Algebraic Control Theory for Linear Multivariable Systems

Pernebo, Lars

1978

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Lars Pernebo Algebraic Control Theory for Linear Multivariable Systems

CODEN: LUTFD2/(TFRT-1016)/1-307/(1978)

Algebraic Control Theory for Linear Multivariable Systems

LARS PERNEBO

Department of Automatic Control · Lund Institute of Technology

Dokumentutgivare
Lund Institute of Technology
Handläggare Dept of Automatic Control
Karl Johan Åström
Författare
Lars Pernebo

Dokumentnamn
REPORT
Utgivningsdatum
May 1978

Dokumentbeteckning
LUTFD2/(TFRT-1016)/1-307/(1978)
Ärendebeteckning

Dokumenttitel och undertitel

Algebraic Control Theory for Linear Multivariable Systems

Referat (sammandrag)

This thesis deals with linear, time invariant, finite dimensional, multivariable systems. Discrete time systems are analysed in part I. A Laurent polynomial system matrix is introduced to describe the system. An equivalence relation is defined and the invariants under equivalence are investigated. A theory for design of controllers for dynamical systems is presented in part II. The approach is purely algebraic. Both continuous and discrete time systems are considered. The servo and regulator problems are formulated. Necessary and sufficient conditions for solutions are given.

Referat skrivet av
Author

Förslag till ytterligare nyckelord

Klassifikationssystem och -klass(er)

Indextermer (ange källa)
Control theory, Linear systems, Difference equations, Equivalence, Feedback control, Transfer functions (Thesaurus of Engineering and Scientific Terms, Engineers Joint Council, N.Y., USA).

Omfång
307 pages

Övriga bibliografiska uppgifter

Språk
English

Sekretessuppgifter

ISSN

ISBN

Dokumentet kan erhållas från
Department of Automatic Control
Lund Institute of Technology
Box 725, S-220 07 Lund 7, Sweden

Mottagarens uppgifter

Pris

DOKUMENTDATABLAD enligt SIS 62 10 12

SIS-DB 1

Blankett LU 11:25 1976-07

Lars Pernebo

Algebraic Control Theory for Linear Multivariable Systems

Lund 1978

ALGEBRAIC CONTROL THEORY

FOR

LINEAR MULTIVARIABLE SYSTEMS

av

Lars Pernebo

Civ ing, Hb

Akademisk avhandling som för avläggande av teknisk doktorsexamen vid tekniska fakulteten vid universitetet i Lund kommer att offentligen försvaras i sal M:A, Maskinhuset, Lunds Tekniska Högskola, torsdagen den 1 juni 1978 kl 10.15.

Thesis for the degree of "Tekn. Dr."
in Automatic Control
at Lund Institute of Technology

TABLE OF CONTENTS

Algebraic Control Theory

The theory of multivariable, linear, time invariant systems has developed rapidly during the last two decades. The theory may be divided into analysis of dynamical systems and design of controllers for dynamical systems. Algebraic concepts are playing an important role in the theory for both analysis and design. One advantage of the algebraic approach is that it often allows a simultaneous treatment of continuous and discrete time systems. The algebraic approach was introduced by Kalman in a series of papers leading to part 4 of Kalman, Falb, Arbib (1969).

In part 4 of Kalman, Falb, Arbib (1969) Kalman presents an abstract theory for the analysis of linear, multivariable systems. He regards the state space of the system as a module over the polynomials and studies problems concerning realizability, controllability and observability. Another approach is presented in Rosenbrock (1970). Rosenbrock starts with a polynomial matrix description of a system. The system is analysed with algebraic methods and the concepts of controllability and observability appear in the form of decoupling zeros. The connection between the works of Kalman and Rosenbrock is not immediately clear. It has, however, to some extent been clarified by Fuhrmann (1976) and (1977). A third approach to the theory for analysis of linear, time invariant, multivariable systems was presented in a series of papers by Wonham and Morse, leading to Wonham (1975). In Wonham (1975) the system is supposed to be described in state space form and linear vector space algebra is used for the analysis.

A variety of design methods for multivariable systems has been developed. The linear quadratic control method, originated in Kalman (1960), is one of the most widely spread. It is described in many books, e.g. in Anderson, Moore (1971). The method leads to a state feedback control law and has been followed by other design methods that use state feedback, e.g. pole placement and the method in Wonham (1975). Wonham studies the output regulation problem and the noninteraction problem from a purely algebraic point of view. The design method is based on the vector space algebra of Wonham (1975). All the methods that lead to state feedback use dynamic observers if the state cannot be measured. In Wolovich (1974) the polynomial matrix analogue of linear state feedback and observer theory is developed.

The design methods, developed by Rosenbrock (1974) and Mac Farlane, Belletrutti (1973), are of a completely different type. Both methods are generalizations of the frequency domain methods for scalar systems by Nyquist and Bode. The idea is to use compensators which make the system "almost diagonal" and use classical theory to design scalar controllers in each loop.

In this thesis a purely algebraic approach is taken and contributions are made to the theory of both analysis and design. Part I and part II of this thesis can be read independently of each other.

Analysis

The two main concepts in the theory for analysis of dynamical systems are the system description and the equivalence relation. Examples of system descriptions are state equations and higher order differential equations. The system description must be general enough to cover all systems of interest. When a system description has been chosen a fundamental problem is to determine if two different sets of equations represent two different systems or if they can be regarded as two representations of the same dynamical system. In other words, an equivalence relation should be defined. When this is done the equivalence classes can be identified with the dynamical systems. It is of great interest to determine the invariants under equivalence because invariants represent properties of the system and not only properties of its representations. It is of course desirable to determine a complete set of invariants, because this set would completely describe the dynamical system. This seems to be a difficult problem that is not yet solved for the most commonly used equivalence relations like e.g. state transformations.

In the so called classical control theory a system was described by a transfer function. The equivalence relation was trivial and never mentioned formally. Each equivalence class consists of only one transfer function.

