s^t \Delta_k(\tau,s)d\tau \qquad \text{for all } k \geq k_0$$

and by Gronwall's inequality

$$\Delta_k(t,s) \ \leq \ e^{-c(t_1-t_0)}\varepsilon\ e^{c(t-s)} \ \leq \ \varepsilon \quad \text{for all } k \geq k_0.$$

This proves the lemma. □

We are now in a position to solve the optimal control problem $(OCP)_\rho$ on finite intervals.

Theorem 6.7 Suppose $\rho < \| L_{t_0}^{\Sigma} \|^{-2}$, $0 \le t_0 < t_1 < \infty$. Then:

(i) There exists a (unique) Hermitian solution $P^{t_1}(\cdot) \in C^1(t_0, t_1;\ \mathbb{C}^{n \times n})$ of $(DRE)_\rho$ with $P^{t_1}(t_1) = 0$.

(ii) If $\rho \ge 0$ (resp. $\rho \le 0$) then $P^{t_1}(t)$ is nonpositive (resp. nonnegative) for all $t \in [t_0, t_1]$.

(iii) The minimal cost of $(OCP)_\rho$ is

$$\inf_{u \in L_2(t_0, t_1;\ \mathbb{C}^m)} J_\rho(x_0, [t_0, t_1), u(\cdot)) = < x_0, P^{t_1}(t_0) x_0 > . \tag{6.9}$$

(iv) The optimal control is given by

$$u^*(t) = -B^*(t) P^{t_1}(t) x(t).$$

Proof: Starting with $P_0(\cdot) \equiv 0$ we recursively define a sequence $P_k(\cdot) \in C^1(t_0, t_1;\ \mathbb{C}^{n \times n})$, $k \in \mathbb{N}$ by the following sequence of differential Lyapunov equations

$$\begin{aligned} \dot{P}_k(t) \quad &+ \quad A_{k-1}^* P_k(t) \quad + \quad P_k(t) A_{k-1}(t) - \rho C^*(t) C(t) \\ & \qquad\qquad\qquad + \quad P_{k-1}(t) B(t) B^*(t) P_{k-1}(t) = 0, \qquad t \in [t_0, t_1], \\ P_k(t_1) \quad &= \quad 0 \end{aligned} \tag{6.10}$$

where

$$A_{k-1}(t) = A(t) - B(t) B^*(t) P_{k-1}(t), \qquad t \in [t_0, t_1],\ k \ge 1.$$

We will show:

(a) $P^{t_1}(t) = \lim_{k \to \infty} P_k(t)$ exists for all $t \in [t_0, t_1]$

(b) $P^{t_1}(\cdot)$ is the unique Hermitian solution of $(DRE)_\rho$ on $[t_0, t_1]$ with $P^{t_1}(t_1) = 0$.
After establishing (a), (b) we have by Lemma 6.2 and Remark 6.3

$$J_\rho(x_0, [t_0, t_1), -B^*(\cdot) x(\cdot)) = < x_0, P^{t_1}(t_0) x_0 >$$

and applying Lemma 6.4 with $F(t) = B^*(t) P^{t_1}(t)$ yields

$$J_\rho(x_0, [t_0, t_1), u(\cdot)) = < x_0, P^{t_1}(t_0) x_0 > + \int_{t_0}^{t_1} \| u(s) - u^*(s) \|^2 \, ds.$$

This shows (iii) and (iv) so it remains to prove (a), (b) and (ii). Note that by (6.10), $P_k(t) = P_F(t)$ where $F(t) = B^*(t) P_{k-1}(t)$. Set

$$u_k(t) = -B^*(t) P_k(t) x_k(t), \ \ \bar{u}_k(t) = u_k(t) - u_{k-1}(t), \quad t \in [t_0, t_1], k \ge 1$$

where $x_k(\cdot)$ solves $\dot{x}_k(t) = A_k(t) x_k(t)$, $x_k(t_0') = x_0$ and $t_0' \in [t_0, t_1]$ is arbitrary. By Lemma 6.2 and Lemma 6.4

$$\begin{aligned} < x_0, P_{k+1}(t_0') x_0 > - < x_0, P_k(t_0') x_0 > \quad &= \quad J_\rho(x_0, [t_0', t_1), u_k(\cdot)) \ - < x_0, P_k(t_0') x_0 > \\ &= \quad - \int_{t_0'}^{t_1} \| [B^*(s) P_{k-1}(s) - B^*(s) P_k(s)] x_k(s) \|^2 \, ds \le 0 \end{aligned}$$

for all $k \le 1, t_0' \in [t_0, t_1]$. But by Lemma 6.5(ii)

$$< x_0, P_k(t_0') x_0 > \ \ge \inf_{u \in L_2(t_0', t_1; m)} J_\rho(x_0, [t_0', t_1), u(\cdot)) \ \ge \ -c_\rho \| x_0 \|^2 .$$

So $(P_k(t_0'))_{k\geq 1}$ is a decreasing sequence uniformly bounded from below and the limit

$$\lim_{k\to\infty} P_k(t_0') = P^{t_1}(t_0') = (P^{t_1}(t_0'))^*$$

exists for every $t_0' \in [t_0, t_1]$. This proves (a). Moreover

$$\lim_{k\to\infty} A_k(t) = A(t) - B(t)B^*(t)P^{t_1}(t) \qquad \text{for all } t \in [t_0, t_1] \tag{6.11}$$

and since $P_k(t), k \geq 1$ is monotonically decreasing and bounded from below we see that $\| A_k(t) \|$ is uniformly bounded on $[t_0, t_1]$. Thus by Lemma 6.6 $\Phi_k(\cdot,\cdot)$ converges uniformly on $[t_0, t_1]$ to $\Phi^{t_1}(\cdot,\cdot)$ the evolution operator generated by $A(\cdot) - B(\cdot)B^*(\cdot)P^{t_1}(\cdot)$. Next we apply Lebesgue's dominated convergence theorem to the sequence

$$P_k(t) = -\int_{t_0}^{t_1} \Phi_{k-1}^*(s,t)[\rho C^*(s)C(s) - P_{k-1}(s)B(s)B^*(s)P_{k-1}(s)]\Phi_{k-1}(s,t)ds$$

to obtain

$$P^{t_1}(t) = -\int_{t}^{t_1} \Phi_{t_1}^*(s,t)[\rho C^*(s)C(s) - P^{t_1}(s)B(s)B^*(s)P^{t_1}(s)]\Phi_{t_1}(s,t)ds$$

Thus $P^{t_1}(\cdot)$ satisfies $(DRE)_\rho$ on $[t_0, t_1]$ and $P^{t_1}(t_1) = 0$. The uniqueness of the solution $P^{t_1}(\cdot)$ of $(DRE)_\rho$ with $P^{t_1}(t_1) = 0$ follows from general theorems. This proves (a) and (b). Applying Lemma 6.2 and Remark 6.3 to the above equation yields

$$P^{t_1}(t) = -\int_{t}^{t_1} \Phi(s,t)[\rho C^*(s)C(s) + P^{t_1}(s)B(s)B^*(s)P^{t_1}(s)]\Phi(s,t)ds$$

from which (ii) is obvious. (Note that $P_{k+1}^{t_1}(t) \geq P_k^{t_1}(t)$ holds for $k \geq 1$ and not for $k = 0$, if $\rho < 0$.) This shows (ii) and completes the proof. □

Corollary 6.8 Suppose $\rho < \| L_{t_0}^{\Sigma} \|^{-2}$, $0 \leq t_0 < t_1 < t_2 < \infty$. Then

$$\begin{array}{lllll} P^{t_2}(t) & \leq & P^{t_1}(t) & \text{for all } t \in [t_0, t_1] & \text{if } \rho \geq 0 \\ P^{t_2}(t) & \geq & P^{t_1}(t) & \text{for all } t \in [t_0, t_1] & \text{if } \rho < 0. \end{array}$$

Proof: Follows from Theorem 6.7 and (6.6). □

We now proceed to examine solutions of $(DRE)_\rho$ on *infinite* intervals and relate them to the *infinite time* optimal control problem $(OCP)_\rho$, $t_1 = \infty$. The following lemma plays a key role.

Lemma 6.9 Suppose $t_0 \geq 0, \rho \in \mathbb{R}$, $u(\cdot) \in L_2(t_0, m)$ and $Q(\cdot) \in C^1(t_0, \infty;\ \mathbb{C}^{n\times n})$ is a bounded Hermitian solution of $(DRE)_\rho$. If $x(\cdot)$ solves

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad t \geq t_0, x(t_0) = x_0 \tag{6.12}$$

Then

$$J_\rho(x_0, [t_0, \infty), u(\cdot)) = \int_{t_0}^{\infty} \| u(s) + B^*(s)Q(s)x(s) \|^2 ds + < x_0, Q(t_0)x_0 > . \tag{6.13}$$

In particular,

$$< x_0, Q(t_0)x_0 > \ \leq \inf_{u\in L_2(t_0,m)} J_\rho(x_0, [t_0, \infty), u(\cdot)), \ \ x_0 \in \mathbb{C}^n. \tag{6.14}$$

Proof: Since $k_B(A) < 0$ we have $x(\cdot) \in L_2(t_0, n)$ and we show that $x(t) \to 0$ as $t \to \infty$. Given $\varepsilon > 0$, choose $t_1 > t_0$ such that $\| u(\cdot) \|_{L_2(t_1,m)} < \varepsilon$, then

$$\| x(t) \| \le \| \Phi(t,t_1)x(t_1) \| + \int_{t_1}^{t} \| \Phi(t,s)B(s)u(s) \| \, ds, \quad t \ge t_1.$$

Hence by the Schwarz inequality

$$\| x(t) \| \le Me^{-\omega(t-t_1)} \| x(t_1) \| + M \| B \|_{L_\infty} (2\omega)^{-\frac{1}{2}} \| u \|_{L_2(t_1,m)} .$$

Now

$$\begin{aligned} \tfrac{d}{dt} < x(t), Q(t)x(t) > &= \rho \| C(t)x(t) \|^2 + \| B^*(t)Q(t)x(t) \|^2 \\ &\quad +2 \, Re < B(t)u(t), Q(t)x(t) > \\ &= \rho \| C(t)x(t) \|^2 + \| u(t) + B^*(t)Q(t)x(t) \|^2 \\ &\quad - \| u(t) \|^2 . \end{aligned}$$

Integrating over $[t_0, t_1]$ and taking limits as $t_1 \to \infty$ yields (6.13).
Since (6.13) holds for all $u(\cdot) \in L_2(t_0, \infty)$, (6.14) follows. □

The above lemma yields immediately the following necessary condition for the existence of bounded Hermitian solutions of $(DRE)_\rho$.

