Geometric Control Theory for Linear Systems: a Tutorial

Giovanni Marro, Fabio Morbidi, Lorenzo Ntogramatzidis*, Domenico Prattichizzo

***Abstract*— This paper reviews in a condensed form the main tools and results of the geometric approach developed in the last forty years. Because of the vastness of the subject, this tutorial does not pretend to be exhaustive, and more emphasis will be given to selected topics and to the related computational tools. The authors hope their effort to provide a unified view of geometric control theory may be profitable to awake renewed interest in this research field.**

I. INTRODUCTION

Forty years after the introduction of the concepts of controlled and conditioned invariance by Basile and Marro [4], the area of systems and control theory known as *Geometric Approach* [8], [35], [33] still attracts a great deal of interest.

However, nowadays more than ever the research in this field seems to be mostly confined to extremely specialised niches. Very few courses in systems and control engineering around the world include geometric notions, even though these have proved to be extremely profitable in delivering a fresh perspective into several classical problems. Indeed, geometry has the potential of offering a very intuitive insight into the properties of systems, that can be hardly achieved with other approaches. In fact, geometric concepts can be easily understood taking advantage of graphical representations of state and output trajectories evolving on suitably defined subspaces; for example, the concepts of controlled and conditioned invariant subspaces, which are the pillars upon which the geometric control theory has been built over the past decades, can be grasped in depth by resorting to an intuitive interpretation in terms of the properties of the state trajectories.

The huge potential of the geometric approach to enhance intuition in solving very difficult analysis and synthesis problems has been greatly underestimated by the control community. In fact, the tools of geometric control theory have been mostly employed not to facilitate insight into the structure of systems and into the solution of fundamental problems, but to create a parallel set of definitions and notions that, while extremely elegant and abstract, cannot be captured without a long and time consuming preliminary study. In addition, in most cases far too much emphasis has been given to rigour and formalism – thus covering the inherent simplicity of geometric tools with unnecessarily heavy mathematics – whereas too little efforts have been devoted to make geometric ideas accessible to the wider community.

This tutorial paper has the ambition of filling this gap, by providing an extremely simplified and intuitive presentation of the basic notions on which geometric control theory relies and a short review of the use of the geometric approach tools to classify, analyse and solve a number of control problems. In fact, some problems already addressed and solved in the literature using a rigorous mathematical formalism can be profitably revisited within the geometric framework to provide an alternative and intuitive insight, more suited to establish connections between apparently different problems. In addition and no less importantly, the tools of the geometric approach benefit of a very efficient computational environment.

In this paper, rigour will not be sacrificed: however special care will be devoted to present the results in an accessible and didactic form. The algebra of matrices will be largely employed, as bases of subspaces, to convey the deep abstract significance of the geometric approach. This will also help keep the theoretical exposition of geometric concepts, notions and ideas significantly closer to their computational implementation.

In this respect, the GA toolbox [19] provides a suite of Matlab functions and routines that allow an immediate application of the developed theory.

To conclude, it is worth emphasizing that the goal of this tutorial paper is not to provide a historical perspective of the main achievements of geometric control in last four decades. However, where possible, a short historical account on the major breakthroughs will be presented.

The rest of the paper is organized as follows. In Section II some preliminary definitions are given. In Section III some geometrically-stated system properties are presented. Section IV reviews geometrically-solved regulation problems. Section V is devoted to H_2-optimal control and filtering. Finally, in Section VI some concluding remarks are given.

Notation: The symbol $\mathbb{R}^{n\times m}$ is used to denote the space of $n \times m$ real constant matrices. The image and the null-space of matrix $M \in \mathbb{R}^{n\times m}$ are respectively denoted by $\operatorname{im} M$ and $\ker M$. Denote by M^{T} and by $M^{\dagger}$ the transpose and the Moore-Penrose pseudo-inverse of M, respectively. The symbol I_d stands for the $d \times d$ identity matrix. The symbol $\mathbb{C}_g$ denotes either the open left-half complex plane

This work was partially supported by the Australian Research Council (Discovery Grant DP0986577).

* Corresponding author and organiser.

G. Marro is with the Dipartimento di Elettronica, Informatica e Sistemistica, Università di Bologna, viale Risorgimento 2, I-40136 Bologna, Italy. giovanni.marro@unibo.it

F. Morbidi and D. Prattichizzo are with the Dipartimento di Ingegneria dell' Informazione, Università di Siena, via Roma 56 - I-53100, Italy. {morbidi,prattichizzo}@ing.unisi.it

L. Ntogramatzidis is with the Department of Mathematics and Statistics, Curtin University of Technology, Perth, 6845 WA, Australia. L.Ntogramatzidis@curtin.edu.au

ISBN 978-963-311-370-7

$\mathbb{C}^-$ in the continuous time or the open unit disc $\mathbb{C}^\circ$ in the discrete time. The symbol $\mathbb{C}_0$ denotes the imaginary axis in the continuous time and the unit circle in the discrete time. The symbol $\sigma(A)$ denotes the spectrum of the square matrix A. If $\sigma(A)\subset\mathbb{C}_g$, matrix A is said to be *stable*. The symbol $\oplus$ denotes the *direct sum* of subspaces. Given a linear map $A:\mathcal{X}\longrightarrow\mathcal{Y}$ and $\mathcal{Z}\subseteq\mathcal{Y}$, the inverse map $A^{-1}\,\mathcal{Z}$ denotes the set of all the points in $\mathcal{X}$ whose image according to A belongs to $\mathcal{Z}$, i.e., $A^{-1}\,\mathcal{Z}=\{x\in\mathcal{X}\,|\,A\,x\in\mathcal{Z}\}$.

II. Geometric approach: foundations

The geometric approach deals with subspaces. Properties of systems are expressed in terms of linear spaces, and the most important analysis and synthesis algorithms are based on operations on them, including sum, intersection, direct and inverse linear transformation, orthogonal complementation, etc. Consider a linear time-invariant (LTI) system Σ, with input $u\in\mathbb{R}^p$, state $x\in\mathbb{R}^n$ and output $y\in\mathbb{R}^q$. For the sake of generality and to achieve a better insight into the meaning of some geometric features of the state trajectories, we will consider both continuous-time systems

$$\begin{aligned}\dot{x}(t) &= A\,x(t)+B\,u(t),\\ y(t) &= C\,x(t),\end{aligned}\qquad(1)$$

and discrete-time systems

$$\begin{aligned}x(k{+}1) &= A\,x(k)+B\,u(k),\\ y(k) &= C\,x(k).\end{aligned}\qquad(2)$$

These systems are often referred to as the *triples* (A,B,C). Matrices B and C are assumed to be of maximum rank. Systems (1) and (2) are said to be *internally stable* if $\sigma(A)\subset\mathbb{C}_g$. The overall theory presented here is grounded on three very intuitive notions: *controlled invariance*, *conditioned invariance* and *invariant zeros*. *Duality* plays an important role in unifying these concepts and it will be frequently used throughout the paper. Before introducing these concepts, we first need to recall some properties of the so-called *invariant subspaces*.

A. Invariant subspaces

Let $\mathcal{X}$ be a vector space and $A:\mathcal{X}\to\mathcal{X}$ a linear map. A subspace $\mathcal{J}\subseteq\mathcal{X}$ is said to be A-*invariant*, if

$$A\,\mathcal{J}\subseteq\mathcal{J}.$$

Consider the change of basis defined by the transformation $T=[T_1\;\;T_2]$ with $\operatorname{im}T_1=\mathcal{J}$ and T_2 such that T is nonsingular. Matrix $A'=T^{-1}A\,T$ can be written as

$$A'=\begin{bmatrix}A'_{11} & A'_{12}\\ 0 & A'_{22}\end{bmatrix}.\qquad(3)$$

The A-invariant subspace $\mathcal{J}$ is said to be *internally stable* if $\sigma(A'_{11})\subset\mathbb{C}_g$ and *externally stable* if $\sigma(A'_{22})\subset\mathbb{C}_g$.

Subspace $\mathcal{J}$ is said to be *complementable* if another A-invariant $\mathcal{J}_c$ exists such that $\mathcal{J}\oplus\mathcal{J}_c=\mathcal{X}$. Let us consider again the change of basis defined by T. Subspace $\mathcal{J}$ is complementable if and only if the *Sylvester equation*

$$A'_{11}\,X-X\,A'_{22}=-A'_{12}$$

admits a solution X. If so, a basis matrix of $\mathcal{J}_c$ is given by $J_c=J\,X+T_2$. The new transformation $T=[T_1\;J_c]$ causes $A'_{12}=0$ in (3).

It is easily proved that the sum and the intersection of A-invariant subspaces is A-invariant; hence any subset of the set of all the A-invariant subspaces contained in a given subspace $\mathcal{C}$ is closed under sum and the set of all the A-invariant subspaces containing a given subspace $\mathcal{B}$ is closed under intersection. Hence, the supremum of the former set is the sum of all the A-invariant subspaces contained in $\mathcal{C}$, and is denoted by $\mathcal{J}^*_\uparrow$; similarly, the infimum of the latter set is the intersection of all the A-invariant subspaces containing $\mathcal{B}$, and is denoted by $\mathcal{J}^*_\downarrow$.