At the beginning of the sixties state space equations were introduced to describe a dynamical system. All state equations that can be obtained from a given set of state equations through coordinate changes in the state space are considered as equivalent. It was found that equivalent state equations have the same transfer function but the converse is not true. It was shown by Kalman that the reason is that there exist uncontrollable and unobservable parts of the state space. These parts are invariant under equivalence, but do not affect the transfer function.

Are the state equations a system description that is general enough to cover all cases of interest? The question was asked by Rosenbrock and answered with "no". State equations are equations of first order, but the differential equations that describe a physical system may be of higher order. In Rosenbrock (1970) a polynomial system matrix is introduced as a system description. The polynomial system matrix is general enough to cover sets of high order linear differential equations as well as state equations. Rosenbrock defined an equivalence relation, called strict system equivalence (s.s.e.), based on the transformations that were usually employed in order to bring a system of high order differential equations to state space form. It was shown that all invariants under coordinate changes in the state space are also invariants under s.s.e. This result gives these invariants a much more general validity. They are not introduced by the transformations that are employed when the high order equations are brought to state space form.

Rosenbrock's polynomial system matrix is general enough to describe linear, time invariant, finite dimensional, continuous time systems. Is the polynomial system matrix general enough to describe discrete time systems? The question is answered with "no" in part I of this thesis. The differential operator in the description of continuous time systems is replaced by the forward shift operator in the description of discrete time systems. A polynomial system matrix can therefore be used only to describe difference equations containing variables that are shifted forward in time. The difference equations that describe a physical system do, however, often contain past values of the variables. Such equations can not be represented by a polynomial system matrix.

In part I of this thesis a Laurent polynomial system matrix is introduced. A Laurent polynomial is a finite linear combination of positive and negative powers of the indeterminate. The Laurent polynomial system matrix is general enough to describe difference equations containing both past and future values of the variables as well as state equations. An equivalence relation is defined. The definition is based on the transformations needed to bring a set of difference equations, containing both past and future values of the variables, to state space form. The equivalence transformations include both forward and backward time shifts of the equations and the variables. Such transformations were introduced and investigated in Sinervo, Blomberg (1971). It is shown that the invariants under coordinate changes in the state space are <u>not</u> invariant under the equivalence relation defined here. This means, for instance, that a part of the uncontrollable or unobservable part of the system, as defined in the state space framework, may have been introduced by the transformations that were used to bring the equations to state space form. This is clearly an unsatisfactory situation. Concepts like for instance the system order, controllability and observability are therefore redefined so that they become invariant under the equivalence relation.

Design

Assume that a dynamical system is given. It is then of interest to construct a theory that shows what can be achieved if a controller is applied to the system. Two issues are of fundamental importance in such a theory. The first issue is to determine a realistic system description. It has to be determined how the system interacts with its environment. It should, for instance, be specified which variables that can be measured and which control variables that are available. The description should be general enough to cover all cases of interest. The second issue is to clearly specify the class of admissible controllers.

When the system description and the class of admissible controllers have been specified a suitable mathematical machinery has to be chosen. The most important requirement on the mathematical machinery is that it should be matched to the class of admissible controllers. It should be such that the admissible controllers are separated from the nonadmissible in a natural way.

A design theory, of the type outlined here, is presented in part II of this thesis. The theory covers both continuous time and discrete time systems. The given dynamical system is assumed to be described by the input-output relation of the box S in figure 1.

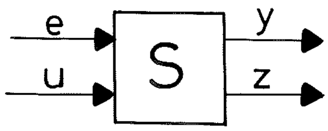


Figure 1. The given dynamical system.

The system is supposed to be linear, time invariant, finite dimensional and causal. It interacts with its environment in the following way. The input u is the control input and e is a nonmeasurable disturbance. The output y consists of the variables to be controlled and z consists of the measured variables. The class of admissible controllers consists of all linear, time invariant, finite dimensional and causal controllers R that stabilize the system in figure 2. The input u_r is a command input.

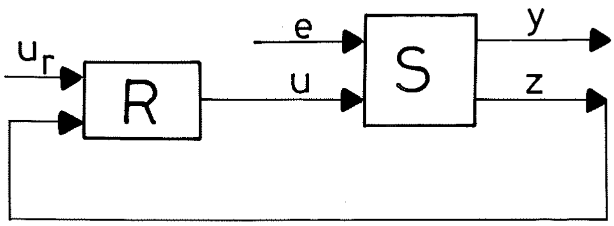


Figure 2. The control configuration.

A mathematical structure that suits the chosen class of admissible controllers is the ring of generalized polynomials. A generalized polynomial can be identified with a stable and causal rational transfer function. The ring of generalized polynomials is a principal ideal domain and many useful results for principal ideal domains are available in the literature on algebra.

In the design theory of part II of this thesis it is examined what can be achieved if admissible controllers R are applied to a given system S as in figure 2. The class $\mathcal{H}$ of achievable transfer functions from u_r to y and the class $\mathcal{F}$ of achievable transfer functions from e to y are determined. It is shown that $\mathcal{H}$ and $\mathcal{F}$ are independent in the following sense. For any $H \in \mathcal{H}$ and $F \in \mathcal{F}$ there is an admissible controller R, such that the transfer function from u_r to y is H and the transfer function from e to y is F. The servo problem is solved in the following way. Necessary and sufficient conditions are given for $\mathcal{H}$ to contain certain types of transfer functions, e.g. diagonal transfer functions. The regulator problem is solved in an analogous way. Necessary and sufficient conditions are given for $\mathcal{F}$ to contain certain types of transfer functions, e.g. transfer functions which do not transmit certain specified disturbances, or transfer functions with poles within a specified region.

Acknowledgements

I wish to express my sincere gratitude to my supervisor Professor Karl Johan Åström for his stimulating guidance and encouraging support throughout this work.

I also wish to thank my colleagues at the Department of Automatic Control in Lund for creating a stimulating environment for research. In particular I want to thank Bo Egardt and Per Hagander for many stimulating discussions and for valuable comments on the manuscript.