Proposition 6.10 Suppose (5.2) and $t_0 \ge 0$. If $Q(\cdot) \in C^1(t_0, \infty;\ \mathbb{C}^{n\times n})$ is a bounded Hermitian solution of $(DRE)_\rho$ on $[t_0, \infty)$ then

$$\rho \le \| L_{t_0}^{\Sigma} \|^{-2} . \tag{6.15}$$

Proof: By (6.14), $0 \le J_\rho(0, [t_0, \infty), u(\cdot))$ for all $u \in L_2(t_0, m)$. This implies (6.15) by Lemma 6.5 (i). □

The following converse result is the main theorem of this section.

Theorem 6.11 Suppose (5.2), $\sum = (A, B, C)$ and $\rho < \| L_{t_0}^{\Sigma} \|^{-2}$, $t_0 \ge 0$. Then we have

(i) There exists a unique stabilizing bounded Hermitian solution

$$P^+(\cdot) \in C^1(t_0, \infty;\ \mathbb{C}^{n\times n}) \quad \text{of } (DRE)_\rho \text{ on } [t_0, \infty).$$

(ii) P^+ is maximal in the sense that, for any bounded Hermitian solution $Q(\cdot) \in C^1(t_0', \infty;\ \mathbb{C}^{n\times n})$

on $[t_0', \infty)$, $t_0' \ge t_0$,

$$Q(t) \le P^+(t) \quad \text{for all } t \ge t_0'.$$

(iii) The minimal costs are

$$\inf_{u \in L_2(t_0,m)} J_\rho(x_0, [t_0, \infty), u(\cdot)) = < x_0, P^+(t_0)x_0 > \tag{6.16}$$

and the optimal control is given by

$$u(t) = -B^*(t)P^+(t)x(t), \quad t \ge t_0 \tag{6.17}$$

where $x(\cdot)$ solves

$$\dot{x}(t) = (A(t) - B(t)B^*(t)P^+(t))x(t), \quad t \ge t_0, \quad x(t_0) = x_0. \tag{6.18}$$

(iv) If the system $\sum$ is uniformly observable and $\rho > 0$ (resp. $\rho < 0$) then

$$P^+(t) \le -\gamma \, I_n \quad (\text{resp. } P^+(t) > \gamma \, I_n) \tag{6.19}$$

for some $\gamma > 0$ and all $t \ge t_0$.

Proof: First, let $\rho \geq 0$. By Lemma 6.5 and Theorem 6.7 there exists $c_\rho > 0$ such that for all $t_1 > t_0$, $t \in [t_0, t_1]$

$$\begin{aligned}
-c_\rho \parallel x_0 \parallel^2 &\leq \inf_{u \in L_2(t,m)} J_\rho(x_0, [t, \infty), u(\cdot)) \\
&\leq \inf_{u \in L_2(t,t_1;\, \mathbb{C}^m)} J_\rho(x_0, [t, t_1), u(\cdot)) \\
&= \langle x_0, P^{t_1}(t) x_0 \rangle .
\end{aligned} \tag{6.20}$$

Thus $P^{t_1}(t)$ is bounded below and since by Corollary 6.8 it is monotonically decreasing we have that

$$P^+(t) = \lim_{t_1 \to \infty} P^{t_1}(t), \tag{6.21}$$

exists for all $t \in [t_0, \infty)$.
Similarly, if $\rho < 0$, exponential stability implies that for every $t_1 > t_0$, $t \in [t_0, t_1]$

$$\begin{aligned}
0 &\leq \langle x_0, P^{t_1}(t) x_0 \rangle \\
&= \inf_{u \in L_2(t,t_1;\, \mathbb{C}^m)} J_\rho(x_0, [t, t_1), u(\cdot)) \\
&\leq \inf_{u \in L_2(t,m)} J_\rho(x_0, [t, \infty), u(\cdot)) \\
&\leq J_\rho(x_0, [t, \infty), 0) < \infty.
\end{aligned}$$

Since $(P^{t_1}(t))$ is monotonically increasing in t_1, the limit (6.21) exists for all $t \in [t_1, \infty)$ in the case $\rho < 0$ as well.

In both cases, $P^{t_1}(\cdot)$ satisfies

$$\begin{aligned}
P^{t_1}(t) = P^{t_1}(t_0) - \int_{t_0}^{t} [A^*(s) P^{t_1}(s) &+ P^{t_1}(s) A(s) - \rho C^*(s) C(s) \\
&- P^{t_1}(s) B(s) B^*(s) P^{t_1}(s)] ds.
\end{aligned}$$

Taking limits (as $t_1 \to \infty$), yields

$$\begin{aligned}
P^+(t) = P^+(t_0) - \int_{t_0}^{t} [A^*(s) P^+(s) &+ P^+(s) A(s) - \rho C^*(s) C(s) \\
&- P^+(s) B(s) B^*(s) P^+(s)] ds
\end{aligned}$$

and differentiation shows that $P^+(\cdot) \in C^1(t_0, \infty;\ \mathbb{C}^{n \times n})$ is a bounded Hermitian solution of $(DRE)_\rho$ on $[t_0, \infty)$.

Before showing that $P^+(\cdot)$ is stabilizing we prove (iii).
If $Q(\cdot) \in C^1(t_0, \infty;\ \mathbb{C}^{n \times n})$ is a bounded Hermitian solution of $(DRE)_\rho$ and $A(\cdot) - B(\cdot) B^*(\cdot) Q(\cdot)$ is the generator of $\Phi_Q(\cdot, \cdot)$, then

$$\begin{aligned}
\frac{d}{ds}[\Phi_Q^*(s, t_0) Q(s) \Phi_Q(s, t_0)] &= \Phi_Q^*(s, t_0)[\rho C^*(s) C(s) \\
&\quad - Q(s) B(s) B^*(s) Q(s)] \Phi_Q(s, t_0).
\end{aligned}$$

Hence

$$\begin{aligned}
\langle x_0, Q(t_0) x_0 \rangle &= \langle \Phi_Q(t, t_0) x_0, Q(t) \Phi_Q(t, t_0) x_0 \rangle + \int_{t_0}^{t} \langle \Phi_Q(s, t_0) x_0, \\
&\quad [Q(s) B(s) B^*(s) Q(s) - \rho C^*(s) C(s)] \Phi_Q(s, t_0) x_0 \rangle ds
\end{aligned} \tag{6.22}$$

First we consider the case $\rho \leq 0$ for which $P^+(t) \geq 0$, $t \geq t_0$. The above equality with $Q(\cdot) = P^+(\cdot)$ yields

$$< x_0, P^+(t_0)x_0 > \;\geq\; J_\rho(x_0, [t_0,\infty), -B^*(\cdot)P^+(\cdot)\Phi_{P^+}(\cdot,t_0)x_0).$$

In particular $\hat{u}(\cdot) := -B^*(\cdot)P^+(\cdot)\Phi_{P^+}(\cdot,t_0)x_0 \in L_2(t_0,m)$ and applying (6.13) with $Q(\cdot) = P^+(\cdot)$ we find

$$\inf_{u\in L_2(t_0,m)} J_\rho(x_0,[t_0,\infty),u(\cdot)) \;=\; J_\rho(x_0,[t_0,\infty),\hat{u}(\cdot)) \;=< x_0, P^+(t_0)x_0 > .$$

The case $\rho > 0$ is more difficult. To do this we extend the finite time optimal control by 0 to $[t_0,\infty)$ and define $u_{t_1}(\cdot) \in L_2(t_0,m)$ by:

$$u_{t_1}(t) = \begin{cases} -B^*(t)P^{t_1}x_{t_1}(t) & for \qquad t_0 \leq t \leq t \\ 0 & for \qquad t_1 < t \end{cases}$$

where $x_{t_1}(\cdot)$ solves

$$\dot{x}(t) = A(t)x(t) + B(t)u_{t_1}(t), \qquad t \geq t_0,\ x(t_0) = x_0.$$

Then by Theorem 6.7

$$\begin{aligned} J_\rho(x_0,[t_0,\infty),u_{t_1}(\cdot)) \;&=\; \textstyle\int_{t_0}^{t_1}[\| u_{t_1}(s)\|^2 \;-\; \rho\,\| C(s)x_{t_1}(s)\|^2]ds \\ &\qquad -\; \textstyle\int_{t_1}^{\infty}\rho\,\| C(s)x_{t_1}(s)\|^2\,ds \\ &=\; < x_0, P^{t_1}(t_0x_0 > \;-\; \rho\textstyle\int_{t_1}^{\infty}\| C(s)x_{t_1}(s)\|^2\,ds \end{aligned} \tag{6.23}$$

By applying (6.14) to $P^+(\cdot)$

$$J_\rho(x_0,[t_0,\infty),u(\cdot)) \;\geq\; < x_0, P^+(t_0)x_0 > \qquad \text{for all } u \in L_2(t_0,m) \tag{6.24}$$

and so

$$\begin{aligned} \lim_{t_1\to\infty} \textstyle\int_{t_1}^{\infty}\| C(s)x_{t_1}(s)\|^2\,ds \;&=\; 0, \\ \lim_{t_1\to\infty} J_\rho(x_0,[t_0,\infty),u_{t_1}(\cdot)) \;&=\; < x_0, P^+(t_0)x_0 > . \end{aligned} \tag{6.25}$$