If $A:\mathbb{F}^n\to\mathbb{F}^n$ (where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$) the subspace $\mathcal{J}^*_\uparrow$ can be computed with the sequence

$$\begin{aligned}\mathcal{J}_1 &= \mathcal{C},\\ \mathcal{J}_i &= \mathcal{C}\cap A^{-1}\,\mathcal{J}_{i-1},\qquad(i=2,3,\ldots),\end{aligned}\qquad(4)$$

and $\mathcal{J}^*_\downarrow$ with the sequence

$$\begin{aligned}\mathcal{J}_1 &= \mathcal{B},\\ \mathcal{J}_i &= \mathcal{B}+A\;\mathcal{J}_{i-1},\qquad(i=2,3,\ldots),\end{aligned}\qquad(5)$$

Both the above sequences converge in a finite number of steps (at most n).

B. Controlled invariant subspaces

Given a linear map $A:\mathcal{X}\to\mathcal{X}$ and a subspace $\mathcal{B}\subseteq\mathcal{X}$, a subspace $\mathcal{V}\subseteq\mathcal{X}$ is $(A,\mathcal{B})$-*controlled invariant* if the following inclusion holds

$$A\,\mathcal{V}\subseteq\mathcal{V}+\mathcal{B}.$$

Controlled invariant subspaces have the following properties:

1) The sum of any two $(A,\mathcal{B})$-controlled invariant subspaces is $(A,\mathcal{B})$-controlled invariant.
2) Let $\mathcal{B}=\operatorname{im}B$ and $\mathcal{C}=\ker C$. A state trajectory $x(\cdot)$ of the continuous or discrete-time LTI system Σ can be made invisible at the output by means of a suitable control action if and only if the initial state $x(0)$ belongs to an $(A,\mathcal{B})$-controlled invariant subspace contained in $\mathcal{C}$.
3) For any $(A,\mathcal{B})$-controlled invariant subspace $\mathcal{V}$ there exists a matrix F (called a *friend* of $\mathcal{V}$) such that $(A+BF)\,\mathcal{V}\subseteq\mathcal{V}$. If, furthermore, a matrix F exists such that $\mathcal{V}$ is an internally stable and/or an externally stable $(A+BF)$-invariant, $\mathcal{V}$ is said to be *internally stabilizable* and/or *externally stabilizable*, respectively.

Property 1 implies that the set of all the controlled invariant subspaces contained in a given subspace $\mathcal{C}\subseteq\mathcal{X}$ admits a supremum, that coincides with their sum. It can be computed with the sequence

$$\begin{aligned}\mathcal{V}_1 &= \mathcal{C},\\ \mathcal{V}_i &= \mathcal{C}\cap A^{-1}\,(\mathcal{V}_{i-1}+\mathcal{B}),\qquad(i=2,3,\ldots),\end{aligned}\qquad(6)$$

that converges in a finite number of steps. Referring to the system Σ (continuous or discrete-time), the sequence (6) converges to $\mathcal{V}^*$, the maximal $(A,\mathcal{B})$ controlled invariant subspace contained in $\mathcal{C}$, that is the locus of all the possible state trajectories of Σ invisible at the output. For this reason, $\mathcal{V}^*$ is called *the maximal output-nulling controlled invariant subspace of* Σ, [1].

A state trajectory that crosses a controlled invariant subspace $\mathcal{V}$ can be maintained on it using a suitable control action that can always be expressed as a static state-to-input feedback. Since $\mathcal{V} \subseteq \mathcal{V}^* \subseteq \mathcal{C}$, the segment of trajectory on $\mathcal{V}$ is invisible at the output.

C. Conditioned invariant subspaces

Conditioned invariant subspaces are dual to controlled invariant subspaces. Given a linear map $A : \mathcal{X} \to \mathcal{X}$ and a subspace $\mathcal{C} \subseteq \mathcal{X}$, a subspace $\mathcal{S} \subseteq \mathcal{X}$ is $(A, \mathcal{C})$-*conditioned invariant* if

$$A\,(\mathcal{S} \cap \mathcal{C}) \subseteq \mathcal{S}.$$

Conditioned invariant subspaces enjoy the following properties:

1) The intersection of any two $(A, \mathcal{C})$-conditioned invariant subspaces is $(A, \mathcal{C})$-conditioned invariant.
2) The state trajectories of the discrete-time system Σ that originate at $x(0) = 0$ can be made invisible at the output for a certain number of steps ρ while lying on a given subspace $\mathcal{S}$ if and only if $\mathcal{S}$ is an $(A, \mathcal{C})$-conditioned invariant subspace containing $\mathcal{B}$.
3) For any $(A, \mathcal{C})$-conditioned invariant subspace $\mathcal{S}$ there is a matrix G (called a *friend* of $\mathcal{S}$) such that $(A + GC)\,\mathcal{S} \subseteq \mathcal{S}$. If, furthermore, a matrix G exists such that $\mathcal{S}$ is an externally stable and/or an internally stable $(A + GC)$-invariant subspace, $\mathcal{S}$ is said to be *externally stabilizable* and/or *internally stabilizable*, respectively.

Property 1 implies that the set of all the conditioned invariant subspaces containing a given subspace $\mathcal{B} \subseteq \mathcal{X}$ admits an infimum, that coincides with their intersection. It can be computed with the sequence

$$\begin{aligned} \mathcal{S}_1 &= \mathcal{B}, \\ \mathcal{S}_i &= \mathcal{B} + A\,(\mathcal{S}_{i-1} \cap \mathcal{C}), \qquad (i = 2, 3, \ldots), \end{aligned} \tag{7}$$

that converges to it in a finite number of steps ρ. Referring to the system Σ, the sequence (7) yields $\mathcal{S}^*$, which is the minimal $(A, \mathcal{C})$ conditioned invariant subspace containing $\mathcal{B}$, that for discrete-time systems is the maximal subspace of the state space reachable from the origin in at most ρ steps with trajectories that have all the states except the last one invisible at the output. Subspace $\mathcal{S}^*$ is also called *the minimal input-containing conditioned invariant subspace of* Σ.

Property 2 has no counterpart in the continuous-time case unless control functions, that are usually considered to be piecewise continuous, are extended to include distributions. This aspect is beyond the scope of this work and is therefore omitted. A state trajectory of a discrete-time system starting from the origin can be maintained on a subspace $\mathcal{S}$ for a certain number of steps with a suitable control action if and only if $\mathcal{S}$ is a conditioned invariant subspace containing $\mathcal{B}$.

D. Reachable subspace on $\mathcal{V}^*$ and invariant zeros of Σ

Still referring to Σ, consider the previously defined maximal output-nulling subspace $\mathcal{V}^*$ and define the subspace $\mathcal{R}^* \subseteq \mathcal{V}^*$ as the maximum subspace reachable from the origin with state trajectories completely belonging to $\mathcal{V}^*$. It has been shown in [25] that

$$\mathcal{R}^* = \mathcal{V}^* \cap \mathcal{S}^*.$$

Hence $\mathcal{R}^*$ can be computed by using again the sequences (6) and (7). Note that $\mathcal{R}^*$ is $(A, \mathcal{B})$-controlled invariant. In fact, it is easily proved that the intersection of $(A, \mathcal{B})$-controlled invariant subspaces contained in $\mathcal{C}$ and $(A, \mathcal{C})$-conditioned invariant subspaces containing $\mathcal{B}$ is $(A, \mathcal{B})$-controlled invariant. While any state in $\mathcal{R}^*$ is completely reachable from any other state in $\mathcal{R}^*$ with state trajectories

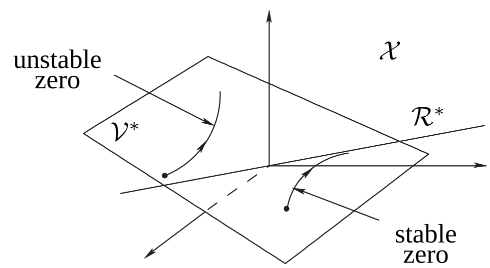

Fig. 1. $\mathcal{R}^*$ and some trajectories corresponding to invariant zeros.

invisible at the output, the trajectories on $\mathcal{V}^*$ corresponding to initial states not belonging to $\mathcal{R}^*$ are strictly constrained to be linear combinations of modes of the type $t^{m_k-1}\,e^{\zeta_i t}$, called *fixed modes*. The values of ζ_i, in general complex, are called the *invariant zeros* of Σ and the maximum value of the integer m_k is said the *multiplicity of* ζ_i, see Fig. 1.

Let F be a friend of $\mathcal{V}^*$. Consider the change of basis defined by the transformation $T = [T_1\; T_2\; T_3]$ with $\operatorname{im} T_1 = \mathcal{R}^*$, $\operatorname{im}[T_1\; T_2] = \mathcal{V}^*$ and T_3 such that T is nonsingular. The new matrices $A' = T^{-1}(A + BF)T$, $B' = T^{-1}B$, $C' = CT$ can be written as

$$\begin{aligned} A' &= \begin{bmatrix} A'_{11} & A'_{12} & A'_{13} \\ 0 & A'_{22} & A'_{23} \\ 0 & 0 & A'_{33} \end{bmatrix}, \quad B' = \begin{bmatrix} B'_1 \\ 0 \\ B'_3 \end{bmatrix}, \\ C' &= \begin{bmatrix} 0 & 0 & C'_3 \end{bmatrix}. \end{aligned} \tag{8}$$

The invariant zeros of Σ correspond to the eigenvalues of matrix A'_{22}. Both the pairs (A'_{11}, B'_1) and (A'_{33}, B'_3) are stabilizable if such is Σ. The *self-bounded controlled invariant subspaces* of Σ are defined, in the new basis, as the sum of $\mathcal{R}^*$ with the invariant subspaces of A'_{22}, [7], [31]. The system is said to be of *minimum phase* if all its invariant zeros are in $\mathbb{C}_g$. Hence, the invariant zeros of Σ are the internal unassignable eigenvalues of $\mathcal{V}^*$ and will be referred to with the symbol $\mathcal{Z}(\mathcal{V}^*)$ or $\mathcal{Z}(\Sigma)$. By extension, we will denote with $\mathcal{Z}(\mathcal{V})$ the internal unassignable eigenvalues of any controlled invariant subspace $\mathcal{V}$. If $\mathcal{V}$ is self-bounded, i.e., satisfies $\mathcal{R}^* \subseteq \mathcal{V} \subseteq \mathcal{V}^*$, $\mathcal{Z}(\mathcal{V})$ is a subset of the invariant zeros of Σ.