Many thanks also to Gudrun Christensen, Eva Dagnegård, Gunnel Sotiriou and Kerstin Ulveland for typing the manuscript and to Britt-Marie Carlsson for preparing the figures.

This work was partly supported by the Swedish Board for Technical Development under contracts 75-3776 and 77-3548.

References

Anderson B D O, Moore J B (1971): Linear Optimal Control. Prentice-Hall, Englewood Cliffs, N J.

Fuhrmann P A (1976): Algebraic System Theory: An Analyst's Point of View. Journal of the Franklin Institute 301, 521-540.

Fuhrmann P A (1977): On Strict System Equivalence and Similarity. International Journal of Control 25, 5-10.

Kalman R E (1960): Contributions to the Theory of Optimal Control. Boletin de la Sociedad Matematica Mexicana, 102-119.

Kalman R E, Falb P L, Arbib M A (1969): Topics in Mathematical System Theory. McGraw-Hill, New York.

Mac Farlane A G J, Belletrutti J J (1973): The Characteristic Locus Design Method. Automatica 9, 575-588.

Rosenbrock H H (1970): State-space and Multivariable Theory. Nelson, London.

Rosenbrock H H (1974): Computer-aided Control System Design. Academic Press, London.

Sinervo J, Blomberg H (1971): Algebraic Theory for Ordinary Linear Time-Invariant Difference Systems. Acta Polytechnica Scandinavica MA 21, Helsinki, Finland.

Wolovich W A (1974): Linear Multivariable Systems. Springer-Verlag, New York.

Wonham W M (1975): Linear Multivariable Control. Springer-Verlag, Berlin.

Part I - Analysis of Discrete Time Systems

1. INTRODUCTION

The idea of using polynomial matrices to describe linear, time invariant, multivariable systems was introduced by Rosenbrock. In Rosenbrock (1970) the analysis is mainly done for continuous time systems. Discrete time systems are only discussed briefly. In part 1 of this thesis discrete time systems are analysed from a point of view that is slightly different from Rosenbrock's.

Many authors have considered discrete time systems. Kalman, in Kalman, Falb, Arbib (1969), introduces a module theoretic approach and the analysis is made from an input-output point of view. Fuhrmann (1976) and (1977) generalizes this to internal system descriptions and gives the connection to Rosenbrock's work. The algebra, used by Fuhrmann to describe discrete time systems, is the same as the algebra for continuous time systems.

There is a connection between the spaces used to describe the input and output signals and the system algebra. This was emphasized in Sinervo, Blomberg (1971). It is shown that it is reasonable to use different algebras for discrete and continuous time systems. This point of view is adopted in this thesis.

Consider a linear, time invariant, discrete time system, described by a set of difference equations. The equations may contain both future and past values of the variables. Therefore, the system can be written as

$$T(q)\xi = U(q)u \tag{1.1 a}$$

$$y = V(q)\xi + W(q)u, \tag{1.1 b}$$

where q is the forward shift operator, defined through

$$qx(t) = x(t+1). \tag{1.2}$$

Furthermore, T(q), U(q), V(q) and W(q) are matrices, with entries that are Laurent polynomials. A Laurent polynomial is a finite linear combination of positive and negative powers of q. The formal definition is made in Chapter 2. The matrix T(q) is assumed to be square and have a nonzero determinant.

The system (1.1) is not included in the polynomial matrix description in Rosenbrock (1970) since negative powers of q are allowed. It is natural to allow negative powers of q because they give a causal relationship between the variables. Observe that if the entries of V(q) are polynomials of non-zero degree, then y does not depend causally on ξ.

For polynomial systems of the form (1.1) Rosenbrock (1970) defines concepts like the order of the system and decoupling zeros, i.e. uncontrollable or unobservable modes. A system matrix is introduced and an equivalence relation, called strict system equivalence (s.s.e.), is defined. It is desirable to make corresponding definitions for the Laurent polynomial description (1.1). This is done in this thesis in the following way.

Before the system order and the controllability and observability concepts are defined a state of the system will be defined. The state of the system is "the least amount of information that is needed to determine the future behaviour of the system if the future input is known". It has to be decided what is meant by "the future behaviour of the system". This can only be done if the vector spaces for the time sequences u, ξ, and y have been specified. Different possible vector spaces are discussed in Sinervo, Blomberg (1971). In Chapter 3 of this thesis it is shown that it is reasonable to use the space of all R^n-valued time sequences, defined on all positive and negative integers. When this choice has been made the solutions to the difference equations (1.1) can be analysed and a state can be defined.

Based on the definition of the state the definitions of the system order, decoupling zeros, controllability and observability indices are made in Chapters 3 and 4. Stability and causality are also defined from the solutions of (1.1). It is shown how all these definitions can be expressed in terms of the matrices T(q), U(q), V(q) and W(q).

In Chapter 4 the system matrix is introduced as in Rosenbrock (1970) and an equivalence relation is defined. The equivalence relation is defined so that there is an isomorfism between the sets of solutions to equivalent systems. This implies that the definitions above, which are based on the solutions to (1.1), are invariant under equivalence. It is shown that the equivalence relation can be described in the same way as s.s.e. The only difference is that Laurent polynomials are substituted for the ordinary polynomials. The analysis of the equivalence relation is very similar to the analysis of s.s.e. The reason is that the Laurent polynomials as well as the ordinary polynomials form a principal ideal domain.

The equivalence relation is more powerful than s.s.e. in the sense that the equivalence classes are larger. It is shown in Chapter 4 that each equivalence class contains systems in polynomial form and in state space form. It also includes systems where the entries of T(q), U(q), V(q) and W(q) are polynomials in the backward shift operator q^{-1}. In some cases it is suitable to use the polynomial form in q^{-1}. An example is the development of the minimum variance controller in Åström (1970). In other cases the polynomial form in q or the state space form is more suitable. It is therefore valuable to have an equivalence relation where each equivalence class contains all these forms.