Now from (6.23) we have for every $\alpha > 0$

$$\begin{aligned} 0 \;&\geq\; < x_0, P^{t_1}(t_0)x_0 > \\ &\geq\; J_\rho(x_0,[t_0,\infty),u_{t_1}(\cdot)) \\ &=\; \int_{t_0}^{\infty}[\| u_{t_1}(s)\|^2 - \rho\,\| C(s)\Phi(s,t_0)x_0 + (L_{t_0}^{\Sigma}\,u_{t_1})\|^2]ds \\ &\geq\; (1-\rho(1+\alpha)\,\| L_{t_0}^{\Sigma}\|^2)\cdot\| u_{t_1}\|^2_{L_2(t_0,m)} - \rho(1+\alpha^{-1})\,\| C(\cdot)\Phi(\cdot,t_0)x_0\|^2_{L_2(t_0,p)} \end{aligned} \tag{6.26}$$

by the same estimate as that used in establishing Lemma 6.5. Choosing $\alpha > 0$ small enough we see there exists a constant K independent of t_0, so that for all $t_0 \geq 0$

$$\| u_{t_1}\|^2_{L_2(t_0,m)} \;\leq K\,\| x_0\|^2 . \tag{6.27}$$

Hence $\{u_{t_1}, t_1 \geq t_0\}$ is bounded in $L_2(t_0,m)$, so there exists a sequence $(u_{t_k})_{k\in\mathbb{N}}, t_k \to \infty$ which converges weakly to some $\hat{u}(\cdot) \in L_2(t_0,m)$. By (6.24) and (6.25), (u_{t_k}) is a mininmizing sequence. It is easy to see that J_ρ is strictly convex. Moreover it follows from the last inequality in (6.26) - which holds for arbitrary $u \in L_2(t_0,m)$ instead of u_{t_1} - that $u \mapsto J_\rho(x_0,[t_0,\infty),u(\cdot))$

is coercive. Hence (*Ekeland and Temam* (1976) p.35), $\hat{u}(\cdot)$ is the unique optimal control and the minimum cost is

$$J_\rho(x_0,[t_0,\infty),\hat{u}(\cdot)) = < x_0, P^+(t_0)x_0 >$$

Lemma 6.9 implies for $Q(\cdot) = P^+(\cdot)$

$$J_\rho(x_0,[t_0,\infty),\hat{u}(\cdot)) = \int_{t_0}^{\infty} \| \hat{u}(s) + B^*(s)P^+(s)x(s) \|^2 ds + < x_0, P^+(t_0)x_0 >$$

and so

$$\hat{u}(t) = -B^*(t)P^+(t)x(t), \qquad t \geq t_0. \tag{6.28}$$

To prove the uniqueness and maximality, assume that $Q(\cdot)$ is a bounded Hermitian solution of $(DRE)_\rho$ on $[t_0',\infty)$. Using Lemma 6.9 and (6.16) we obtain

$$< x_0, Q(t)x_0 > \leq \inf_{u\in L_2(t,m)} J_\rho(x_0,[t,\infty),u(\cdot)) = < x_0, P^+(t)x_0 >$$

for all $t \geq t_0'$ and all $x_0 \in \mathbf{C}^n$. Hence the maximality of $P^+(\cdot)$. Now assume that $Q(\cdot)$ is stabilizing, then for every $\hat{t}_0 \geq t_0'$ the feedback control $u(t) = -B^*(t)Q(t)x(t)$, $t \geq \hat{t}_0$ is in $L_2(\hat{t}_0,\infty;\ \mathbf{C}^m)$ and so by Lemma 6.9

$$J_\rho(x_0,[\hat{t}_0,\infty),u(\cdot)) = < x_0, Q(\hat{t}_0)x_0 > \leq < x_0, P^+(\hat{t}_0)x_0 > .$$

Hence by (6.16) uniqueness holds.
To prove the feedback system(6.18) is exponentially stable we note that by (6.16) when $\rho \leq 0$ and via (6.27) and the weak convergence when $\rho > 0$, we have

$$\| \hat{u} \|^2_{L_2(t_0,m)} \leq K \| x_0 \|^2$$

for some constant K. Then it is easy to show, that the solution $x(\cdot)$ of (6.18) satisfies $\| x(\cdot) \|^2_{L_2(t_0,n)} < \overline{K} \| x_0 \|^2$, with $\overline{K}$ independent of t_0. The exponential stabilization then follows from Theorem 3.7.
It only remains to prove (iv). Suppose that $\sum$ is uniformly observable and $\rho > 0$, then

$$\begin{aligned} < x_0, P^+(t)x_0 > &= \inf_{u\in L_2(t,m)} J_\rho(x_0,[t_0,\infty),u(\cdot)) \\ &\leq J_\rho(x_0,[t_0,\infty),0) = -\rho \int_t^\infty \| C(s)\Phi(s,t)x_0 \|^2 ds \\ &\leq -\rho\, \beta_0 \| x_0 \|^2 . \end{aligned}$$

The proof for the case $\rho < 0$ is more difficult. Using (6.22) with $Q(\cdot) \equiv P^+(\cdot)$, $t = \infty$, yields

$$\begin{aligned} < x_0, P^+(t_0)x_0 > = \int_{t_0}^{\infty} &< \Phi_{P^+}(s,t_0)x_0, [P^+(s)B(s)B^*(s)P^+(s) \\ &-\rho C^*(s)C(s)]\Phi_{P^+}(s,t_0)x_0 > ds \end{aligned} \tag{6.29}$$

where $\Phi_{P^+}(\cdot,\cdot)$ is generated by $A(\cdot) - B(\cdot)B^*(\cdot)P^+(\cdot)$ and hence

$$\Phi_{P^+}(t,s) = \Phi(t,s) - \int_s^t \Phi(t,\sigma)B(\sigma)B^*(\sigma)P^+(\sigma)\Phi_{P^+}(\sigma,s)d\sigma. \tag{6.30}$$

Given $\varepsilon > 0$, suppose there exist t, x_0 such that $< x_0, P^+(t)x_0 > \leq \varepsilon \| x_0 \|^2$, then from (6.29) with $t_0 = t$, since $\rho < 0$, we have

$$\textstyle\int_t^\infty \| B^*(s)P^+(s)\Phi_{P^+}(s,t)x_0 \|^2 ds \leq \varepsilon \| x_0 \|^2$$

$$\textstyle\int_t^\infty \| C(s)\Phi_{P^+}(s,t)x_0 \|^2 ds \leq \frac{\varepsilon}{|\rho|} \| x_0 \|^2 .$$

Let $M, \omega > 0$ be such that $\| \Phi(t,s) \| \leq Me^{-\omega(t-s)}$ for all $t \geq s \geq 0$. Then the first inequality and (6.30) imply

$$\| \Phi_{P^+}(\cdot,t)x_0 - \Phi(\cdot,t)x_0 \|^2_{L_2(t,n)} \leq \frac{M^2\varepsilon \| B(\cdot) \|^2_{L_\infty}}{\omega^2} \| x_0 \|^2 .$$

But then

$$\begin{aligned} \| C(\cdot)\Phi(\cdot,t)x_0 \|^2_{L_2(t,p)} &\leq 2[\| C(\cdot)(\Phi(\cdot,t) - \Phi_{P^+}(\cdot,t))x_0 \|^2_{L_2(t,p)} \\ &\quad + \| C(\cdot)\Phi_{P^+}(\cdot,t)x_0 \|^2_{L_2(t,p)}] \\ &\leq 2\varepsilon[\frac{M^2}{\omega^2} \| C(\cdot) \|^2_{L_\infty} \| B(\cdot) \|^2_{L_\infty} + \frac{1}{|\rho|}] \| x_0 \|^2 . \end{aligned}$$

For ε sufficiently small this contradicts the assumption that $\sum$ is uniformly observable and this completes the proof. □

Proposition 6.10 and Theorem 6.11, together, imply the following characterization of $\| L^\Sigma_{t_0} \|$ in terms of the solvability of $(DRE)_\rho$:

$$\| L^\Sigma_{t_0} \| = \sup\{\rho \in \mathbb{R}; (DRE)_\rho \text{ has a bounded Hermitian solution on } [t_0,\infty)\}. \tag{6.31}$$

More precisely, if $\rho < \| L^\Sigma_{t_0} \|^{-2}$, then $(DRE)_\rho$ possesses a bounded Hermitian solution on $[t_0,\infty)$ whereas for $\rho > \| L^\Sigma_{t_0} \|^{-2}$ there does not exist such a solution. However, there may exist solutions on some smaller interval $[t'_0,\infty)$, $t'_0 > t_0$. The following corollary shows that $\ell(A;B,C)$ is a tight upper bound for those $\rho \in \mathbb{R}$ for which there is a bounded Hermitian solution of $(DRE)_\rho$ on some interval unbounded to the right.

Corollary 6.12 Suppose (5.2). If $\rho < \ell(A;B,C)^2$ then there exists a bounded Hermitian solutions of $(DRE)_\rho$ on some interval $[t_0,\infty)$, $t_0 > 0$.

Remark 6.13 The above results are not applicable to the limiting parameters value $\rho^* = \| L^\Sigma_{t_0} \|^{-2}$ (resp. $\rho^* = \ell(A;B,C)^2$). In the time-invariant case it is known that $(ARE)_\rho$ has a Hermitian solution for $\rho^* = \| L_0 \|^{-2}$ but the corresponding closed loop system is no longer exponentially stable and there may not exist a solution of the corresponding optimal control problem $(OCP)_{\rho^*}$ (see *Hinrichsen and Pritchard* (1986b)). So the differential Riccati equation $(DRE)_\rho$ and the optimal control problem $(OCP)_\rho$ are decoupled at the parameter value $\rho^* = \| L_0 \|^{-2}$.