E. Extension to quadruples

For the sake of simplicity, the geometric concepts introduced in the previous sections have been defined for LTI systems represented by a triple (A, B, C), without any *feedthrough matrix* D. However, these ideas can easily be extended to systems represented by a quadruple (A, B, C, D), i.e., ruled by

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,u(t), \\ y(t) &= C\,x(t) + D\,u(t), \end{aligned} \tag{9}$$

and

$$\begin{aligned} x(k+1) &= A\,x(k) + B\,u(k), \\ y(k) &= C\,x(k) + D\,u(k), \end{aligned} \tag{10}$$

in the continuous and discrete-time case, respectively.

In these cases, output-nulling subspaces and their friends are defined as pairs ($\mathcal{V}\ F$) such that $(A+BF)\mathcal{V} \subseteq \mathcal{V}$ and $\mathcal{V} \subseteq \ker(C+DF)$. Likewise, input-containing subspaces and their friends are defined as pairs ($\mathcal{S}\ G$) such that $(A+GC)\mathcal{S} \subseteq \mathcal{S}$ and $\mathcal{S} \supseteq \operatorname{im}(B+GD)$.

The computation of the maximum output-nulling subspace $\mathcal{V}^*$ and the minimum input-containing subspace $\mathcal{S}^*$ is still possible with algorithms (6) and (7) applied to a suitably extended system. Refer to the overall system $\widehat{\Sigma}$ shown in

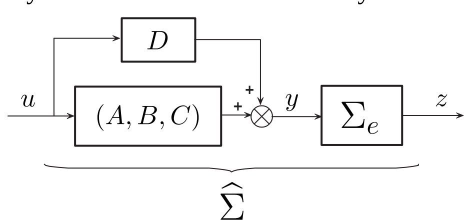

Fig. 2. An artifice to reduce a quadruple to a triple.

Fig. 2, where Σ_e is a set of integrators in the continuous-time case or a set of unit delays in the discrete-time case. It can be described by the extended state $\hat{x}$ and extended triple $(\widehat{A}, \widehat{B}, \widehat{C})$ defined as

$$\hat{x} = \begin{bmatrix} x \\ u \end{bmatrix}, \quad \widehat{A} = \begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix}, \quad \widehat{B} = \begin{bmatrix} B \\ D \end{bmatrix}, \quad \widehat{C} = \begin{bmatrix} 0 & I_q \end{bmatrix}. \tag{11}$$

Let us compute the output nulling $\widehat{\mathcal{V}}^*$ with basis matrix $\widehat{V}$ for system $\widehat{\Sigma}$ and the corresponding friend $\widehat{F}$ with (6), and denote with

$$\widehat{V} = \left[\begin{array}{c} V_1 \\ V_2 \end{array} \right], \quad \widehat{F} = \left[\begin{array}{cc} F_1 & F_2 \end{array} \right],$$

their partitions according to (11). Owing to the structure of $\widehat{C}$, it turns out that $V_2 = 0$ and $F_2 = 0$ and the maximum output nulling subspace of the quadruple (A, B, C, D) is $\mathcal{V}^* = \operatorname{im} V_1$ and that F_1 is a corresponding friend adapted to its basis matrix V_1.

A similar procedure applies to $\mathcal{S}^*$. The definitions of $\mathcal{R}^*$ and invariant zeros in Section II-D are still valid, provided that they are referred to the subspaces $\mathcal{V}^*$ and $\mathcal{S}^*$ defined in this section.

III. Geometrically-stated system properties

In the next two sections, the properties of controllability, observability, left and right invertibility and relative degree, will be expressed in geometric terms.

A. Controllability and observability

The concepts of *controllability* and *observability* were introduced in the early 60s to state mathematical conditions connected with the existence and uniqueness of Kalman regulators and filters. Kalman himself pointed out that the reachable subspace $\mathcal{R}$ of an LTI system Σ is the minimum A-invariant subspace containing $\mathcal{B} = \operatorname{im} B$ and the unobservable subspace $\mathcal{U}$ the maximum A-invariant subspace contained in $\mathcal{C} = \ker C$, [14]. These were the first system properties stated in geometric terms, and the related controllability and observability subspaces can be computed with the recursive algorithms (5) and (4), respectively.

Recall that if Σ is completely reachable (i.e., $\mathcal{R} = \mathcal{X}$), the eigenvalues of $A + BF$ are completely assignable by a suitable choice of F. If Σ is completely observable (i.e., $\mathcal{U} = \{0\}$), the eigenvalues of $A + GC$ are completely assignable by a suitable choice of G. This is the *pole assignability* property of static state feedback and static output injection.

B. Left invertibility, right invertibility, and relative degree

Consider the continuous-time LTI system (1) or (9) with $x(0) = 0$ and assume that its input function $u(t)$ is bounded and piecewise continuous. We introduce the following definitions.

1) The system is said to be *left invertible* if, given any admissible output function $y(t)$, $t \in [0, t_1]$, $t_1 > 0$, there is a unique corresponding input function $u(t)$, $t \in [0, t_1)$ producing that output function $y(t)$.
2) The system is said to be *right invertible* if there is an integer $\rho \geq 1$ such that, given any output function $y(t)$, $t \in [0, t_1]$, $t_1 > 0$ with piecewise continuous ρ-th derivative and such that $y(0) = 0, \ldots, y^{(\rho-1)}(0) = 0$, there is at least one input function $u(t)$, $t \in [0, t_1)$ producing that output function $y(t)$. The minimum value of ρ satisfying the above statement is called the *relative degree* of the system.

Consider the discrete-time LTI system (2) or (10) with $x(0) = 0$ and assume that its input function $u(t)$ is bounded.

1) The system is said to be *left invertible* if, given any admissible output function $y(k)$, $k \in [0, k_1]$, $k_1 \geq n$, there is a unique corresponding input function $u(k)$, $k \in [0, k_1)$ producing that output function $y(k)$.
2) The system is said to be *right invertible* if there is an integer $\rho \geq 1$ such that, given an output function $y(k)$, $k \in [0, k_1]$, $k_1 \geq \rho$ such that $y(k) = 0$, $k \in [0, \rho - 1]$, there is at least one input function $u(k)$, $k \in [0, k_1 - 1]$ producing that output function $y(k)$. The minimum value of ρ satisfying the above statement is called the *relative degree* of the system.

The geometric necessary and sufficient conditions for left and right invertibility of both continuous and discrete-time systems are simple when expressed in terms of $\mathcal{V}^*$ and $\mathcal{S}^*$:

- Left invertibility:

$$\mathcal{V}^* \cap \mathcal{S}^* = \{0\}.$$

- Right invertibility:

$$\mathcal{V}^* + \mathcal{S}^* = \mathcal{X}.$$

- Relative degree:
  For a right invertible system without feedthrough the relative degree is the minimal value of ρ such that $\mathcal{V}^* + \mathcal{S}_\rho = \mathcal{X}$ where $\mathcal{S}_i$ $(i = 1, 2, \ldots)$ is provided by sequence (7). For systems with feedthrough this still holds but referred to a suitably extended system, like the one shown in Fig. 2 (note that in this case the computed relative degree has to be reduced by one).

IV. Geometrically-solved regulation problems

In this section we introduce some basic control problems and present necessary and sufficient constructive conditions for their solvability. For the sake of brevity the proofs are omitted: the interested reader can find them in specialized books [8], [35].

A. Disturbance decoupling

When the output of a system must be decoupled from an input signal we must distinguish three cases

1) (inaccessible) disturbance decoupling
2) measurable signal decoupling
3) previewed signal decoupling

This distinction is basic for a correct statement and solution of the corresponding control problem. Feedback is strictly required only in the first case, since the second and third case are solvable with feedforward, possibly applied to a system which has been pre-strengthened with feedback.

1) Inaccessible disturbance decoupling by state feedback: The disturbance decoupling problem by state feedback is the basic problem of the geometric approach. It was studied in [9] and [36] as one of the earliest applications of these new concepts.