It is possible to find certain simple forms of the system matrix in each equivalence class. In Chapter 4 it is shown how the system order, the decoupling zeros, the controllability and observability indices easily can be computed and how stability and causability can be checked from these forms of the system matrix.

Finally, it should be noted that the theory in part I of this thesis can be generalized in a straightforward manner. Generalized polynomials, as defined in part II, can be used instead of Laurent polynomials to describe the systems. The equivalence relation, as well as all definitions, should then be modified analogously. This would give a unified theory for systems described by system matrices. In particular the theories for analysis in Rosenbrock (1970) and in part I of this thesis would be included as special cases. Furthermore, a theory which disregards stable uncontrollable or stable unobservable modes would be obtained. Such a theory might be useful for design purposes. Some ideas in this direction are used in part II of this thesis.

2. PRELIMINARIES

In section 2.1 we will define Laurent polynomials. It is shown that they form a ring with essentially the same properties as the ring of polynomials. The Laurent polynomials will be used to describe linear dynamical systems. In section 2.2 two different vector spaces of time sequences are introduced. The action of the shift operator on elements in these spaces is analysed.

2.1 Laurent polynomials

Let R denote the field of real numbers, Z the ring of integers, Z_+ the positive integers and Z_{0+} the non-negative integers. Furthermore let R[x] denote the ring of polynomials in the indeterminate x with coefficients in R, i.e.

$$R[x] = \{ p(x) \mid p(x) = \sum_{i=0}^{n} p_i x^i, \quad p_i \in R, \quad n \in Z_{0+} \}$$

Def. 2.1.1 The set of Laurent polynomials R(x] is defined as

$$R(x] = \{ a(x) \mid a(x) = \sum_{i=n}^{m} a_i x^i, \quad a_i \in R, \quad n,m \in Z, \quad m \geq n \}$$

Remark. Observe that R(x] is a subset of R(x), rational functions.

Addition and multiplication in R(x] are defined as in R(x). It is easy to verify that R(x] is a ring.

Theorem 2.1.1 R(x] is an euclidean domain.

Proof. We have to verify that R(x] is an integral domain and that there is a function v from the nonzero elements of R(x] into the nonnegative integers such that

(i) For all pairs a,b from $R(x]$ for which $b \neq 0$ there exist q and r in $R(x]$ such that $a = bq+r$ and either $r = 0$ or $v(r) < v(b)$.

(ii) For all pairs a,b from $R(x]$ for which $a \neq 0$, $b \neq 0$, $v(a) \leq v(ab)$.

That $R(x]$ is an integral domain is obvious. An element a in $R(x]$ can be written

$$a = \sum_{i=n}^{m} a_i x^i$$

with $a_i \in R$ and $a_n \neq 0$ and $a_m \neq 0$. Define $v(a) = m-n$, then (ii) is satisfied. Using the division algorithm for polynomials it is easy to see that also (i) is satisfied.

□

Every Laurent polynomial $a(x)$ can be written $a(x) = x^k p(x)$, where $p(x)$ is a polynomial with $p(0) \neq 0$ and $k \in Z$. Of course $p(x)$ is uniquely determined by $a(x)$.

Def. 2.1.2 The degree of a Laurent polynomial $a(x)$ is defined as the degree of the polynomial $p(x)$ and denoted deg $a(x)$, while deg $p(x)$ denotes the degree of the polynomial $p(x)$.

Remark. The degree of $a \in R(x]$ is equal to $v(a)$ in the proof of theorem 2.1.1.

Def. 2.1.3 The zeros of a Laurent polynomial $a(x)$ are defined as the zeros of the polynomial $p(x)$.

Remark. Note that a Laurent polynomial has no zeros at the origin.

Since every euclidean domain is a principal ideal domain it follows that $R(x]$ is a principal ideal domain.

This insures that matrices with entries in $R(x)$ have properties analogous to those with entries in $R[x]$.

Let $R^{n\times m}(x)$ denote the set of all $n\times m$ matrices with entries in $R(x)$.

Def. 2.1.4 A matrix $A \in R^{n\times n}(x)$ is unimodular if there is a $B \in R^{n\times n}(x)$ such that $AB = I$.

It is well known (e.g. MacDuffee (1946)) that a matrix is unimodular if and only if its determinant is a unit in the ring of its entries. The units of $R(x)$ are Laurent polynomials of the form cx^k, where $c \in R \setminus \{0\}$ and $k \in Z$. Therefore we have:

Theorem 2.1.2 A matrix $A \in R^{n\times n}(x)$ is unimodular if and only if $\det A = cx^k$, where $c \in R \setminus \{0\}$ and $k \in Z$.

Remark. If A is a polynomial matrix, then it is unimodular if and only if $\det A = c$, where $c \in R \setminus \{0\}$. It is thus important to clearly state if a matrix is unimodular as a polynomial matrix or as a Laurent polynomial matrix. The matrix $\mathrm{diag}(x,1)$ is for instance a unimodular Laurent polynomial matrix, but not a unimodular polynomial matrix.

Def. 2.1.5 The matrices $A,B \in R^{n\times m}(x)$ are equivalent if there are unimodular matrices $U \in R^{n\times n}(x)$ and $V \in R^{m\times m}(x)$ such that $A = UBV$.

Since $R(x)$ is a principal ideal domain the following result is true (e.g. MacDuffee (1946)).

Theorem 2.1.3 Every matrix $A \in R^{n\times m}(x)$ is equivalent to a matrix S of the form

$$S = \begin{pmatrix} D & 0 \end{pmatrix} \quad \text{if} \quad m > n$$

$$S = D \quad \text{if} \quad m = n$$

$$S = \begin{pmatrix} D \\ 0 \end{pmatrix} \quad \text{if} \quad m < n$$

where $D = \text{diag}(i_1, i_2, \ldots, i_k, 0, \ldots, 0)$. $\{i_j\}_{j=1}^{k}$ are called invariant factors and have the property $i_j | i_{j+1}$. The invariant factors are unique up to units in $R(x]$.