In the remainder of this section we show that if $\sum$ is uniformly controllable, under the conditions of Theorem 6.11, there exists a solution $P^-(\cdot)$ of $(DRE)_\rho$ on $[t_0+\sigma,\infty)$ such that the closed loop system $\dot{x}(t) = [A(t) - B(t)B^*(t)P^-(t)]x(t)$ is *completely unstable* (i.e. the adjoint system $\dot{x}(t) = -[A(t) - B(t)B^*(t)P^-(t)]^*x(t)$ is exponentially stable).

Proposition 6.14 Suppose (5.2), $\sum = (A,B,C)$, $\rho \leq \| L^\Sigma_{t_0} \|^{-2}$, $t_0 \geq 0$ and $\sum$ is uniformly controllable with controllability interval of length σ. Set

$$Y(t_0,t) = \int_{t_0}^t \Phi_+(t_0,s)B(s)B^*(s)\Phi^*_+(t_0,s)ds \quad \text{for } t \geq t_0 \geq 0$$

where $\Phi_+(\cdot,\cdot)$ denotes the evolution operator generated by $A_+(\cdot) = A(\cdot) - B(\cdot)B^*(\cdot)P^+(\cdot)$. Then

$$P^-(t) = P^+(t) - Y^{-1}(t_0,t) \quad \text{for } t \geq t_0 + \sigma$$

is a bounded Hermitian solution of the $(DRE)_\rho$ on $[t_o+\sigma,\infty)$ and the system

$$\dot{x}(t) = A_-(t)x(t), \quad A_-(t) := A(t) - B(t)B^*(t)P^-(t), \quad t \geq t_0+\sigma$$

is completely exponentially unstable.

Proof: By Lemma 3 in *Silverman and Anderson* (1968) it is clear that (A_+, B, C) is also uniformly completely controllable with the same length of the controllability interval σ. Thus, see *Coppel* (1978), there exist $\delta_0, \delta_1 > 0$ such that

$$-\delta_0 I_n \leq -Y^{-1}(t_0,t) \leq -\delta_1 I_n \qquad \text{for all } t \geq t_0+\sigma.$$

This proves that $P^-(\cdot)$ is bounded. Next we show that it solves the $(DRE)_\rho$. By differentiation of $Y(t_0,\cdot)$ we see

$$\dot{Y}(t_0,t) - A_+^*(t)Y(t_0,t) - Y(t_0,t)A_+(t) - B(s)B^*(s) = 0 \qquad \text{for } t \geq t_0 \tag{6.32}$$

Since $Y(t_0,t)$ is invertible for $t \geq t_0+\sigma$ it follows that $-Y^{-1}(t_0,t) = P^-(t) - P^+(t)$ solves

$$\dot{X}(t) + A_+^*(t)X(t) + X(t)A_+(t) - X(t)B(t)B^*(t)X(t) = 0 \qquad \text{for } t \geq t_0+\sigma.$$

Thus (leaving out the argument t):

$$\dot{P}^- + A^*P^- + P^-A - P^-BB^*P^- = \dot{P}^+ + A^*P^+ + P^+A - P^+BB^*P^+.$$

But P^+ solves the $(DRE)_\rho$ on $[t_0,\infty)$ and so the left hand side is equal to $\rho C^*(\cdot)C(\cdot)$. Therefore P^- is a solution of the $(DRE)_\rho$ on $[t_0+\sigma,\infty)$.
Using (6.32) it is easy to see that $Y(t_0,t)$ solves the following Lyapunov equations for $t \geq t_0+\sigma$

$$\begin{aligned}\dot{Y} &= [A - BB^*(P^- + Y^{-1})]Y + Y[A - BB^*(P^- + Y^{-1})]^* + BB^* \\ &= -(-A_-^*)^*Y - Y(-A_-^*) - BB^*.\end{aligned} \tag{6.33}$$

It now follows by applying Theorem 5.2 in *Megan*(1976a) to (6.33) that the system

$$\dot{x}(t) = -A_-^*(t)x(t), \quad t \geq t_0+\sigma$$

is exponentially stable. Hence the proof is complete. □

The following example will show that, in contrast to the time-invariant case, $P^-(\cdot)$ will *not* in general be a minimal solution of $(DRE)_\rho$ on $[t_0+\sigma,\infty)$.

Example 6.15 Consider

$$\begin{aligned}\dot{x}(t) &= -x(t) + u(t) \\ y(t) &= x(t)\end{aligned}$$

and the associated $(DRE)_\rho$

$$\dot{p}(t) - 2p(t) - \rho - p(t)^2 = 0.$$

Then $\| L_{t_0} \| = 1$ and

$$p(t) = \frac{k\, e^{e\alpha t}(1+\alpha) - (1-\alpha)}{1 - k\, e^{2\alpha t}}, \quad k \in \mathbb{R}, \alpha^2 = 1-\rho$$

whence

$$\begin{aligned}
p^+(t) &= -(1-\alpha), \quad k=0 \\
y(0,t) &= \int_0^t e^{-2\alpha s} ds = \frac{1-e^{-2\alpha t}}{2\alpha} \\
p^-(t) &= -(1-\alpha) - \frac{2\alpha}{1-e^{-2\alpha t}} = -\frac{(1-\alpha)e^{2\alpha t}-(1-\alpha)}{e^{2\alpha t}-1}, \quad k=1.
\end{aligned}$$

Now

$$p(t) \geq p^-(t) \quad \text{for } k \geq 1 \text{ or } k \leq 0, t > 0$$

but for t large

$$p(t) \leq p^-(t) \quad \text{for } 0 < k < 1.$$

4.7 Dependence of the solution $P_\rho^+(t)$ on the parameter ρ

In this section we examine properties of the maps

$$\begin{aligned}
P_\cdot^+(t) : (-\infty, \| L_{t_0}^\Sigma \|^{-2}) &\rightarrow \mathbb{C}^{n\times n} \qquad , t \geq t_0 \\
\rho &\mapsto P_\rho^+(t)
\end{aligned}$$

and analyze how the norm of the perturbation operator $L_{t_0}^{\Sigma_\rho}$ of the closed loop system $\sum_\rho = (A_\rho, B, C)$ changes with ρ ,

$$A_\rho(t) = A(t) - B(t)B^*(t)P_\rho^+(t), \quad t \geq t_0.$$

Let $\Phi_\rho(\cdot,\cdot)$ denote the evolution operator generated by $A_\rho(\cdot)$.

Proposition 7.1 Assume (4.2).

(i) The map $\rho \mapsto P_\rho^+(t)$ is differentiable on $(-\infty, \| L_{t_0}^\Sigma \|^{-2})$ for every $t \geq t_0$ and

$$\frac{d}{d\rho} P_\rho^+(t) = -\int_t^\infty \Phi_\rho^*(s,t)C^*(s)C(s)\Phi_\rho(s,t)ds.$$

(ii) If $\rho_1 \leq \rho_2 < \| L_{t_0}^\Sigma \|^{-2}$ then

$$P_{\rho_1}^+(t) \geq P_{\rho_2}^+(t) \qquad \text{for all } t \geq t_0.$$

(iii) If $\rho_1 \leq \rho < \| L_{t_0}^\Sigma \|^{-2}$, then $P_\rho^+(\cdot) - P_{\rho_1}^+(\cdot)$ is the maximal bounded Hermitian solution of the differential Riccati equation

$$\dot{X}(t) + A_{\rho_1}^*(t)X(t) + X(t)A_{\rho_1}(t) - (\rho-\rho_1)C^*(t)C(t) - X(t)B(t)B^*(t)X(t) = 0 \qquad (7.1)$$

on $[t_0, \infty)$.

(iv) If $\rho < \| L_{t_0}^\Sigma \|^{-2}$, then

$$\| L_{t_0}^{\Sigma_\rho} \|^{-2} = \| L_{t_0}^\Sigma \|^{-2} - \rho.$$

Proof: Subtracting $(DRE)_{\rho_1}$ form $(DRE)_\rho$, yields

$$\frac{d}{dt}(P_\rho^+ - P_{\rho_1}^+) + A_{\rho_1}^*(P_\rho^+ - P_{\rho_1}^+) + (P_\rho^+ - P_{\rho_1}^+)A_{\rho_1} - (\rho-\rho_1)C^*C - (P_\rho - P_{\rho_1})BB^*(P_\rho - P_{\rho_1}) = 0.$$

Hence by Lemma 6.1 (ii)

$$\begin{aligned}(P_\rho^+ - P_{\rho_1}^+)(t) &= -\int_t^\infty \Phi_{\rho_1}^*(s,t)[(\rho-\rho_1)C^*(s)C(s) \\ &\quad +(P_\rho^+ - P_{\rho_1}^+)(s)B(s)B^*(s)(P_\rho^+ - P_{\rho_1}^+)(s)]\Phi_{\rho_1}(s,t)ds \qquad (7.2)\\ &\le 0 \quad \text{if } \rho \ge \rho_1.\end{aligned}$$

Hence (ii) is proved. To prove (i) note that for $\rho \ne \rho_1$ (7.2) is equivalent to

$$\begin{aligned}\Delta_\rho(t) - (\rho-\rho_1)\int_t^\infty \Phi_{\rho_1}^*(s,t) \quad & \Delta_\rho(s)B(s)B^*(s)\Delta_\rho(s)\Phi_{\rho_1}(s,t)ds \\ &= \int_t^\infty \Phi_{\rho_1}^*(s,t)C^*(s)C(s)\Phi_{\rho_1}(s,t)ds \qquad (7.3)\end{aligned}$$

where

$$\Delta_\rho(t) = -\frac{P_\rho^+(t) - P_{\rho_1}^+(t)}{\rho-\rho_1}.$$

Now assume $\rho < \rho_1 <\| L_{t_0}^\Sigma \|^{-2}$. Since $k_B(A_{\rho_1}) < 0$ the right hand term in (7.3) is uniformly bounded in t and then by (ii) and (7.3)

$$0 \le \Delta_\rho(t) < \alpha\, I_n, \qquad t \ge t_0 \quad \text{for some real constant } \alpha > 0.$$