Consider the continuous-time LTI system Σ with two inputs shown in Fig. 3, described by

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,u(t) + H\,h(t), \\ y(t) &= C\,x(t), \end{aligned} \qquad (12)$$

where u denotes the manipulable input and h the disturbance input. Let $\mathcal{B} = \mathrm{im}B$, $\mathcal{H} = \mathrm{im}H$ and $\mathcal{C} = \ker C$. The (inaccessible) disturbance decoupling problem is stated as follows: *Determine, if possible, a state feedback matrix F such that the disturbance h has no influence on the output y.* In

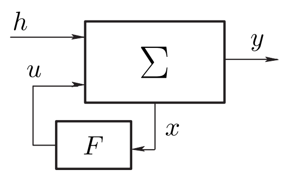

Fig. 3. The inaccessible disturbance decoupling problem.

spite of its apparent simplicity, the disturbance decoupling problem was not completely solved at the outset. The system with state feedback is described by

$$\begin{aligned} \dot{x}(t) &= (A + BF)\,x(t) + H\,h(t), \\ y(t) &= C\,x(t). \end{aligned}$$

It behaves as requested if and only if its reachable set by h, i.e., the minimum $(A + BF)$-invariant subspace containing $\mathcal{H}$, is contained in $\mathcal{C}$. Denote by $\mathcal{V}^*_{(\mathcal{B},\mathcal{C})}$ the maximum output-nulling $(A, \mathcal{B})$-controlled invariant subspace contained in $\mathcal{C}$. Since any $(A + BF)$-invariant subspace is $(A, \mathcal{B})$-controlled invariant, the inaccessible disturbance decoupling problem has a solution if and only if

$$\mathcal{H} \subseteq \mathcal{V}^*_{(\mathcal{B},\mathcal{C})}. \qquad (13)$$

This is a necessary and sufficient *structural condition* and does not ensure internal stability. If stability is required, we have the inaccessible disturbance decoupling problem *with stability*. Stability is conveniently handled by using the *self-bounded controlled invariant subspaces*, introduced in Section II-D. Recall that any $(A, \mathcal{B})$-controlled invariant subspace $\mathcal{V}$ which is self-bounded with respect to $\mathcal{V}^*_{(\mathcal{B},\mathcal{C})}$ satisfies the relation

$$\mathcal{R}^*_{(\mathcal{B},\mathcal{C})} \subseteq \mathcal{V} \subseteq \mathcal{V}^*_{(\mathcal{B},\mathcal{C})},$$

and the set of all the self-bounded controlled invariant subspaces is closed with respect to both sum and intersection: it is therefore possible to define the minimum self-bounded controlled invariant subspaces containing $\mathcal{H}$, provided that (13) holds. This is the reachable subspace with both inputs u and h, that clearly contains $\mathcal{H}$. It is defined as

$$\mathcal{V}_m = \mathcal{V}^*_{(\mathcal{B},\mathcal{C})} \cap \mathcal{S}^*_{(\mathcal{C},\mathcal{B}+\mathcal{H})}. \qquad (14)$$

The change of basis shown in (8) also holds for $\mathcal{V}_m$ instead of $\mathcal{V}^*$ with the stabilizability property of (A'_{33}, B'_3). Hence there is a state feedback F_m achieving disturbance decoupling and making the overall system stable if and only if $\mathcal{V}_m$ is internally stabilizable. Thus, the necessary and sufficient conditions for the solvability of the disturbance decoupling problem with stability are both the structural condition (13) and the following *stabilizability condition*:

$$\mathcal{V}_m \;\; \text{internally stabilizable}, \qquad (15)$$

that is equivalent to

$$\mathcal{Z}(\mathcal{V}_m) \subseteq \mathbb{C}_g. \qquad (16)$$

Since $\mathcal{Z}(\mathcal{V}_m)$ is a part of $\mathcal{Z}(\mathcal{V}^*_{(\mathcal{B},\mathcal{C})})$, phase minimality ensures stabilizability.

2) Measurable signal decoupling: This problem can be stated as follows: *Determine, if possible, a static or dynamic compensator such that the measurable signal h has no influence on output y.* The system is still ruled by (12). This problem appears as a slight extension of the previous inaccessible disturbance decoupling problem, but indeed it is very different and opens out to many types of solution. The first solution considered in the literature, based on state feedback and static feedforward, is illustrated in Fig. 4. The

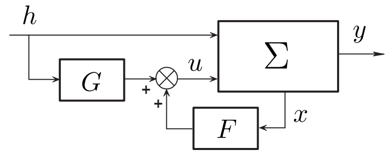

Fig. 4. Measurable signal decoupling with state feedback.

structural necessary and sufficient condition for its solvability [11], is

$$\mathcal{H} \subseteq \mathcal{V}^*_{(\mathcal{B},\mathcal{C})} + \mathcal{B}. \tag{17}$$

In fact if (17) holds, by assuming as matrix G in Fig. 4 a projection of Hh on $\mathcal{V}^*_{(\mathcal{B},\mathcal{C})}$ along $\mathcal{B}$, the action of input h is driven on $\mathcal{V}^*_{(\mathcal{B},\mathcal{C})}$, hence it is invisible at the output. It can be proved that the stabilizability condition for this problem is again given by (15) or (16), provided that (17) is satisfied [8]. However, the block scheme shown in Fig. 4 is not the most convenient for achieving decoupling of a measurable signal since it requires full access to the state. Let us consider instead the layout shown in Fig. 5 where Σ_c is a replica of the system with feedback shown in Fig. 4. It is worth noting that the dynamics of Σ_c can be restricted to $\mathcal{V}_m$, that is an internally stable $(A+BF_m)$-invariant subspace containing the projection of Hh on $\mathcal{V}_m$ along $\mathcal{B}$. However

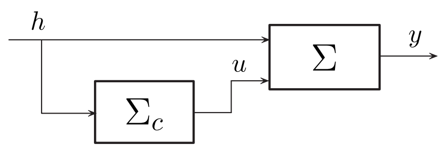

Fig. 5. Measurable signal decoupling with a feedforward compensator.

in this case Σ must be stable. But, if it is unstable, feedback can be avoided as shown in Fig. 6. If Σ is stabilizable

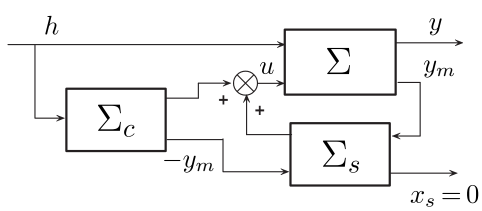

Fig. 6. Using both a compensator and a stabilizer.

from u and detectable from a suitable measurable output y_m (possibly coinciding with y, but, in general, provided by a second output matrix C_1), Σ can be stabilized with a feedback unit Σ_s which can be maintained at zero by the pre-compensator since this reproduces the state evolution (hence the output) of Σ, restricted to $\mathcal{V}_m$. If $y_m = y$ this connection is not necessary. The stabilizer is based on the state feedback matrix F_s such that $A+BF_s$ is strictly stable and the output injection G_s such that $A+G_sC_1$ is strictly stable, i.e., $\Sigma_s\,:\,(A+BF_s+G_sC_1\,,\,-G_s\,,\,F_s)$. The input matrix $-G_s$ is referred to both inputs. Note that the output of the stabilizer is identically zero, since its inputs due to the action of h on Σ and Σ_c cancel each other.

Now let us consider the dual problem of the measurable signal decoupling with feedforward control architecture: the so-called unknown-input observation. Consider the continuous-time LTI system Σ with two outputs shown in Fig. 7 and described by

$$\begin{aligned}\dot{x}(t) &= A\,x(t) + B\,u(t),\\ e(t) &= E\,x(t),\\ y(t) &= C\,x(t).\end{aligned} \tag{18}$$

The unknown-input observation problem of a linear function of the state (possibly the whole state) is stated as follows: *Design a stable feedforward unit that, connected to the output y, provides an exact estimation of the output e.* This problem was the object of very early investigation

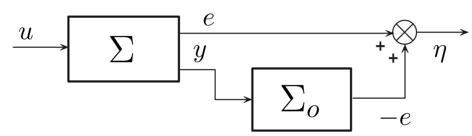

Fig. 7. Unknown-input observer of a linear function of the state.

in [3], [5]. More recently, owing to its connection with the fault detection problem, it has been the subject of hundreds of papers relying on convoluted matrix manipulations, but duality with the measurable signal decoupling problem has never been adequately recognized.

The overall system in Fig. 7 is clearly the dual of measurable signal decoupling considered in Fig. 5 (the summing junction on the right is only explanatory), so the design of an unknown-input observer can be considered a very standard problem in the geometric approach context. Obviously, the necessary and sufficient conditions to build a stable unknown-input observer Σ_o are still (13), (15) for the dual of system (18), obtained with the substitutions $A^{\mathrm{T}} \to A$, $B^{\mathrm{T}} \to C$, $C^{\mathrm{T}} \to B$, $E^{\mathrm{T}} \to H$. However, if an equivalent set of geometric conditions directly referred to the system matrices in (18) is sought after, we have that the structural condition is

$$\mathcal{S}^*_{(\mathcal{C},\mathcal{B})} \cap \mathcal{C} \subseteq \mathcal{E},$$

with $\mathcal{C} = \ker C$, $\mathcal{E} = \ker E$, and the stability condition is

$$\mathcal{S}_M \text{ externally stabilizable,}$$

where $\mathcal{S}_M$ is defined as

$$\mathcal{S}_M = \mathcal{S}^*_{(\mathcal{C},\mathcal{B})} + \mathcal{V}^*_{(\mathcal{B},\mathcal{C}\,\cap\,\mathcal{E})}. \tag{19}$$

Also in this case, a stabilizer can be used if the system

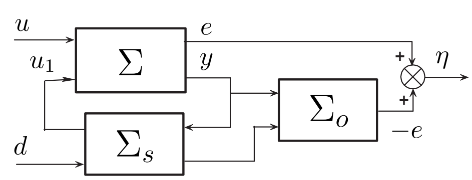

Fig. 8. Unknown-input observer with a stabilizer.

is unstable, as shown in Fig. 8. By duality, the output η is independent of both u and d (any disturbance acting on the stabilizer).