The matrix S is called the Smith form of A. Let us require that $\{i_j\}$ are polynomials with leading coefficient 1 and $i_j(x) \neq 0$ for $x = 0, \forall j$. The Smith form is then unique.

□

Let A belong to $R^{n \times m}(x]$. The following operations on A are called elementary row operations.

- Multiply a row by a unit in $R(x]$, i.e. by cx^k, where $c \in R \setminus \{0\}$, $k \in Z$. (2.1.1)

- Add a multiple by a Laurent polynomial of one row to another row. (2.1.2)

- Interchange any two rows. (2.1.3)

It is clear that any elementary row operation can be obtained by multiplying A from the left by a unimodular matrix. Conversely, since $R(x]$ is a euclidean domain, multiplication of A from the left by a unimodular matrix gives a matrix which can be obtained by performing elementary row operations on A (MacDuffee (1946)). Corresponding results are true for elementary column operations and multiplication from the right by a unimodular matrix.

The Laurent polynomial matrices $T(x)$ and $U(x)$ are said to be relatively left prime if the only common left divisors are unimodular Laurent polynomial matrices.

The following theorem is formulated and proven for polynomial matrices in Rosenbrock (1970).

<u>Theorem 2.1.4</u> Let $T(x)$ and $U(x)$ be respectively $r \times r$ and $r \times \ell$ Laurent polynomial matrices with $\det T(x) \neq 0$. The following conditions are equivalent.

(i) $T(x)$ and $U(x)$ are relatively left prime.

(ii) The rank of $(T(x) \; U(x))$ is r for all nonzero complex x.

(iii) The Smith form of $(T(x) \; U(x))$ is $(I \;\; 0)$.

(iv) There exist Laurent polynomial matrices $V(x)$ and $W(x)$, respectively $\ell \times r$ and $\ell \times \ell$ such that the matrix

$$\begin{pmatrix} T(x) & U(x) \\ -V(x) & W(x) \end{pmatrix}$$

is unimodular.

(v) There exist relatively right prime Laurent polynomial matrices $X(x)$ and $Y(x)$, respectively $r \times r$ and $\ell \times r$, such that

$$T(x)\,X(x) + U(x)\,Y(x) = I$$

<u>Proof</u>. The theorem is proven as in Rosenbrock (1970), theorem 6.1 with the difference that all concepts related to polynomial matrices should be substituted by the corresponding concepts for Laurent polynomial matrices.

□

2.2 Shift operators and signal spaces

Let R^n be the vector space of n-tuples with components in R.

Def. 2.2.1 Let R^n_z be the set of all functions from Z to R^n.

Remark. R^n_z is a vector space over R if all operations are defined pointwise.

Def. 2.2.2 Let $\overline{R^n_z}$ be a subspace of R^n_z given by

$$\overline{R^n_z} = \{\ u \in R^n_z \mid \exists\ t_0 \in Z \text{ such that } t < t_0 \Rightarrow u(t) = 0\ \}.$$

Def. 2.2.3 The shift operator q is defined as a map from R^n_z to R^n_z by

$$qu(t) = u(t+1) \qquad u \in R^n_z, \quad t \in Z$$

Positive and negative powers of the shift operator are defined in the obvious way

$$q^k u(t) = u(t+k), \quad k \in Z.$$

A Laurent polynomial in the shift operator is now defined as

$$\left(\sum_{i=n}^{m} a_i\ q^i\right) u(t) = \sum_{i=n}^{m} a_i u(t+i) \qquad n,m \in Z.$$

The sets of polynomials and Laurent polynomials in the shift operator are denoted $R[q]$ and $R(q]$ respectively. Observe that $R(q^{-1}]$ is equal to $R(q]$. If $a(q) \in R(q]$ then $a(q)$ is a linear map from R^n_z to R^n_z and the restriction of $a(q)$ to $\overline{R^n_z}$ is a linear map from $\overline{R^n_z}$ to $\overline{R^n_z}$. It is of fundamental importance to determine which of these linear maps that are bijective. In Sinervo, Blomberg (1971) the following two theorems are proved.

<u>Theorem 2.2.1</u> Regard $a(q) \in R(q]$ as a map from R_z^n to R_z^n. Then $a(q)$ is bijective if and only if it is of the form $a(q) = c\, q^k$, $c \in R \setminus \{0\}$ and $k \in Z$.

<u>Remark</u>. The theorem says that $a(q) \in R(q]$ is bijective if and only if it is a unit in $R(q]$.

<u>Theorem 2.2.2</u> Regard $a(q) \in R(q]$ as a map from $\overline{R}_z^n$ to $\overline{R}_z^n$. Then $a(q)$ is bijective if and only if it is nonzero.

<u>Remark</u>. The last result may by a first sight seem strange. Observe, however, that the equation $(q+b)u=0$, $b \in R \setminus \{0\}$, has only the solution $u = 0$ in $\overline{R}_z^n$.

The last two theorems show how important it is to specify the signal spaces. Let $a_1(q)$ be a bijective map. Then the two equations

$$a_1(q)\, a_2(q)\, y = a_1(q)\, b(q)\, u \qquad u \text{ given}$$

and

$$a_2(q)\, y = b(q)\, u$$

have exactly the same solutions. In other words it is possible to cancel common factors if they represent bijective maps. Because of theorem 2.2.2 it also makes sense to represent the solution y to the equation

$$a(q)\, y = b(q)\, u \qquad u \text{ given}$$

as

$$y = \frac{b(q)}{a(q)}\, u$$

if we are working with the space $\overline{R}_z^n$.

A matrix $A(q) \in R^{n\times n}(q]$ can be regarded as a linear map from R_z^n to R_z^n or from $\overline{R}_z^n$ to $\overline{R}_z^n$. We will now generalize theorems 2.2.1 and 2.2.2 to matrices $A(q) \in R^{n\times n}(q]$.