Hence

$$\lim_{\substack{\rho\to\rho_1\\ \rho<\rho_1}} -\Delta_\rho(t) = -\int_t^\infty \Phi_{\rho_1}^*(s,t)C^*(s)C(s)\Phi_{\rho_1}(s,t)ds.$$

Since this limit is continuous in ρ_1 it follows from *Kato* (1976) p. 494, that

$$\frac{d}{d\rho}P_\rho^+(t)\,|_{\rho=\rho_1} = -\int_t^\infty \Phi_{\rho_1}^*(s,t)C^*(s)C(s)\Phi_{\rho_1}(s,t)ds.$$

To prove (iii) note that since

$$\dot{x}(t) = [A_{\rho_1}(t) - B(t)B^*(t)(P_\rho^+ - P_{\rho_1}^+)(t)]x(t) = [A(t) - B(t)B^*(t)P_\rho^+(t)]x(t)$$

is exponentially stable, Theorem 6.11 applied to (7.1) yields that $(P_\rho^+ - P_{\rho_1}^+)(\cdot)$ is its maximal bounded Hermitian solution.
It remains to show (iv). Now since $(P_\rho^+ - P_{\rho_1}^+)(\cdot)$ solves (7.1) on $[t_0,\infty)$, it follows from Proposition 6.10 that

$$\rho - \rho_1 \le \| L_{t_0}^{\Sigma_{\rho_1}} \|^{-2} \qquad \text{for all } \rho \in [\rho_1, \| L_{t_0}^\Sigma \|^{-2}).$$

Hence

$$\| L_{t_0}^\Sigma \|^{-2} \le \| L_{t_0}^{\Sigma_{\rho_1}} \|^{-2} + \rho_1.$$

Now suppose that

$$\| L_{t_0}^\Sigma \|^{-2} + \varepsilon = \| L_{t_0}^{\Sigma_{\rho_1}} \|^{-2} + \rho_1 \qquad \text{for some } \varepsilon > 0.$$

Then by Theorem 6.11 there exists a bounded Hermitian solution $\overline{P}(\cdot) \in C^1(t_0, \infty;\ \mathbb{C}^{n\times n})$ of the differential Riccati equation

$$\dot{P} + A_{\rho_1}^* P + PA_{\rho_1} - (\| L_{t_0}^{\Sigma} \|^{-2} - \rho_1 + \frac{\varepsilon}{2})C^*C - PBB^*P = 0 \tag{7.4}$$

with the property that $\dot{x}(t) = [A_{\rho_1}(t) - B(t)B^*(t)\overline{P}(t)]x(t)$ is exponentially stable. Adding (7.4) to the $(DRE)_{\rho_1}$ shows that $(P_{\rho_1} + \overline{P})(\cdot) \in C^1(t_0, \infty;\ \mathbb{C}^{n\times n})$ is a bounded Hermitian solution of

$$\dot{X} + A^*X + XA - (\| L_{t_0}^{\sigma} \|^{-2} + \frac{\varepsilon}{2})C^*C - PBB^*P = 0$$

on $[t_0, \infty)$ with the property that

$$\begin{aligned} \dot{x}(t) &= [A(t) - B(t)B^*(t)(P_{\rho_1} + \overline{P})(t)]x(t) \\ &= [A_{\rho_1}(t) - B(t)B^*(t)\overline{P}(t)]x(t) \end{aligned}$$

is exponentially stable. This is a contradiction by Proposition 6.10. □

4.8 Nonlinear perturbations and robust Lyapunov functions

In this section we extend our robustness analysis to nonlinear perturbations of the form $\Delta(t) = B(t)N(C(t)x(t), t)$ so that the perturbed system is

$$\dot{x}(t) = A(t)x(t) + B(t)N(C(t)x(t), t), \qquad t \geq t_0,\ \ x(t_0) = x_0 \tag{8.1}$$

where (A, B, C) satisfies (5.2) and $N : \mathbb{R}^p \times \mathbb{R}_+ \to \mathbb{R}^m$ is continuously differentiable. We assume $N(0, t) = 0$ so that 0 is an equilibrium state of (8.1). Our aim is to determine conditions on the "norm" of the nonlinear perturbation such that exponential stability of (8.1) is preserved. To this end we need the following lemmata.

Lemma 8.1 Set $\sum^{\varepsilon} = (A + \varepsilon I_n, B, C)$. Then for all $\varepsilon \in [0, \varepsilon_0)$, $\varepsilon_0 < -k_B(A)$, there exists κ independent of ε such that

$$\| L_{t_0}^{\sum^{\varepsilon}} - L^{\sum_{t_0}} \|_{\mathcal{L}(L_2(t_0,m), L_2(t_0,p))} \leq \kappa\, \varepsilon. \tag{8.2}$$

Proof: Let $A(\cdot) + \varepsilon I_n$ generate $\Phi_\varepsilon(\cdot,\cdot)$, then $\Phi_\varepsilon(t,s) = e^{\varepsilon(t-s)}\Phi(t,s)$ and hence

$$\| \Phi_\varepsilon(t, t_0) - \Phi(t, t_0) \| \leq | 1 - e^{\varepsilon(t-t_0)} | \cdot \| \Phi(t, t_0) \| .$$

Since there exist M and $\omega \in (\varepsilon_0, -k_B(A))$ such that $\| \Phi(t, t_0) \| \leq Me^{-\omega(t-t_0)}$, $t \geq t_0$, we obtain

$$\begin{aligned} \| \Phi_\varepsilon(t, t_0) - \Phi(t, t_0) \| &\leq M\, | 1 - e^{\varepsilon(t-t_0)} |\, e^{-\omega(t-t_0)} \\ &\leq N\, \varepsilon\, e^{-(\omega-\varepsilon)(t-t_0)} \\ &\leq N\, \varepsilon\, e^{-\omega'(t-t_0)} \end{aligned}$$

for $\omega' := \omega - \varepsilon_0$ and some $N > 0$. Then for $f(t) := e^{-\omega' t}$ and every $u(\cdot) \in L_2(t_0, \infty;\ \mathbb{C}^m)$

$$\begin{aligned} \| (L_{t_0}^{\Sigma^{\varepsilon}} &- L_{t_0}^{\Sigma})(u(\cdot)) \|^2_{L_2(t_0,\infty;\ \mathbb{C}^p)} \\ &= \int_{t_0}^{\infty} \| \int_{t_0}^{t} C(t)[\Phi_\varepsilon(t,s) - \Phi(t,s)]B(s)u(s)ds \|^2\, dt \\ &\leq \| C(\cdot) \|^2_{L_\infty} \| B(\cdot) \|^2_{L_\infty}\ N^2\varepsilon^2 \int_{t_0}^{\infty} (f\star \| u \|)^2(t)dt \\ &\leq \| C(\cdot) \|^2_{L_\infty} \| B(\cdot) \|^2_{L_\infty}\ N^2\, \varepsilon^2\omega'^{-2}\, \| u(\cdot) \|^2_{L_2(t_0,\infty;\ \mathbb{C}^m)} \end{aligned}$$

which proves the lemma. □

By P_ρ^ε we denote the maximal bounded Hermitian solution of the Riccati equation

$$\dot{P}(t) + [A(t) + \varepsilon\, I_n]^* P(t) + P(t)[A(t) + \varepsilon\, I_n] - \rho C^*(t)C(t) - P(t)B(t)B^*(t)P(t) = 0. \quad (8.3)$$

Lemma 8.2 Suppose $0 \le \rho < \| L_{t_0}^{\Sigma} \|^{-2}$. Then there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \le \varepsilon_0$

$$P_\rho^\varepsilon(\cdot) \text{ exists on } [t_0, \infty) \quad (8.4)$$

$$P_\rho^{\varepsilon_0}(t) \le P_\rho^\varepsilon(t) \le P_\rho(t) \quad \text{for all } t \ge t_0 \quad (8.5)$$

$$\| P_\rho^\varepsilon(t_0) - P_\rho(t_0) \| \le \kappa\, \varepsilon \quad \text{for some } \kappa \text{ independent of } \varepsilon. \quad (8.6)$$

Proof: For ε_0 sufficiently small $\delta < \| L_{t_0}^{\Sigma^\varepsilon} \|^{-2}$ by Lemma 8.1, hence (8.4) follows from Theorem 6.11. If $\overline{P} := P_\rho^\varepsilon(\cdot) - P_\rho(\cdot)$, then

$$\dot{\overline{P}} + [A + \varepsilon\, I_n - BB^* P_\rho]^* \overline{P} + \overline{P}[A + \varepsilon\, I_n - BB^* P_\rho] + 2\varepsilon P_\rho - \overline{P}BB^*\overline{P} = 0.$$

Since $k_B(A + \varepsilon I_n - BB^* P_\rho) < 0$ for all $\varepsilon \le \varepsilon_0$ and ε_0 sufficiently small, we obtain by Lemma 6.1

$$\overline{P}(t) = \int_t^\infty \Phi_\varepsilon^*(s,t)[2\varepsilon\, P_\rho(s) - \overline{P}(s)B(s)B^*(s)\overline{P}(s)]\Phi_\varepsilon(s,t)ds \quad (8.7)$$

where $[A + \varepsilon\, I_n - BB^* P_\rho](\cdot)$ generates $\Phi_\varepsilon(\cdot,\cdot)$. Now $P_\rho(s) \le 0$ for $s \ge t_0$ and thus $\overline{P}(t) \le 0$ for $t \ge t_0$. This proves the right inequality in (8.5). The left inequality is established in a similar way.