3) Previewed signal decoupling and delayed state estimation: It has been previously pointed out that a sufficient condition for stability in the above disturbance and measurable signal decoupling problems is phase minimality of Σ. If Σ is not minimum-phase, however, it is still possible to obtain decoupling if h is known in advance by a certain amount of time (several times the maximum time constant of the unstable zeros). In this case the necessary structural condition is still (17), while the stabilizability condition is

$$\mathcal{Z}(\mathcal{V}_m) \cap \mathbb{C}_0 = \emptyset.$$

In Fig. 9(a), h denotes the previewed signal and h_p its

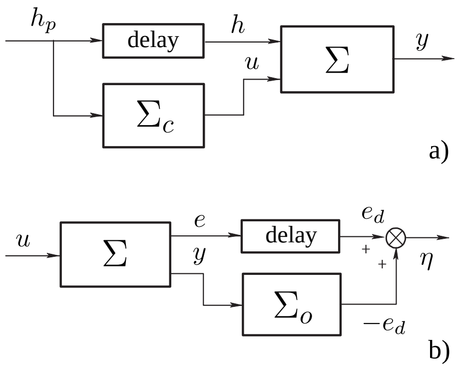

Fig. 9. (a) Previewed decoupling; (b) Delayed unknown-input observation.

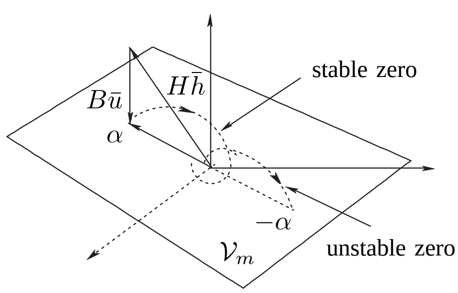

Fig. 10. Preaction along the unstable zeros.

value t_0 seconds in advance, so that $h_p(t) = h(t+t_0)$. The feedforward unit Σ_c includes a convolutor, also called a *finite impulse response* (FIR) system. Refer to Fig. 10 and suppose that a single impulse is applied at input h, i.e., assume $h(t) = \bar{h}\,\delta(t)$, causing an initial state $H\bar{h}$ at time zero. Since $\mathcal{H} \subseteq \mathcal{V}_m + \mathcal{B}$ owing to (17), this initial state can be projected on $\mathcal{V}_m$ along $\mathcal{B}$, and decomposed into three components: a component on $\mathcal{R}^*_{(\mathcal{B},\mathcal{C})}$, a component on the subspace of $\mathcal{V}_m$ corresponding to strictly stable zeros, and a component on the subspace of $\mathcal{V}_m$ corresponding to strictly unstable zeros. While the former two (state α) can be driven to the origin along stable trajectories on $\mathcal{V}_m$, the latter, that corresponds to unstable motions on $\mathcal{V}_m$, can be nulled by a preaction on u prior to its occurrence (in the time interval $[-t_0, 0]$), thus cancelling it at $t = 0$ (state $-\alpha$). This is obtained by means of the aforementioned FIR system, where the convolution profile corresponding to the control action along the unstable zeros, computed backward in time, is suitably stored.

The overall system in Fig. 9(a) is dualized as shown in Fig. 9(b), thus obtaining an unknown-input observer with delay. The FIR system included in Σ_c to steer the system along the unstable zeros is simply dualized by transposing its convolution profile at each time instant.

B. Disturbance decoupling by dynamic output feedback

Let us now focus on the following extension to the inaccessible disturbance decoupling problem with state feedback considered in Section IV-A.1. Consider the continuous-time

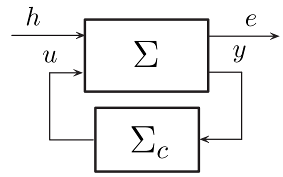

Fig. 11. Disturbance decoupling by dynamic output feedback.

LTI system Σ with two inputs and two outputs shown in Fig. 11 and described by

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,u(t) + H\,h(t), \\ y(t) &= C\,x(t), \\ e(t) &= E\,x(t), \end{aligned}$$

where u denotes the manipulable input and h the disturbance input. Let $\mathcal{B} = \mathrm{im} B$, $\mathcal{H} = \mathrm{im} H$, $\mathcal{C} = \ker C$ and $\mathcal{E} = \ker E$. Σ_c denotes a feedback dynamic compensator, described by

$$\begin{aligned} \dot{z}(t) &= N\,z(t) + M\,y(t), \\ u(t) &= L\,z(t) + K\,y(t). \end{aligned}$$

The disturbance decoupling problem by dynamic output feedback is set as follows: *Design, if possible, a dynamic compensator (N, M, L, K) such that the disturbance h has no influence on the regulated output e and the overall system is stable.*

This problem is of key importance in control. The structural conditions for its solution were investigated in [3], [16] and stated in precise terms in [30], while conditions including stability were first stated in [34] and restated in terms of self-bounded controlled invariant subspaces and their duals in [10], [8]. The structural condition is

$$\mathcal{S}^*_{(\mathcal{C},\mathcal{H})} \subseteq \mathcal{V}^*_{(\mathcal{B},\mathcal{E})}. \tag{20}$$

Let us recall definition (14) of $\mathcal{V}_m$, the minimum self-bounded controlled invariant of Σ containing $\mathcal{H}$, and definition (19) of $\mathcal{S}_M$, its dual. The stabilizability conditions are

$$\begin{aligned} &\mathcal{S}_M \text{ externally stabilizable}, \\ &\mathcal{V}_M = \mathcal{V}_m + \mathcal{S}_M \text{ internally stabilizable.} \end{aligned} \tag{21}$$

Conditions (20), (21) are constructive, since they trace the way to build a full-order unknown-input observer providing information to make $\mathcal{V}_M$ a locus of state trajectories due to h while stabilizing the overall system.

C. Model following

The model following problem has a long history in control theory. In the geometric framework, it was first addressed in [26] and refined in [12], [18]. Model following is a particular case of measurable signal decoupling, and its solution in this context is straightforward and appealing. The use of self-bounded controlled invariant subspaces and the geometric interpretation of invariant zeros has made the most recent contributions very complete. In fact, they consider both feedforward and feedback architectures, and the nonminimum-phase case. In what follows, the theory developed in [22] will be briefly reviewed. With reference to Fig. 12, the model

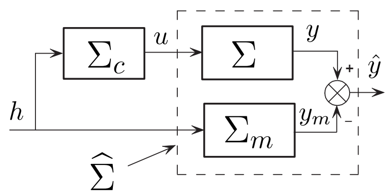

Fig. 12. Model following.

following problem can be stated as follows: *Determine, if possible, a dynamic feedforward compensator Σ_c such that the output of the system Σ strictly follows (is equal to) the output of a given model Σ_m*. The block diagram shown in Fig. 12 is equivalent to that in Fig. 5, provided the model is considered as part of the controlled system.

Assume that Σ is described by the triple (A, B, C) and Σ_m by the triple (A_m, B_m, C_m). The overall system $\widehat{\Sigma}$ is then described by

$$\hat{A} = \begin{bmatrix} A & 0 \\ 0 & A_m \end{bmatrix},\ \hat{B} = \begin{bmatrix} B \\ 0 \end{bmatrix},\ \hat{H} = \begin{bmatrix} 0 \\ B_m \end{bmatrix},\ \hat{C} = \begin{bmatrix} C & -C_m \end{bmatrix}.$$

Both the system and the model are assumed to be stable, square, left and right invertible. The structural condition expressed by inclusion (17) is satisfied if

$$\rho(\Sigma) \leq \gamma(\Sigma_m), \tag{22}$$

i.e., if the relative degree ρ of Σ is less or equal to the minimum delay γ of Σ_m. The minimum delay of a triple (A, B, C) is defined as the minimum value of i such that $C\,A^i\,B$ is nonzero. Hence the structural condition is satisfied if a model is chosen with a sufficiently high minimum delay. Consider the stabilizability condition (16). If Σ and Σ_m have no coincident invariant zeros, it can be shown that the internal eigenvalues of $\widehat{\mathcal{V}}_m$ are the union of the invariant zeros of Σ and the eigenvalues of A_m, so that in general model following with stability is not achievable if Σ is nonminimum-phase. Hence the stabilizability condition to be considered together with (22) is

$$\mathcal{Z}(\Sigma) \subseteq \mathbb{C}_g. \tag{23}$$

Condition (23) can be evaded by properly replicating in Σ_m the unstable zeros of Σ. This can be achieved, for instance, by assuming a model Σ_m consisting of q independent single-input single-output systems all having the unstable invariant zeros of Σ as zeros, so that these

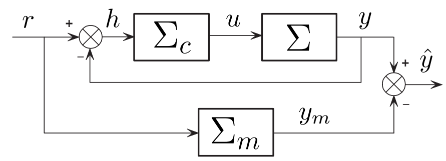

Fig. 13. Model following with feedback.

are cancelled as internal eigenvalues of $\widehat{\mathcal{V}}_m$. This allows to achieve both input-output decoupling and internal stability, but restricts the model choice. If Σ is nonminimum phase, perfect or almost perfect following of a minimum phase model may also be achieved if h is previewed by a significant time interval t_0, as pointed out in Section IV-A.3. In the geometric approach context, the model following problem with feedback, corresponding to the block diagram shown in Fig. 13, is also easily solvable. As in the feedforward case, both Σ and Σ_m are assumed to be stable and Σ_m to have at least the same relative degree as Σ. Replacing the feedback

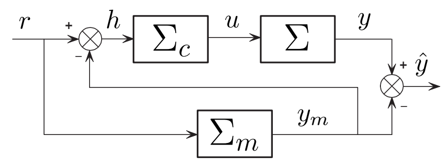

Fig. 14. A structurally equivalent connection.