<u>Theorem 2.2.3</u> Regard $A(q) \in R^{n\times n}(q]$ as a map from R_z^n to R_z^n. Then $A(q)$ is bijective if and only if it is unimodular.

<u>Proof</u>. Suppose $A(q)$ is unimodular. Then by definition $A^{-1}(q) \in R^{n\times n}(q]$. Therefore $A^{-1}(q)$ is a linear map from R_z^n to R_z^n. $A(q)$ is surjective because for any v in R_z^n the element $A^{-1}(q)v \in R_z^n$ is mapped by $A(q)$ to v. $A(q)$ is injective because if $A(q)u = 0$ then $A^{-1}(q)A(q)u = 0$, i.e. $u = 0$.

Conversely suppose that $A(q)$ is not unimodular. Then there is at least one $z \in C \setminus \{0\}$ such that $\det A(z) = 0$ and an $u_0 \neq 0$ such that $A(z)u_0 = 0$.
We have $A(z)u_0 = 0 \Leftrightarrow \overline{A(z)u_0} = 0 \Leftrightarrow A(\bar{z})\bar{u}_0 = 0$.
Choose $u = u_0 z^t + \bar{u}_0 \bar{z}^t \in R_z^n$ then $A(q)u = A(z)u_0 z^t + A(\bar{z})\bar{u}_0\bar{z}^t =$ $= 0$. I.e. $A(q)$ is not injective.

□

<u>Theorem 2.2.4</u> Regard $A(q) \in R^{n\times n}(q]$ as a map from $\overline{R}_z^n$ to $\overline{R}_z^n$. Then $A(q)$ is bijective if and only if $\det A(q) \neq 0$.

<u>Proof</u>. Suppose $\det A(q) \neq 0$. Then $A^{-1}(q)$ can be computed and $A^{-1}(q) \in R^{n\times n}(q)$, matrices whose elements are rational functions in q, which by theorem 2.2.2 are well defined linear maps from $\overline{R}_z$ to $\overline{R}_z$. Therefore $A^{-1}(q)$ is a well defined map from $\overline{R}_z^n$ to $\overline{R}_z^n$. That $A(q)$ is bijective follows as in the proof of theorem 2.2.3.

Conversely suppose that $\det A(q) = 0$. Suppose $A(q)$ has rank m, which is less than n. Then there is a nonzero minor of order m. Without loss of generality we can suppose that the upper left $m\times m$ submatrix has nonzero determinant. Then $A(q)$ can be partitioned as

$$A(q) = \begin{pmatrix} A_{11}(q) & A_{12}(q) \\ A_{21}(q) & A_{22}(q) \end{pmatrix}$$

where $A_{11}(q)$ is an $m \times m$ matrix and $\det A_{11}(q) \neq 0$. Furthermore the last n-m rows are linear combinations of the first m rows. I.e. there is a rational $(n-m) \times m$ matrix $B(q)$ such that

$$B(q) \begin{pmatrix} A_{11}(q) & A_{12}(q) \end{pmatrix} = \begin{pmatrix} A_{21}(q) & A_{22}(q) \end{pmatrix} \tag{2.2.1}$$

Let u_2 be an arbitrary nonzero element in ΓR_z^{n-m}. By the first part of the proof there is a unique solution $u_1 \in \Gamma R_z^m$ to the equation

$$A_{11}(q)\ u_1 = -A_{12}(q)\ u_2 \tag{2.2.2}$$

Now, $B(q)$ is a well defined linear map from ΓR_z^m to ΓR_z^{n-m}. Therefore (2.2.2) implies

$$B(q)\ A_{11}(q)\ u_1 = -B(q)\ A_{12}(q)\ u_2$$

or by (2.2.1)

$$A_{21}(q)\ u_1 = -A_{22}(q)\ u_2. \tag{2.2.3}$$

From (2.2.2) and (2.2.3) it now follows that

$$\begin{pmatrix} A_{11}(q) & A_{12}(q) \\ A_{21}(q) & A_{22}(q) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 0$$

Since $u_2 \neq 0$ it follows that $A(q)$ is not injective.

□

Theorem 2.2.5 If $A(q) \in R^{n \times m}(x]$ then

$$A(q)\ u = 0 \quad \forall\, u \in \Gamma R_z^m \quad \Rightarrow \quad A(q) = 0$$

Proof. Choose the i:th component of u equal to 1 at $t = 0$ and equal to 0 for all other t. Let all other components of u be equal to 0 for all $t \in Z$. Then $A(q)\ u = 0$ implies that the i:th column of $A(q)$ is zero. Repeat the argument for every i.

□

Remark. The theorem is true even if R_z^m is substituted for $\overline{R_z^m}$ since $\overline{R_z^m} \subset R_z^m$.

3. LINEAR TIME INVARIANT SYSTEMS

The class of systems to be considered is defined in section 3.1. In the following sections the concepts of system order, stability, causality, controllability and observability are defined. The definitions are made using only the time sequences of the system, i.e. the input, the internal variable and the output. It is then shown how the definitions can be expressed in terms of the matrices used to describe the difference equations.

3.1 The system description

The systems to be considered here are described by the difference equations

$$T(q)\xi = U(q)u \tag{3.1.1a}$$

$$y = V(q)\xi + W(q)u, \tag{3.1.1b}$$

where $T(q) \in R^{r\times r}(q]$, $U(q) \in R^{r\times \ell}(q]$, $V(q) \in R^{m\times r}(q]$, $W(q) \in R^{m\times \ell}(q]$, $u \in \overline{R}_z^{\ell}$, $\xi \in R_z^r$, and $y \in R_z^m$. Furthermore $\det T(q) \neq 0$. Here u is regarded as a given input and ξ is a solution to (3.1.1a) and is called the internal variable. Finally y, called the output, is uniquely given by ξ and u.