By

$$\overline{P}(t) \;\le\; 2\varepsilon \int_t^\infty \| \Phi_\varepsilon(s,t) \|^2 \| P_\rho(s) \|\, ds \quad (8.8)$$

and since $P_\rho(\cdot)$ is bounded it remains to show the existence of some $\kappa' > 0$ such that

$$\int_t^\infty \| \Phi_\varepsilon(s,t) \|^2\, ds \le \kappa' \quad \text{for all } t \ge t_0, \varepsilon \le \varepsilon_0. \quad (8.9)$$

If $\hat{\Phi}(\cdot,\cdot)$ is generated by $[A - BB^* P_\rho](\cdot)$ and $\| \hat{\Phi}(s,t) \| \le \hat{M} e^{-\hat{\omega}(s-t)}$ for some $\hat{M}, \hat{\omega} > 0$ then Variation-of-Constants yields

$$\| \Phi_\varepsilon(s,t) \| \le \hat{M}\, e^{-\hat{\omega}(s-t)} + \int_t^s \hat{M}\, e^{-\hat{\omega}(s-t)}\varepsilon\, \| \Phi_\varepsilon(\tau,t) \|\, d\tau.$$

Multiplying this inequality by $e^{\hat{\omega}s}$ and applying Gronwall's Lemma gives

$$\| e^{\hat{\omega}s} \Phi_\varepsilon(s,t) \| \le \hat{M}\, e^{\hat{\omega}t}\, e^{\varepsilon \hat{M}(s-t)}.$$

Thus (8.8) holds for ε_0 sufficiently small and the proof is complete. □

If a global bound for the nonlinear perturbation is known we obtain the following result.

Theorem 8.3 Suppose (5.2), $\sum = (A,B,C)$, $t_0 \ge 0$ and for some $\gamma < \| L_{t_0}^{\Sigma} \|^{-1}$

$$\| N(y,t) \| \le \gamma \| y \| \quad \text{for all } t \ge t_0, y \in \mathbb{C}^p. \quad (8.10)$$

Then the origin is globally exponentially stable for the system (8.1).

Proof: Choose $\rho \in (\gamma^2, \| L_{t_0}^{\Sigma} \|^{-2})$ and $\varepsilon \geq 0$ sufficiently small such that $P_\rho^\varepsilon(\cdot)$ exists on $[t_0, \infty)$. (Note that, by assumption (8.10), the right hand side of (8.1) is linearly bounded and so the solutions of (8.1) exist on $[t_0, \infty)$. Consider the functional

$$V(t,x) = - < x, P_\rho^\varepsilon(t)x >, \quad t \geq t_0, x \in \mathbb{C}^n.$$

Its derivative along any solution $x(\cdot)$ of (8.1) is

$$\begin{aligned}\dot V(t,x(t)) &= - 2\varepsilon V(t,x(t)) - \rho \| C(t)x(t) \|^2 - \| B^*(t)P_\rho^\varepsilon(t)x(t) \|^2 \\ &\quad - 2\, Re < P_\rho^\varepsilon(t)x(t), B(t)N(C(t)x(t),t) > \\ &= - 2\varepsilon V(t,x(t)) - \| B^*(t)P_\rho^\varepsilon(t)x(t) + N(C(t)x(t),t) \|^2 \\ &\quad - [\rho \| C(t)x(t) \|^2 - \| N(C(t)x(t),t) \|^2].\end{aligned}$$

Hence

$$\dot V(t,x(t)) \leq -2\,\varepsilon V(t,x(t)) - \delta \| C(t)x(t) \|^2, \quad t \geq t_0$$

where $\delta = \rho - \gamma^2$. Integrating yields

$$V(t_1,x(t_1))e^{2\varepsilon t_1} - V(t_0,x(t_0))e^{2\varepsilon t_0} \leq -\delta \int_{t_0}^{t_1} e^{2\varepsilon t} \| C(t)x(t) \|^2 \, dt$$

for all $t_1 > t_0$ and since $V(t_1, x(t_1)) \geq 0$

$$\int_{t_0}^{\infty} e^{2\varepsilon(t-t_0)} \| C(t)x(t) \|^2 \, dt \leq -\delta^{-1} < x_0, P_\rho^\varepsilon(t_0)x_0 > . \tag{8.11}$$

Now if $A(\cdot)$ generates $\Phi(\cdot,\cdot)$

$$\| x(t) \| \leq \| \Phi(t,t_0)x_0 \| + \int_{t_0}^{t} \| \Phi(t,s)B(s)N(C(s)x(s),s) \| \, ds.$$

But there exists $M, \omega > 0$ such that $\| \Phi(t,s) \| \leq Me^{-\omega(t-s)}, t \geq s$. Hence

$$\begin{aligned}e^{\varepsilon(t-t_0)} \| x(t) \| &\leq Me^{-(\omega-\varepsilon)(t-t_0)} \| x_0 \| \\ &\quad +\gamma \int_{t_0}^{t} M \| B \|_{L_\infty} e^{-(\omega-\varepsilon)(t-s)} e^{\varepsilon(s-t_0)} \| C(s)x(s) \| \, ds \\ &\geq Me^{-(\omega-\varepsilon)(t-t_0)} \| x_0 \| \\ &\quad +\gamma M \| B \|_{L_\infty} [\int_{t_0}^{t} e^{-2(\omega-\varepsilon)(t-s)} ds]^{\frac{1}{2}} [\int_{t_0}^{t} e^{2\varepsilon(s-t_0)} \| C(s)x(s) \|^2 \, ds]^{\frac{1}{2}}.\end{aligned}$$

So, by (8.10), there exists a constant $K > 0$ such that

$$\| x(t) \| \leq Ke^{-\varepsilon(t-t_0)} \| x(t_0) \| \qquad \text{for all } t \geq t_0 \geq 0$$

This concludes the proof. □

If condition (8.9) is required only locally then the following local version of Theorem 8.3 is available. However, one needs to know $P_\rho(t)$ for all $t \geq t_0$.

Theorem 8.4 Suppose (5.2) , $\sum = (A,B,C)$, $t_0 \geq 0$ and $0 < \rho < \| L_{t_0}^{\Sigma} \|^{-2}$. If there exist $d > 0$ and $\delta > 0$, such that

$$- < x, P_\rho(t)x > < d \;\Rightarrow\; \| N(C(t)x,t) \|^2 \leq (\rho - \delta) \| C(t)x \|^2, \quad t \geq t_0 \tag{8.12}$$

then $\{x_0 \in \mathbb{C}^n \mid - < x_0, P_\rho^\varepsilon(t_0)x_0 > < d\}$ is a region of exponential stability at t_0, for all ε sufficiently small.

Proof: Suppose that $-<x(t),P^{\varepsilon}_{\rho}(t)x(t)> < d$ then by Lemma 8.2 for ε sufficiently small $-<x(t),P_{\rho}(t)x(t)> < d$. Now the proof follows in exactly the same way as in Theorem 8.3. $\square$

The dependency on ε of the region of exponential stability in the above theorem is not very satisfactory. However as a consequence of Lemma 8.2 we have the following corollary.

Corollary 8.5 Every bounded set in $\{x_0 \in \mathbb{C}^n \mid -<x_0,P_{\rho}(t_0)x_0> < d\}$ is a region of exponential stability at t_0. This follows from Lemma 8.3 since for all $x_0 \in \mathbb{C}^n$ and $\varepsilon \leq \varepsilon_0$

$$-<x_0,P^{\varepsilon}_{\rho}(t_0)x_0> + <x_0,P_{\rho}(t_0)x_0> \;\leq\; \kappa\,\varepsilon \parallel x_0 \parallel^2$$

The proof of Theorem 8.3 shows that $V(t,x) =< x, P^{\varepsilon}_{\rho}(t)x >$ is a joint Lyapunov function for all the systems (8.1) satisfying (8.9) with $\gamma <\parallel L^{\Sigma}_{t_0} \parallel^{-1}$. In the linear case one has

Proposition 8.6 Suppose (5.2), $\sum = (A,B,C)$ and $\rho <\parallel L^{\Sigma}_{t_0} \parallel^{-2}$ for some $t_0 \geq 0$. Then

$$V(t,x) \;=\; -<x,P_{\rho}(t)x>, \qquad t \geq t_0, x \in \mathbb{C}^n$$

is a common Lyapunov function guaranteeing the exponential stability of all the systems

$$\Sigma_D:\; \dot{x}(t) = [A(t)+B(t)D(t)C(t)]x(t), \qquad t \geq t_0,\; x(t_0) = x_0$$

with $D(\cdot) \in PC(t_0,\infty;\; \mathbb{C}^{m\times p})$ and $\parallel D(\cdot) \parallel^2_{L_\infty(t_0,\infty;\; \mathbb{C}^{m\times p})} < \rho$.