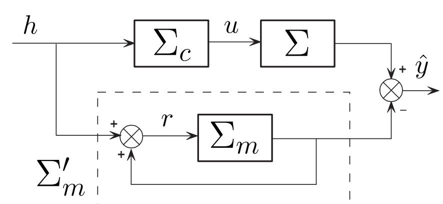

Fig. 15. Another structurally equivalent connection.

connection with that shown in Fig. 14 does not affect the structural properties of the system. However, it may affect stability. The new block diagram represents a feedforward model following problem. In fact, note that h is obtained as the difference of r (applied to the input of the model) and y_m (the output of the model). This corresponds to the parallel connection of Σ_m and a diagonal static system with gain -1, that is invertible, having zero relative degree. Its inverse is Σ_m with a feedback connection through the identity matrix, as shown in Fig. 15. Let the model consist of q independent single-input single-output systems all having the unstable invariant zeros of Σ as zeros. Since the invariant zeros of a system are preserved under state feedback connection, a feedforward model following compensator designed with reference to the block diagram in Fig. 15 does not include them as poles. It is also possible to include multiple internal models in the feedback connection shown in the figure (this is well known in the single input/output case), that are repeated in the compensator, so that both Σ'_m and the compensator may be unstable. In fact, zero output in the modified system may be obtained as the difference of diverging signals.

However, stability is recovered when going back to the original feedback connection illustrated in Fig. 13.

D. Noninteraction, fault detection and isolation

Another problem that in the geometric approach context was originally approached with state feedback (even though using measurable exogenous signals), is noninteracting control.

The noninteracting control problem is stated as follows: *Given an LTI system* Σ *whose output is partitioned by blocks* $(y_1, y_2, \ldots)$*, derive a controller with the same number of inputs* $(\alpha_1, \alpha_2, \ldots)$ *such that* α_i *allows complete reachability of output* y_i *while maintaining at zero all the other outputs*.

Only two output blocks y_1 and y_2 will herein be considered for the sake of simplicity. Therefore Σ is described by

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,u(t), \\ y_1(t) &= C_1\,x(t), \\ y_2(t) &= C_2\,x(t). \end{aligned}$$

This problem was first approached by Wonham and Morse in their first paper on geometric approach [36]. The solution was based on state feedback and static feedforward units, and was probably suggested by the measurable signal decoupling layout with static feedforward and feedback shown in Fig. 4. Achieving the most complete noninteraction with this technique is more restrictive than with other methods since, in general, the same state feedback cannot transform any two controlled invariant subspaces into simple $(A{+}BF)$-invariant subspaces. This drawback can be overcome by extending the state with a suitable bank of integrators as proposed by Wonham and Morse in their second paper [27].

An alternative solution, inspired by the measurable signal decoupling by means of a dynamic feedforward unit whose layout is shown in Fig. 5, was proposed in [6]. Refer to Fig. 16 and suppose now that the controlled system Σ is stable. Let $y_i \in \mathbb{R}^{q_i}$ and $\mathcal{C}_i = \ker C_i$ $(i = 1, 2)$. The maximum subspace that can be reached from the origin while being invisible at output y_2 is $\mathcal{R}^*_{(\mathcal{B},\mathcal{C}_2)}$ and the maximum subspace that can be reached from the origin while being invisible at output y_1 is $\mathcal{R}^*_{(\mathcal{B},\mathcal{C}_1)}$. These subspaces are internally stabilizable (or, more exactly, pole assignable) controlled invariant subspaces whose dynamics can be reproduced in the feedforward units Σ_1 and Σ_2. Hence noninteraction is possible if and only if

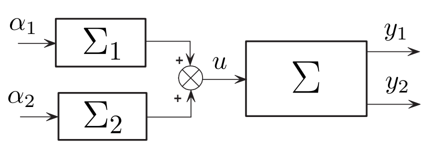

Fig. 16. Noninteracting control with dynamic feedforward units.

$$\begin{aligned} C_1\,\mathcal{R}^*_{(\mathcal{B},\mathcal{C}_2)} &= \mathbb{R}^{q_1}, \\ C_2\,\mathcal{R}^*_{(\mathcal{B},\mathcal{C}_1)} &= \mathbb{R}^{q_2}. \end{aligned}$$

There is a certain degree of freedom in the choice of inputs α_1 and α_2. They can be assumed of full dimension, i.e., corresponding to input matrices spanning the whole subspaces $\mathcal{R}^*_{(\mathcal{B},\mathcal{C}_2)}$ and $\mathcal{R}^*_{(\mathcal{B},\mathcal{C}_1)}$. On the other hand, since only dynamic reachability is required, owing to the well known Heymann lemma [13], they can also be assumed to be scalar without affecting the solvability of the problem. It is not necessary that the controlled system Σ be stable, but, similarly to the measurable disturbance decoupling problem, only stabilizability and detectability are required. The overall block diagram for this case, similar to that in Fig. 6, is shown in Fig. 17.

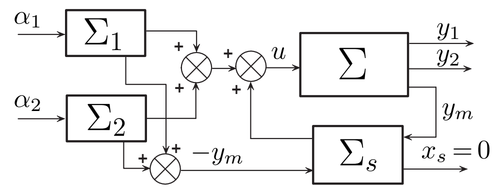

Fig. 17. Using a stabilizer in feedforward noninteracting control.

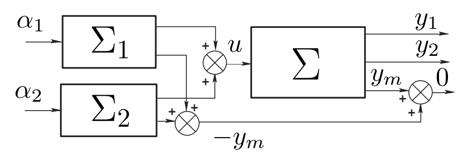

Fig. 18. Nulling a measurable output.

Note that the stabilizer shown in Fig. 17 is not influenced by the inputs α_1 and α_2 since the measured output y_m (from which Σ is detectable) can be nulled by a signal $-y_m$ generated in the feedforward units, where a replica of the state evolution produced by α_1 and α_2 is available. This idea is illustrated in Fig. 18.

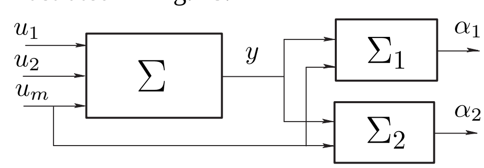

Fig. 19. A block diagram for fault detection and isolation.

Let us now consider the block diagram in Fig. 19, which clearly is the dual of that in Fig. 18. The model of Σ is

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B_m\,u_m(t) + B_1\,u_1(t) + B_2\,u_2(t), \\ y(t) &= C\,x(t), \end{aligned}$$

where u_m refers to a measurable input, while both u_1 and u_2 are assumed to be inaccessible. The *fault detection and isolation* (FDI) problem is stated as follows: *Given an LTI system* Σ *having an inaccessible input partitioned into blocks* $(u_1, u_2, \ldots)$*, derive an observer with the same number of scalar outputs* $(\alpha_1, \alpha_2, \ldots)$ *such that* α_i *is different from zero if any component of* u_i *is different from zero while all the other outputs are maintained at zero.*

A geometric solution to this problem was first proposed in [23] and restated in improved terms in [24].

V. Geometric approach to H_2-optimal regulation and filtering

The study of the Kalman *linear-quadratic regulator* (LQR) is the central topic of most courses and textbooks on advanced control systems. See, for instance, the books [15], [2], [17].

Recently, a special attention has been devoted to *H_2-optimal control*, which is substantially a reformulation of the LQR as a standard and well settled problem of the geometric approach (e.g., disturbance decoupling with output feedback). Feedthrough is not present in general, so that the standard Riccati equation-based solutions are not implementable and the existence of an optimal solution is not ensured. Contributions on this subject are [32] and [29]. The computational tools used to solve the H_2-optimal problem are linear matrix inequalities (LMI), supported by a "special coordinate basis" that points out the geometric features of the systems at hand.

An alternative approach is to treat the singular and cheap problems, where feedthrough is not present, by directly referring to the LTI system obtained by differentiating the Hamiltonian function, which can be considered as a generic dynamic system, with all the previously described features. This approach, developed in [21], [28], is herein briefly recalled.

A. Disturbance decoupling in H_2-norm

The H_2-norm of a continuous-time LTI system Σ represented by the differential equations (1) is defined as

$$\|\Sigma\|_2 = \sqrt{\mathrm{trace}\Big(\int_0^\infty g(t)\, g^{\mathrm{T}}(t)\, dt\Big)},$$

where $g(t)$ denotes the impulse response of the system.

Likewise, for a discrete-time LTI system Σ represented by the difference equations (2), the H_2-norm is

$$\|\Sigma\|_2 = \sqrt{\mathrm{trace}\Big(\sum_{k=0}^{\infty} g(k)\, g^{\mathrm{T}}(k)\Big)},$$

where the impulse response $g(k)$ is defined like in the continuous-time case, but referring to the autonomous system $x(k+1) = A\,x(k)$, $y(k) = C\,x(k)$.

The continuous-time case will be primarily considered herein. Only the main distinguishing differences will be added on for the discrete-time case.