The system description (3.1.1) differs from that of Rosenbrock (1970) in that we allow the matrices $T(q)$, $U(q)$, $V(q)$, and $W(q)$ to be Laurent polynomial matrices while Rosenbrock demands them to be polynomial matrices. The restriction to polynomial matrices may cause trouble if we want to write down the equations for a given physical system. It may very well happen that the output depends on old values of the internal physical variable. This situation cannot be handled if $V(q)$ is restricted to be polynomial. The problem may be circumvented if we make a time translation of the internal variable. This is however not a transformation of strict system equivalence and not studied by Rosenbrock. It will be studied in chapter 4 of this work, where an equivalence transformation, slightly more general than strict system equivalence, is defined.

The question now arises why ξ and y are allowed to belong to R_z^r and R_z^m. An argument against this is that in a physical system all signals must start at some finite time. The consequence is that ξ and y should belong to $\overline{R}_z^r$ and $\overline{R}_z^m$. This would by theorem 2.2.4 imply that equation (3.1.1a) has a unique solution, which can be written as

$$\xi = T^{-1}(q)\ U(q)\ u.$$

Inserting this into (3.1.1b) gives

$$y = (V(q)T^{-1}(q)\ U(q) + W(q))\ u.$$

This means that the internal variable and the output are uniquely given by the input.

The system (3.1.1) is supposed to be a mathematical description of what we may call a process or a physical system. If the equations (3.1.1) describe the process exactly then the motivation above would imply that ξ belongs to $\overline{R}_z^r$ and therefore $y \in \overline{R}_z^m$. In most cases the process is not exactly described by (3.1.1). Some parameters in (3.1.1) may have incorrect values, noise may act on the process or some initial value may be given to the process variable ξ. We will show that one way to take such phenomena into account in the mathematical description (3.1.1) is to allow the solution ξ to belong to the larger space R_z^r. This will then imply that y belongs to R_z^m. Let us start with a simple example.

<u>Example 3.1.1</u> Let the mathematical description of a system be

$$(q-2)\xi=(q-2)u \tag{3.1.2a}$$

$$y = \xi \tag{3.1.2b}$$

where $u \in \overline{R}_z$. The solution to (3.1.2a) in $\overline{R}_z$ is

$$\xi(t) = u(t) \qquad t \in Z, \tag{3.1.3}$$

while the solution in R_z is

$$\xi(t) = u(t) + a\cdot 2^t \qquad t \in Z, \qquad a \in R. \tag{3.1.4}$$

Suppose that there is some kind of noise acting on the <u>process</u>, so that it is described by

$$(q-2)\xi = (q-2)u + e \tag{3.1.5a}$$

$$y = \xi \tag{3.1.5b}$$

If for instance $e(t)$ is equal to 1 for $t = t_0 \leq 0$ and 0 for all other t, then the solution for $t \geq 0$ is given by

$$\xi(t) = u(t) + b\cdot 2^t \qquad t \geq 0 \tag{3.1.6}$$

where $b = 2^{-t_0}$. Observe that $\xi(t) = 0$ for sufficiently large negative values of t.

Suppose that we, at time $t = 0$, want to solve equation (3.1.2a) to find out what might happen to the internal variable in the future. If we solve (3.1.2a) in $\overline{R}_z$ then the solution (3.1.3) gives the answer that ξ will follow u exactly in the future. The solution in R_z, i.e. (3.1.4), says that there might be a difference between ξ and u, which will tend to infinity as 2^t. Comparing with (3.1.6) we find that the last answer is correct if there has been disturbances acting on the system in the past.

□

The idea of this example can clearly be generalized to the system (3.1.1). A disturbance on the process in the past may drive the internal variable outside the set of solutions to (3.1.1a) in $\overline{R}_z^r$. Since we do not know how the disturbances influence the internal variable we have to consider the largest existing solution space, i.e. R_z^r.

For (3.1.1) to describe a physical system there has to be a causal relationship between u and y. We will later give necessary and sufficient algebraic conditions for this. If ξ is describing a set of physical variables then they have to depend causally on u, and y has to depend causally on ξ. We will however allow ξ to be mathematical variables with no physical significance. In this case there does not have to be a causal relationship between u and ξ or between ξ and y, but still between u and y.

3.2 The order of the system

Suppose that a system, described by equations (3.1.1), is observed at time $t = 0$. Then knowledge of the future input sequence $u(t)$, $t > 0$ is not sufficient to uniquely determine the future sequences of the internal variable ξ and the output y. A certain amount of information about the past behaviour is also needed. The least "number of parameters" needed together with the input $u(t)$, $t > 0$ to uniquely determine the future behaviour of the system is called the order of the system. The set of parameters is called the state of the system. We will later specify what shall be meant by "the future behaviour of the system".

The following lemma shows that the influence on the solutions from future inputs can be disregarded.

<u>Lemma 3.2.1</u> For any $u \in \overline{R}^{\ell}_{z}$ put $u_0(t) = \begin{cases} u(t), & t > 0 \\ 0 & t \leq 0 \end{cases}$ and $u_1 = u - u_0$. Then any solution $\xi \in R^r_z$ to (3.1.1a) can be written $\xi = \xi_0 + \xi_1$, where ξ_0 is the unique solution to

$$T(q)\ \xi_0 = U(q)\ u_0$$

in ${}^{\ulcorner}R^r_z$ and ξ_1 is some solution to

$$T(q)\ \xi_1 = U(q)\ u_1$$

in R^r_z.

<u>Proof</u>. The result follows from theorem 2.2.4 and the linearity of (3.1.1a).

□

In this section $u(t)$, $t > 0$ is supposed to be known and we want to examine how the solutions in the future are influenced by the behaviour of the system for $t \leq 0$. It therefore follows from lemma 3.2.1 that we can put $u(t) = 0$, $t > 0$ without loss of generality. Accordingly <u>it will be supposed throughout this section that $u(t) = 0$ for $t > 0$</u>, unless otherwise is stated.

<u>Def. 3.2.1</u> ${}^{\ulcorner}R^{n\urcorner}_{z} = \{u \in \overline{R}^n_z \mid u(t) = 0 \quad \text{for} \quad t > 0\}$.