Proof: Suppose that $D \in PC(t_0,\infty;\; \mathbb{C}^{m\times p})$ and $\gamma :=\parallel D \parallel_{L_\infty} < \rho^{\frac{1}{2}}$. $V(t,x)$ is non-negative and its derivative $\dot{V}(t,x)$ along the trajectories of $\sum_D$ is non-positive by the previous proof. Hence V is a Lyapunov function for $\sum_D$. Moreover, applying (8.4) with $\varepsilon = 0$, $\delta = \rho - \gamma^2$ we obtain

$$\int_{t_0'}^{\infty} \parallel C(t)x(t) \parallel^2 dt \;\leq\; \delta^{-1}V(t_0',x_0)$$

for all $t_0' \leq t_0$ and trajectories $x(\cdot) = x_D(\cdot;t_0',x_0)$ of $\sum_D$. Thus, for any $x_0 \in \mathbb{C}^n$, we have for some $K > 0$

$$\sup_{t_0' \geq t_0} \parallel C(\cdot)x_D(\cdot;t_0',x_0) \parallel_{L_2(t_0',\,\mathbb{C}^p)} < K \parallel x_0 \parallel$$

and this implies the exponential stability of $\sum_D$ (see proof of Proposition 5.2 (iii)). $\square$

A Lyapunov function could be called of *maximal robustness* with respect to perturbations of the structure $\Delta(t) = B(t)D(t)C(t)$ if it guarantees the exponential stability of *all* the perturbed systems $\sum_D$ with $\parallel D \parallel_{L_\infty} < r_{\mathbb{C}}(A;B,C)$. In the time-invariant case a Lyapunov function of maximal robustness can in fact be constructed using the maximal solution of the $(ARE)_\rho$ with $\rho^{\frac{1}{2}} = r_{\mathbb{C}}(A;B,C)$, see *Hinrichsen and Pritchard* (1986b). The time-varying case is more complicated since $\parallel L^{\Sigma}_{t_0} \parallel^{-1}$ does not equal $r_{\mathbb{C}}(A;B,C)$ in general. In fact one can improve the result of Proposition 8.6 by using scalar Bohl transformations. If $\parallel D \parallel^2_{L_\infty} < \rho$ and $\rho^{\frac{1}{2}} < \sup_{\Theta} \ell(A - \frac{\dot{\Theta}}{\Theta} I_n;B,C)$, then the perturbed system $\sum_D$ will be exponentially stable. We will now construct a joint Lyapunov function for all these perturbed systems. Let $\sum_\theta = (A - \frac{\dot{\theta}}{\theta}I_n, B, C)$ and consider the differential Riccati equation

$$\begin{aligned}\dot{P}(t) + [A(t) \;&-\; \theta(t)^{-1}\dot{\theta}(t)I_n]^*P(t) + P(t)[A(t) - \theta(t)^{-1}\dot{\theta}(t)I_n] \\ &-\; \rho C^*(t)C(t) - P(t)B(t)B^*(t)P(t) = 0.\end{aligned}$$

Suppose $\rho^{\frac{1}{2}} < \ell(A - \frac{\dot{\theta}}{\theta} I_n; B, C)$. If $P^{\theta}(\cdot)$ is the maximal bounded Hermitian solution, then differentiation of

$$V^{\theta}(t,x) = - < x, P^{\theta}(t)x >$$

along any solution of $\sum_{\theta}$ yields

$$\begin{aligned} \dot{V}^{\theta}(t,x(t)) = & - \| B^{\star}(t)P^{\theta}(t)x(t) + D(t)C(t)x(t) \|^2 \\ & - (\rho - \| D(t) \|^2) \| C(t)x(t) \|^2 . \end{aligned}$$

Now in a similar manner to Proposition 8.6 it follows that $\sum_{\theta}$ is exponentially stable and since $\theta(\cdot)$ is a Bohl transformation $\sum_{D}$ is exponentially stable as well.

References

ANDERSON, B.D.O. (1972) *External and internal stability of linear system - a new connection;* IEEE Trans. Autom. Control 17, 107 - 111

ANDERSON, B.D.O. and **MOORE, J.B.** (1969) *New results in linear system stability;* SIAM J. Control Optimization 7, 398 - 414

BASILE, G. and **MARRO, G.** (1969) *Controlled and conditioned invariant subspaces in linear system theory;* J. Optimization Theory Appl. 3, 306-315

BOHL, P. (1913) *Über Differentialungleichungen;* J. Reine Angew. Math. 144, 284 - 313

BROCKETT, R.W. (1970) *Finite Dimensional Linear Systems;* John Wiley and Sons Inc., New York

BRUNOVSKÝ, P. (1970) *A classification of linear controllable system;* Kybernetika Čislo 3, 173 - 187

CODDINGTON, E.A. and **LEVINSON, N.** (1955) *Theory of Ordinary Differential Equations;* Mc Graw-Hill, New York

COHN, P.M. (1971) *Free Rings and their Relations;* Academic Press, London and New York

COPPEL, W.A. (1971) *Disconjugacy;* Lect. Notes Math.220, Springer-Verlag, Berlin et al.

COPPEL, W.A. (1978) *Dichotomies in Stability Theory;* Lect. Notes Math.629, Springer-Verlag, Berlin et al.

COZZENS, J. and **FAITH, C.** (1975) *Simple Noetherian Rings;* Cambridge University Press

CURTAIN, R. (1985) *Decoupling in infinite dimensions;* Syst. Control Lett. 5, 249-254

CURTAIN, R. (1986) *Invariance concepts in infinite dimensions;* SIAM J. Control Optimization 24, 1009 - 1030

DALECKII, Ju.L. and **KREIN, M.G.** (1974) *Stability of Solutions of Differential Equations in Banach Spaces;* AMS, Providence, Rhode Island

DOYLE, J.C. and **STEIN. G.** (1981) *Concepts for a classical/modern synthetics;* IEEE Trans. Autom. Control 26, 4 - 16

EKELAND,I. and **TEMAM, R.** (1976) *Convex Analysis and Variational Problems;* North-Holland, Amsterdam

FORNEY, G.D. Jr. (1975) *Minimal bases of rational vector spaces, with applications to multivariable linear systems;* SIAM J. Control Optimization 13, 493-520

FRANCIS, B: and **ZAMES, G.** (1983) *Feedback minimax sensitivity and optimal robustness;* IEEE Trans. Autom. Control 28, 585 - 601

FUHRMANN, P.A. (1976) *Linear algebra and finite dimensional linear systems;* Ben Gurion University of Negev, Math. Report No. 143

FUHRMANN, P.A. (1977) *On strict system equivalence and similarity;* Int. J. Control 25, 5-10

GANTMACHER, F.R. (1959) *The Theory of Matrices;* Vol. 1, Chelsea Publ., New York

GLÜSING-LÜERŒEN, H. (1987) *Invariants of discrete time-varying linear systems;* in preparation

GOHBERG, I., LANCASTER, P. and **RODMAN, L.** (1983) *Matrices and Indefinite Scalar Products;* Birkhäuser-Verlag, Basel et al.

HAHN, W. (1967) *Stability of Motion;* Die Grundlehren der math. Wissenschaften Bd. 138, Springer-Verlag, New York

HAUTUS, M.L.J. (1980) *(A,B)-invariant and stabilizability subspaces, a frequency domain description;* Automatica 16, 703-707

HEARN,A.C. (1985) *Reduce user's manual;* Version 3.2, Rand Publication CP 78, The Rand Corporation, Santa Monica, CA 90406

HINRICHSEN, D., ILCHMANN, A. and **PRITCHARD, A.J.** (1987) *Robustness of stability of time-varying systems;* Report No. 161, Institut für Dynamische Systeme, Bremen, submitted to the Journal Differ. Equations

HINRICHSEN, D. and **PRÄTZEL-WOLTERS, D.** (1980) *Solution modules and system equivalence;* Int. J. Control 32, 777-802

HINRICHSEN, D. and **PRITCHARD, A.J.** (1986a) *Stability radii of linear systems;* Syst. Control Lett. 7, 1-10

HINRICHSEN, D. and **PRITCHARD, A.J.** (1986b) *Stability radius for structural perturbations and the algebraic Riccati equation;* Syst. Control Lett. 8, 105-113

HIRSCHHORN, R.M. (1981) *(A,B)-invariant distributions and disturbance decoupling of nonlinear systems;* SIAM J. Control Optimization 19, 1-19

HOPPENSTAED, F.C. (1966) *Singular perturbations on the infinite interval,* Trans. Am. Math. Soc. 123, 521-535

IKEDA,M., MAEDA, H.and KODAMA, S. (1972) *Stabilization of linear systems;* SIAM J. Control Optimization 10, 716-729

IKEDA, M., MAEDA, H. and **KODAMA, S.** (1975)*Estimation and feedback in linear time-varying systems: a deterministic theory,* SIAM J. Control Optimization 13, 304 - 326

ILCHMANN, A. (1985a) *Time-varying linear systems and invariants of system equivalence;* Int. J. Control 42, 759-790

ILCHMANN, A. (1985b) *Disturbance decoupling for time-varying linear systems: a geometric approach;* Report No. 149, Forschungsschwerpunkt dynamische Systeme, Bremen

ILCHMANN, A. (1986) *Time-varying linear control systems: a geometric approach;* submitted for publication to IMA J. of Math. Control. and Inform.

ILCHMANN, A. (1987) *Correspondence, Comments on 'Canonical transformation for a class of time-varying multivaribale systems';* Int. J. Control 45, 365

ILCHMANN, A, and **KERN, G.** (1987) *Stabilizability of systems with exponential dichotomy;* Syst. Control Lett. 8, 211 - 220

ILCHMANN, A., NÜRNBERGER, I. and **SCHMALE, W.** (1984) *Time-varying polynomial matrix systems;* Int. J. Control 40, 329-362

ILCHMANN, A., OWENS, D.H. and **PRÄTZEL-WOLTERS, D.** (1987a) *High gain robust adaptive controllers for multivariable systems;* Syst. Control Lett. 8, 397-404

ILCHMANN, A., OWENS, D.H. and **PRÄTZEL-WOLTERS, D.** (1987b) *Sufficient conditions for stability of linear time-varying systems;* Syst. Control Lett. 9, 157-163

ISIDORI, A. (1985) *Nonlinear Control Systems: An Introduction;* Lect. Notes in Control Inf. Sci. 72, Springer-Verlag, Berlin, et al.

ISIDORI, A., KRENER,A.J., GORI-GIORI,G. and **MONACO, S.** (1981) *Nonlinear decoupling via feedback: a differential geometric approach;* IEEE Trans. Autom. Control 26, 331 - 345

JOHNSON, G.W. (1969) *A deterministic theory of estimation and control;* IEEE Trans. Autom. Control 14, 380 - 384

KAILATH, T. (1980) *Linear Systems;* Prentice-Hall, Englewood Cliffs N.J.

KALMAN, R.E. (1960) *Contributions to the theory of optimal control;* Bull. Soc. Math. Mexico 5, 102 - 119

KALMAN, R.E., HO, Y.C. and **NARENDA, K.S.** (1963) *Controllability of linear dynamical systems;* Contributions to Differ. Equations 1, 189 - 213

KAMEN, E.W. (1976) *Representation and realization of operational differential equations with time-varying coefficients;* J. Franklin Inst. 301, 559 - 571

KAMEN, E.W. (1979) *New results in realization theory for linear time-varying analytic systems;* IEEE Trans. Autom. Control 24, 866 - 877

KATO, T. (1976) *Perturbation Theory for Linear Operators;* Springer-Verlag, Berlin et al.