The standard *linear quadratic regulator problem* (LQR) is stated as follows: *Given a stabilizable LTI system whose state evolution is described by*

$$\dot{x}(t) = A\,x(t) + B\,u(t)\,, \quad x(0) = x_0,$$

determine a control function $u(\cdot)$ *such that the corresponding state trajectory minimizes the* performance index

$$J = \int_0^\infty \big(x(t)^{\mathrm{T}} Q\, x(t) + u(t)^{\mathrm{T}} R\, u(t) + \, 2\, x(t)^{\mathrm{T}} S\, u(t)\big)\; dt, \tag{24}$$

where $\left[\begin{smallmatrix} Q & S \\ S^{\mathrm{T}} & R \end{smallmatrix}\right]$ *is symmetric positive semidefinite.* If $R > 0$ (positive definite), the problem is said to be *regular* and its solution is standard, if $R \geq 0$ (positive semidefinite) the problem is said to be *singular*, while if $R = 0$ the problem is said to be *cheap*.

It is well known that the regular LQR problem is solved by a state feedback F, independent of the initial state x_0.

The equivalence of the minimum H_2-norm disturbance rejection problem and the classical Kalman regulator problem is a simple consequence of the expression of the H_2-norm in terms of the impulse response of the triple (A, B, C). In fact, there exist matrices C and D such that

$$\begin{bmatrix} C^{\mathrm{T}} \\ D^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} C & D \end{bmatrix} = \begin{bmatrix} Q & S \\ S^{\mathrm{T}} & R \end{bmatrix}. \tag{25}$$

Consider the two-input system (12) with also a possible feedthrough, i.e.,

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,u(t) + H\,h(t), \\ y(t) &= C\,x(t) + D\,u(t), \end{aligned} \tag{26}$$

where C and D are defined in (25), see Fig. 3. System (26) is assumed to be left-invertible, not necessarily stabilizable, and with no zeros on the imaginary axis. It is easily seen that in our case the H_2-norm is the square root of (24) for the system $(A+BF, H, C+DF)$ described by

$$\begin{aligned} \dot{x}(t) &= (A+BF)\,x(t)\,, \quad x(0) = H, \\ y(t) &= (C+DF)\,x(t), \end{aligned} \tag{27}$$

where state and output are now matrices instead of vectors, so that it is minimized for any H by the Kalman feedback matrix F.

The LQR problem is solvable with the standard geometric tools. According to the classical optimal control approach, consider the *Hamiltonian function*

$$\begin{aligned} M(t) \,=\, & x(t)^{\mathrm{T}} Q\, x(t) + u(t)^{\mathrm{T}} R\, u(t) + 2\, x(t)^{\mathrm{T}} S\, u(t) \\ & + p(t)^{\mathrm{T}} \big(A\,x(t) + B\,u(t)\big), \end{aligned}$$

and set the state, costate equations and stationary condition as

$$\dot{x}(t) = \frac{\partial M(t)}{\partial p(t)}, \quad \dot{p}(t) = -\frac{\partial M(t)}{\partial x(t)}, \quad 0 = \frac{\partial M(t)}{\partial u(t)}.$$

We derive the following *Hamiltonian system*

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,u(t)\,, \quad x(0) = h_i, \\ \dot{p}(t) &= -2\,Q\,x(t) - A^{\mathrm{T}} p(t) - 2\,S\,u(t), \\ 0 &= 2\,S^{\mathrm{T}} x(t) + B^{\mathrm{T}} p(t) + 2\,R\,u(t), \end{aligned} \tag{28}$$

where h_i denotes a generic column of H, that can be rewritten in the more compact form as

$$\begin{aligned} \dot{\hat{x}}(t) &= \widehat{A}\,\hat{x}(t) + \widehat{B}\,u(t) + \widehat{H}\,h(t), \\ 0 &= \widehat{C}\,\hat{x}(t) + \widehat{D}\,u(t), \end{aligned} \tag{29}$$

with

$$\hat{x} = \begin{bmatrix} x \\ p \end{bmatrix}, \quad \widehat{H} = \begin{bmatrix} H \\ 0 \end{bmatrix},$$

$$\begin{aligned} \widehat{A} &= \begin{bmatrix} A & 0 \\ -2\,Q & -A^{\mathrm{T}} \end{bmatrix}, \quad \widehat{B} = \begin{bmatrix} B \\ -2\,S \end{bmatrix}, \\ \widehat{C} &= \begin{bmatrix} 2\,S^{\mathrm{T}} & B^{\mathrm{T}} \end{bmatrix}, \quad \widehat{D} = 2\,R. \end{aligned}$$

Equations (29) can be considered as referring to an LTI dynamic system whose output is constrained to be at zero. It follows that minimizing the H_2-norm of system (27) is equivalent to the perfect decoupling problem for the quadruple $(\widehat{A},\widehat{B},\widehat{C},\widehat{D})$, that admits a solution if and only if there is an internally stable $(\widehat{A},\widehat{\mathcal{B}})$-controlled invariant subspace $\widehat{\mathcal{V}}^*$, output-nulling for the overall extended system whose projection on the state space of the original system, defined as

$$P\,\widehat{\mathcal{V}}^* = \left\{ x \in \mathcal{X} \,:\, \begin{bmatrix} x \\ p \end{bmatrix} \in \widehat{\mathcal{V}}^* \right\},$$

contains the image of the matrix initial state H. It can be proved that the internal unassignable eigenvalues of $\widehat{\mathcal{V}}^*$ having nonzero real parts are stable-unstable by pairs. Hence a solution of the LQR problem is obtained through the following steps:

1) compute $\widehat{\mathcal{V}}^*$;
2) compute a matrix $\widehat{F}$ such that $(\widehat{A}+\widehat{B}\widehat{F})\widehat{\mathcal{V}}^* \subseteq \widehat{\mathcal{V}}^*$;
3) compute $\widehat{\mathcal{V}}_s$, the maximum internally stable $(\widehat{A}+\widehat{B}\widehat{F})$-invariant contained in $\widehat{\mathcal{V}}^*$ (this is a standard eigenvalue-eigenvector problem) and define $\mathcal{V}^*_{H_2}$ as $P\,\widehat{\mathcal{V}}_s$;
4) if $\mathcal{H} \in \mathcal{V}^*_{H_2}$, the problem admits a solution F that is computable directly from $\mathcal{V}^*_{H_2}$, which is $(A,\mathcal{B})$-controlled invariant.

The above procedure provides a state feedback matrix F corresponding to the minimum H_2-norm of the LTI system with input h and output y. This immediately follows from the previously recalled expression of the H_2-norm in terms of the impulse response. In fact, the impulse response corresponds to the set of initial states defined by the column vectors of matrix H.

Let us now briefly consider the extension to the discrete-time case, corresponding to the two-input system

$$\begin{aligned} x(k+1) &= A\,x(k) + B\,u(k) + H\,h(k), \\ y(k) &= C\,x(k) + D\,u(k), \end{aligned} \tag{30}$$

In this case the Hamiltonian function is

$$\begin{aligned} M(k) = x(k)^{\mathrm{T}} Q\,x(k) + u(k)^{\mathrm{T}} R\,u(k) + 2\,x(k)^{\mathrm{T}} S\,u(k) \\ + p(k+1)^{\mathrm{T}}\big(A\,x(k) + B\,u(k)\big), \end{aligned}$$

and the state, costate equations and stationary condition are

$$x(k+1) = \frac{\partial M(k)}{\partial p(k+1)}, \quad p(k) = \frac{\partial M(k)}{\partial x(k)}, \quad 0 = \frac{\partial M(k)}{\partial u(k)},$$

i.e.

$$\begin{aligned} x(k+1) &= A\,x(k) + B\,u(k), \quad x(0) = h_i, \\ p(k) &= 2\,Q\,x(k) + A^{\mathrm{T}}\,p(k+1) + 2\,S\,u(k), \\ 0 &= 2\,S^{\mathrm{T}}\,x(k) + B^{\mathrm{T}}\,p(k+1) + 2\,R\,u(k). \end{aligned}$$

Like in the continuous-time case, it is convenient to rewrite this system in the compact form

$$\begin{aligned} \hat{x}(k+1) &= \widehat{A}\,\hat{x}(k) + \widehat{B}\,u(k) + \widehat{H}\,h(k), \\ 0 &= \widehat{C}\,\hat{x}(k) + \widehat{D}\,u(k), \end{aligned}$$

with

$$\hat{x} = \begin{bmatrix} x \\ p \end{bmatrix}, \quad \widehat{H} = \begin{bmatrix} H \\ 0 \end{bmatrix},$$

$$\widehat{A} = \begin{bmatrix} A & 0 \\ -2\,A^{-\mathrm{T}} Q & -A^{-\mathrm{T}} \end{bmatrix}, \quad \widehat{B} = \begin{bmatrix} B \\ -2\,A^{-\mathrm{T}} S \end{bmatrix},$$

$$\begin{aligned} \widehat{C} &= \begin{bmatrix} -2\,B^{\mathrm{T}} A^{-\mathrm{T}} Q + 2\,S^{\mathrm{T}} & B^{\mathrm{T}} A^{-\mathrm{T}} \end{bmatrix}, \\ \widehat{D} &= 2\,R - 2\,B^{\mathrm{T}} A^{-\mathrm{T}} S. \end{aligned}$$

The drawback due to $A^{-\mathrm{T}}$ when A is singular can be overcome by using a stabilizing state feedback to be subtracted to the final state feedback solving the problem. The solution is obtained again with a geometric procedure, but, unlike the continuous-time case, this time a dead-beat like motion is also feasible and $\mathcal{V}^*_{H_2}$ covers the whole state space of system (30) if (A,B) is stabilizable. Hence the problem of minimizing the H_2-norm from h to y is always solvable in the discrete-time case.

1) The Kalman regulator: The Kalman regulator is the minimum H_2 norm extension of the exact disturbance decoupling problem with stability considered in Section IV-A.

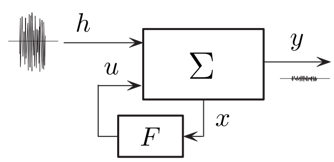

Fig. 20. Minimal H_2-norm decoupling (Kalman regulator).

Referring to Fig. 20, recall that the H_2 norm of a system $\Sigma : (A,B,C)$ is the mean power of the output signal when the input is white noise with zero mean and unitary variance. The problem of minimizing the H_2 norm from h to y has a solution if and only if

$$\mathcal{H} \subseteq \mathcal{V}^*_{H_2}. \tag{31}$$

Condition (31) replaces (13) and (16) for a minimal norm solution. It always holds in the regular case (since $\mathcal{V}^*_{H_2} = \mathbb{R}^n$ in this case), but also applies to the singular and cheap cases, where the dimension of $\mathcal{V}^*_{H_2}$ is always reduced. Notice that $\mathcal{V}^*_{H_2}$ is an internally stabilizable controlled invariant subspace, but for synthesis purposes it may be replaced with the minimum controlled invariant contained in it and containing $\mathcal{H}$ by using (14) with $\mathcal{V}^*_{H_2}$ instead of $\mathcal{C}$, to reduce the number of fixed modes.

2) The Kalman dual filter and the Kalman filter: Refer now to the problems considered in Section IV-A.2, i.e., the measurable signal decoupling and the unknown-input observation of a linear function of the state; let us consider their H_2-norm extensions, represented by the block diagrams shown in Fig. 21.

The necessary and sufficient condition for the solution of the H_2-norm problem shown in Fig. 21(a) is

$$\mathcal{H} \subseteq \mathcal{V}^*_{H_2} + \mathcal{B}, \tag{32}$$

similar to condition (17) for the exact decoupling. It directly ensures stability and also holds in the singular and cheap

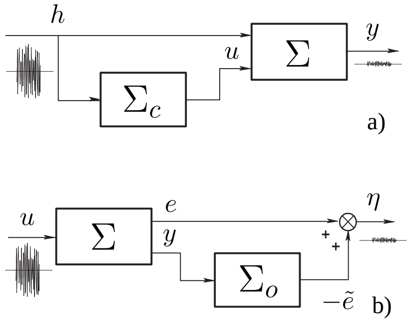

Fig. 21. (a) Kalman dual filter; (b) Kalman filter.

cases. The block diagram in Fig. 21(b) refers to the Kalman filter, which is here deduced by duality. The transpose of the matrix on the right-hand side of (25) represents the covariance matrix of a global white noise injected into the state and the output of Σ. The singular and cheap cases correspond to incomplete or absent measurement noise (noise injected at the output).

VI. Concluding remarks

In this tutorial paper, we have provided a broad overview of the main tools, problems and solutions of the geometric approach. The presentation of the results and selection of the arguments reflect the personal view of the authors. The main goal of our work is not completeness: rather, we aimed at offering a higher perspective of the key topics of geometric control theory, in order to encourage discussion and stimulate interest. To conclude, we wish to emphasize that research on geometric control theory is not in its twilight. In fact, a number of challenging open problems are still on the table. It lacks, for instance, a general theory accounting for input/state/output constraints. Preliminary steps have been recently taken to address the spectral and J-spectral factorization problems from a geometric viewpoint [20].

References

[1] B. D. O. Anderson. Output-nulling invariant and controllability subspaces. In *Proc. 6th IFAC Congress*, Boston, 1975. Paper 43.6.

[2] B. D. O. Anderson and J. B. Moore. *Linear Optimal Control*. Prentice-Hall, Englewood Cliffs, N.J., 1971.

[3] G. Basile, R. Laschi, and G. Marro. Invarianza controllata e non interazione nello spazio degli stati. *L'Elettrotecnica*, 51(1):1–5, 1969.

[4] G. Basile and G. Marro. Controlled and conditioned invariant subspaces in linear system theory. *J. Optimiz. Theory Appl.*, 3(5):306–315, 1969.

[5] G. Basile and G. Marro. On the observability of linear, time-invariant systems with unknown inputs. *J. Optimiz. Theory Appl.*, 3(6):410–415, 1969.

[6] G. Basile and G. Marro. A state space approach to noninteracting controls. *Ricerche di Automatica*, 1(1):68–77, 1970.

[7] G. Basile and G. Marro. Self-bounded controlled invariant subspaces: A straightforward approach to constrained controllability. *J. Optimiz. Theory Appl.*, 38(1):71–81, 1982.

[8] G. Basile and G. Marro. *Controlled and Conditioned Invariants in Linear System Theory*. Prentice Hall, Englewood Cliffs, NJ, 1992.

[9] G. Basile and G. Marro. L'invarianza rispetto ai disturbi studiata nello spazio degli stati. In *Rendiconti della LXX Riunione Annuale AEI*, Rimini, Italy, 1969. Paper 1.4.01.

[10] G. Basile, G. Marro, and A. Piazzi. Stability without eigenspaces in the geometric approach: Some new results. In C. A. Byrnes and A. Lindquist, editors, *Frequency Domain and State Space Methods for Linear Systems*, pages 441–450. North-Holland (Elsevier), Amsterdam, 1986.

[11] S. P. Bhattacharyya. Disturbance rejection in linear systems. *Int. J. Syst. Sci.*, 5(7):633–637, 1974.

[12] J. Descusse and M. Malabre. Model-following using measured output feedback. *IEEE Trans. Automat. Contr.*, AC-26(3):791–795, 1981.

[13] M. Heymann. Comments "On pole assignment in multi-input controllable linear systems". *IEEE Trans. Automat. Contr.*, AC-13(6):748–749, 1968.

[14] R. E. Kalman. On the general theory of control systems. In *Proc. 1st IFAC Congress*, volume 1, pages 481–492, 1961, Butterworths.

[15] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems*. John Wiley & Sons, New York, 1972.

[16] R. Laschi and G. Marro. Alcune considerazioni sull'osservabilità dei sistemi dinamici con ingressi inaccessibili. In *Rendiconti della LXX Riunione Annuale AEI*, Rimini, Italy, 1969. Paper 1.1.06.

[17] F. L. Lewis and V. L. Syrmos. *Optimal Control*. John Wiley & Sons, New York, 1995.

[18] M. Malabre and V. Kučera. Infinite structure and exact model matching problem: a geometric approach. *IEEE Trans. Automat. Contr.*, AC-29(3):266–268, 1984.

[19] G. Marro. The Geometric Approach Tools, 2004. Available from the web page: `http://www3.deis.unibo.it/Staff/FullProf/GiovanniMarro/downloads.htm`.

[20] G. Marro, F. Morbidi, and D. Prattichizzo. A Geometric Solution to the Cheap Spectral Factorization Problem. In *Proc. European Contr. Conf.*, pages 814–819, 2009.

[21] G. Marro, D. Prattichizzo, and E. Zattoni. A geometric insight into the discrete-time cheap and singular (LQR) problems. *IEEE Trans. Automat. Contr.*, 47(1):102–107, 2002.

[22] G. Marro and E. Zattoni. Self-bounded controlled invariant subspaces in model following by output feedback: minimal-order solutions for nonminimum-phase systems. *J. Optimiz. Theory Appl.*, 125(2):409–429, 2005.

[23] M. Massoumnia. A geometric approach to the synthesis of failure detection filters. *IEEE Trans. Automat. Contr.*, AC-31(9):839–846, 1986.

[24] M. Massoumnia, G. C. Verghese, and A. S. Willsky. Failure detection and identification. *IEEE Trans. Automat. Contr.*, 34(3):316–321, 1989.

[25] A. S. Morse. Structural invariants of linear multivariable systems. *SIAM J. Control*, 11(3):446–465, 1973.

[26] A. S. Morse. Structure and design of linear model-following systems. *IEEE Trans. Automat. Contr.*, AC-18(4):346–354, 1973.

[27] A. S. Morse and W. M. Wonham. Decoupling and pole assignment by dynamic compensation. *SIAM J. Control*, 8(3):317–337, 1970.

[28] D. Prattichizzo, L. Ntogramatzidis, and G. Marro. A new approach to the cheap LQ regulator exploiting the geometric properties of the Hamiltonian system. *Automatica*, 44(11):2834–2839, 2008.

[29] A. Saberi, P. Sannuti, and B.M. Chen. *H_2 Optimal Control*. Prentice Hall, New York, 1995.

[30] J. M. Schumacher. Compensator synthesis using (C, A, B)-pairs. *IEEE Trans. Automat. Contr.*, AC-25(6):1133–1138, 1980.

[31] J. M. Schumacher. On a conjecture of Basile and Marro. *J. Optimiz. Theory Appl.*, 41(2):371–376, 1983.

[32] A. Stoorvogel. The singular H_2 control problem. *Automatica*, 28(3):627–631, 1992.

[33] H. L. Trentelman, A. A. Stoorvogel, and M. Hautus. *Control Theory for Linear Systems*. Communications and Control Engineering. Springer, 2001.

[34] J. C. Willems and C. Commault. Disturbance decoupling by measurement feedback with stability or pole placement. *SIAM J. Control Optim.*, 19(4):490–504, 1981.

[35] W. M. Wonham. *Linear Multivariable Control: A Geometric Approach*. Springer-Verlag, New York, 3rd edition, 1985.

[36] W. M. Wonham and A. S. Morse. Decoupling and pole assignment in linear multivariable systems: a geometric approach. *SIAM J. Control*, 8(1):1–18, 1970.