KNOBLOCH, H.W. and KAPPEL, F. (1974) *Gewöhnliche Differentialgleichungen;* Teubner-Verlag, Stuttgart

KNOBLOCH, H.W. and **KWAKERNAAK, H.** (1985) *Lineare Kontrolltheorie;* Springer-Verlag, Berlin et al.

KRAUSE. J.M. and **KUMAR, K.S.P.** (1986) *An alternative stability analysis framework for adaptive control;* Syst. Control Lett. 7, 19-24

KREISSELMEIER, G. (1985) *An approach to stable indirect adaptive control;* Automatica 21, 425 - 431

KWAKERNAAK, H. (1984) *Minimax frequency domain optimization of multivariable linear feedback systems;* Proc. 9th IFAC World Congress, Budapest (Pergamon Press, Oxford-New York)

KWAKERNAAK, H. and **SIVAN, R.** (1972) *Linear Optimal Control Systems;* Wiley, New York

LA SALLE, J. and **LEFSCHETZ, S.** (1961) *Stability by Lyapunov's Direct Methods;* Academic Press, New York

LYAPUNOV, A.M. (1893) *Problème général de la stabilité du mouvement;* Comm. Soc. Math. Kharkov (in Russian); translated in: Ann. Math. Stud. 17 (1949)

MAC DUFFEE, C.C. (1956) *The Theory of Matrices;* Chealsea Publ., New York

MÅRTENSON, B. (1986) *Adaptive Stabilization;* Doctoral Dissertation, Lund Institute of Technology, Lund

MEGAN, M. (1976a) *On the input-output stability of time varying linear control systems;* Semin. Ecuat. Funct., Univ. Timisoara 37

MEGAN, M. (1976b) *On the controllability, stability and Lyapunov differential equation for linear control systems;* Semin. Ecuat. Funct., Univ. Timisoara 39

MÜNZNER, H.F. and **PRÄTZEL-WOLTERS, D.** (1979) *Minimal bases of polynomial modules, structural indices and Brunovský-transformation;* Int. J. Control 30, 291 - 318

NARASIMHAN, R. (1985)*Complexe Analysis in One Variable;* Brikhäuser-Verlag, Boston

NEWMAN, M. (1972) *Integral Matrices;* Academic Press, New York

NGUYEN, C. (1986) *Canonical transformation for a class of time-varying multivaribale systems;* Int. J. Control 43, 1061 - 1074

NOBLE, B. and **DANIEL, J.W.** (1977) *Applied Linear Algebra;* Prentice-Hall, Englewood Cliffs, N.J.

ORE, O. (1933) *Theory of non-commutative polynomials;* Ann. Math. II. Ser. 34, 480 - 508

PEETRE, J. (1960) *Réctification a l'article "Une caractérisation abstraite des opérateurs différentiels";* Math. Scand. 8, 116 - 120

PERNEBO, L. (1977) *Notes on strict system equivalence;* Int. J. Control 25, 21 - 38

PERRON, O. (1930) *Die Stabilitätsfrage der Differentialgleichungen;* Math. Z. 32, 703 - 728

POPOV, V.M. (1972) *Invariant description of linear time-invariant controllable systems;* SIAM J. Control Optimization 10, 252 - 264

POSTLETHWAITE, I., EDMUNDS, J.M. and **MAC FARLANE, A.G.J.** (1981) *Principal gains and principal phases in the analysis of linear multivariable feedback systems;* IEEE Trans. Autom. Control 26, 32 - 46

PRÄTZEL-WOLTERS, D. (1981) *Brunovský equivalence of system matrices: the reachable case;* IEEE Trans. Autom. Control 26, 429 - 434

REGHIS, M. and **MEGAN, M.** (1977) *Riccati equation, exact controllability and stabilizability;* Semin. Ecuat. Funct., Univ. Timisoara 42

ROSENBROCK, H.H. (1963) *The stability of linear time-dependent control systems;* Int. J. Electr. Control 15, 73 - 80

ROSENBROCK, H.H. (1970) *State-Space and Multivariable Theory;* Nelson and Sons Ltd., London

ROSENBROCK, H.H. (1977a) *A comment on three papers;* Int. J. Control 25, 1 - 3

ROSENBROCK, H.H. (1977b) *The transformation of strict systems equivalence;* Int. J. Control 25, 11 - 19

SCHLESINGER, L. (1895) *Handbuch der Theorie der linearen Differentialgleichungen* Bd.1; Leipzig

SCHUMACHER, J.M. (1979) *(C, A)-invariant subspaces: some facts und uses;* Rapport 110, Wiskundig Seminarium, Vrije Universiteit, Amsterdam

SILVERMAN, L.M. (1971) *Realization of linear dynamical systems;* IEEE Trans. Autom. Control 16, 554 - 567

SILVERMAN, L.M. and **ANDERSON, B.D.O.** (1968) *Controllability, observability and stability of linear systems;* SIAM J. Control Optimization 6, 121 - 130

SILVERMAN, L.M. and **BUCY, R.S.** (1970) *Generalizations of a theorem of Doležal,* Math. Syst. Theory 4, 334 - 339

SILVERMAN, L.M. and **MEADOWS, H.E.** (1967) *Controllability and observability in time-variable linear systems;* SIAM J. Control Optimization 5, 64 - 73

WEDDERBURN, J.H.M. (1915) *On matrices whose coefficients are functions of a simple variable;* Trans. Am. Math. Soc. 16, 328 - 332

WILLEMS, J.C. (1971) *Least squares stationary optimal control and the algebraic Riccati equation;* IEEE Trans. Autom. Control 16, 621 - 634

WILLEMS, J.L. (1970) *Stability Theory of Dynamical Systems;* Nelson, London

WOLOVICH, W.A. (1974) *Linear Multivariable Systems;* Springer-Verlag, New York et al.

WONHAM, W.M. (1974) *Linear Multivariable Control: A Geometric Approach;* Lect. Notes Econ. Math. Syst. 101, Springer-Verlag, New York et al.

WONHAM. W.M. (1985) *Linear Multivariable Control: a Geometric Approach;* 3rd ed., Springer-Verlag, New York et al.

WONHAM, W.M. and **MORSE, A.S.** (1970) *Decoupling and pole assignment in linear multivariable systems: a geometric approach;* SIAM J. Control Optimization 8, 1 - 18

WU, M.Y. (1974) *A note on stability of linear time-varying systems;* IEEE Trans. Autom. Control 19, 162

YLINEN, R. (1975) *On the algebraic theory of linear differential and difference systems with time-varying or operator coefficients;* Helsinki University Techn., System theory laboratory report B 23

YLINEN, R. (1980) *An algebraic theory for analysis and synthesis of time-varying linear differential systems;* Acta Polytech. Scand., Math. Comput. Sci. Ser. 32

ZAMES, G. (1981) *Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms and approximate inverses;* IEEE Trans. Autom. Control 26, 301 - 320

Subject Index

Symbol Index

$\mathcal{V} = (\mathcal{V}(t))_{t\in\mathbf{R}}$	time-varying subspace	24
$\mathcal{V}^a(t), \mathcal{V}^d(t)$	time-varying subspace associated with the adjoint resp. dual system	17
$\mathcal{V}^\perp$	$(\mathcal{V}(t)^\perp)_{t\in\mathbf{R}}$	
$\overline{\mathcal{V}}$	equivalence class of time-varying subspaces	25
$\mathcal{V}_1(t) \overset{ae}{=} \mathcal{V}_2(t)$	equal almost everywhere	25
$\mathcal{V}_1(t) \overset{\subseteq}{\overset{ae}{=}} \mathcal{V}_2(t)$	included almost everywhere	25
$\overline{\mathcal{V}_1} < \overline{\mathcal{V}_2}$	inclusion of equivalence classes	83
$\overline{\mathcal{V}_*}(\mathcal{L})$	smallest (C, A)-invariant family which contains $\overline{\mathcal{L}}$	83
$\overline{\mathcal{V}^*}(\mathcal{L}')$	largest meromorphically (A, B)-invariant family contained in $\overline{\mathcal{L}'}$	84
$[V(D)]_\ell$	leading coefficient matrix	61
$\mathbf{W} = V(D) \cdot \mathcal{M}[D]^k$	$\mathcal{M}[D]$- right module generated by $V(D) \in \mathcal{M}[D]^{n\times k}$	61
$\mathbf{W}_d$	$\{v(D) \in \mathbf{W} \mid \deg v(D) \leq d\}$	61
$\mathbf{W}_n$	set of time-varying subspaces $\mathcal{V} = (\mathcal{V}(t))_{t\in\mathbf{R}}$ where $\mathcal{V}(t)$ is a subspace of $\mathbb{R}^n$ for every $t \in \mathbb{R}$	24
$\overline{\mathbf{W}_n} = \{\overline{\mathcal{V}} \mid \mathcal{V} \in \mathbf{W}_n\}$	equivalence class of $\mathbf{W}_n$	81
$W(t_0, t_1), W_i(t_0, t_1)$	controllability resp. induced controllability Gramian	9
$X(t), X^a(t), X^d(t)$	fundamental matrix of the nominal, adjoint resp. dual system	17
$Y_i(t_{-1}.t_0)$	induced reachability Gramian	19

$\mu[P, -Q]$	input module	64
$\Phi(t, t_0), \Phi^a(t, t_0), \Phi^d(t, t_0)$	transition matrix of the nominal, adjoint resp. dual system	5,17
$\Phi_F(t, t_0)$	transition matrix associated with the feedback system $\dot{x} = [A + BF]x$	89
$\Phi_\rho(t, t_0)$	transition matrix associated with the generator $A_\rho = A - BB^* P_\rho^+$	143