We will need a few more definitions.

<u>Def. 3.2.2</u> Let X_u and Y_u be the set of internal variables and outputs respectively corresponding to the fixed input $u \in \overline{R}^{\ell}_{z}$. I.e. if

$$T(q)\xi = U(q)\ u \tag{3.2.1a}$$

$$y = V(q)\ \xi + W(q)\ u \tag{3.2.1b}$$

then $\xi \in X_u$ and $y \in Y_u$.

Remark. $\mathcal{X}_u$ and $\mathcal{Y}_u$ are not linear vector spaces (over R) for a general input u, but they are for u = 0, i.e. $\mathcal{X}_0$ and $\mathcal{Y}_0$.

Def. 3.2.3 Let v belong to R^n_z i.e. $v: Z \to R^n$. Then v_+ is the restriction of v to Z_+.

Def. 3.2.4 Let $\mathcal{Y}^+_u$ be the set of all y_+ such that $y \in \mathcal{Y}_u$ with $u \in \overline{R}^{\ell}_z$ and let $\mathcal{Y}^+ = \bigcup_{u \in \overline{R}^{\ell}_z} \mathcal{Y}^+_u$.

Remark. $\mathcal{Y}^+_u$ is generally not a linear vector space but it is easily shown that $\mathcal{Y}^+_0$ and $\mathcal{Y}^+$ are.

Remark. Observe that $\mathcal{X}_u$, $\mathcal{Y}_u$ and $\mathcal{Y}^+_u$ are defined for any u in $\overline{R}^{\ell}_z$, but in the definition of $\mathcal{Y}^+$ it is supposed that $u \in \overline{R}^{\ell}_z$.

Using the previous definitions we can define the order of the system using only properties of the solutions of the difference equations and not the equations themselves.

Def. 3.2.5 Denote $n_0 = \dim \mathcal{X}_0$ and $n_D = \dim (\mathcal{Y}^+/\mathcal{Y}^+_0)$. Then define the order n of the system (3.2.1) to be $n = n_0 + n_D$.

Remark. n_0 and n_D will be shown to be finite for the systems considered.

Remark. Since $\mathcal{Y}^+_0$ is a subspace of $\mathcal{Y}^+$ then the factor space $\mathcal{Y}^+/\mathcal{Y}^+_0$ is well defined and $\dim (\mathcal{Y}^+/\mathcal{Y}^+_0) = \dim \mathcal{Y}^+ - \dim \mathcal{Y}^+_0$.

Example 3.2.1 We want to determine the order of the system

$$(2\, q^3 + q^2)\ \xi = (2 + q^{-1})\ u$$

$$y = \xi + (q^{-1} - q^{-3})\ u$$

The vectorspace X_0 is given by the solutions to

$$(2q^3 + q^2)\ \xi' = 0$$

which has the general solution in R_z.

$$\xi'(t) = a\ (-\tfrac{1}{2})^t \qquad a \in R.$$

Therefore $n_0 = \dim X_0 = 1$.

The vectorspace Y_0^+ is given by

$$y'(t) = a\ (-\tfrac{1}{2})^t \qquad t \geq 1 \qquad a \in R$$

and $\dim Y_0^+ = 1$.

The general solutions to the given equations are

$$\xi(t) = a\ (-\tfrac{1}{2})^t + u\ (t-3)$$

$$y(t) = a\ (-\tfrac{1}{2})^t + u\ (t-1).$$

Therefore Y^+ is given by

$$y_+(t) = a\ (-\tfrac{1}{2})^t + u(t-1) \qquad t \geq 1, \quad u \in R_z, \quad a \in R$$

The two parameters a and $u(0)$ are needed to determine y_+. Therefore $\dim Y^+ = 2$ and $n_D = \dim Y^+ - \dim Y_0^+ = 1$. We find that the order of the system is $n = 1 + 1 = 2$.

□

We will now specify exactly how much of "the future behaviour of the system" that can be determined with $n = n_0 + n_D$ parameters.

Lemma 3.2.2 Let $T(q)$ be an $r \times r$ Laurent polynomial matrix with $\det T(q) \neq 0$ and let ξ be a solution in R_z^r to $T(q)\xi = 0$. If $\xi(t) = 0$ when $t \geq t_0$ for some $t_0 \in Z$, then $\xi = 0$.

Proof. Let $N_1(q)$ be a diagonal polynomial matrix with $\det N_1(q) = a q^p$, $a \in R \setminus \{0\}$, $p \in Z_{0+}$ such that $Q_1(q) = N_1(q)T(q)$ is a polynomial matrix. $Q_1(q)$ can be factored as $Q_1(q) = Q_2(q)Q(q)$, where $Q_2(q)$ and $Q(q)$ are polynomial matrices with $\det Q_2(q) = c\, q^k$, $c \in R \setminus \{0\}$, $k \in Z_{0+}$ and $\det Q(0) \neq 0$. The factorization can be done using the method of Rosenbrock (1970) p. 61-62. We have $Q(q) = N(q)T(q)$, where $N(q) = Q_2^{-1}(q)N_1(q)$ is a unimodular Laurent polynomial matrix. By theorem 2.2.3, the equations

$$T(q)\ \xi = 0 \tag{3.2.2}$$

and

$$Q(q)\ \xi = 0 \tag{3.2.3}$$

have exactly the same solutions in R_z^r.

Define $Q_0 = Q(0)$ and $\bar{Q}(q) = Q(q) - Q_0$. Then equation (3.2.3) can be written

$$\xi = -Q_0^{-1}\bar{Q}(q)\ \xi. \tag{3.2.4}$$

At time t_0-1 this gives

$$\xi(t_0-1) = -Q_0^{-1}\ \bar{Q}(q)\ \xi(t_0-1) \tag{3.2.5}$$

where the right member contains only $\xi(t)$ for $t \geq t_0$. Therefore if $\xi(t) = 0$ for $t \geq t_0$ then $\xi(t) = 0$ for $t \geq t_0-1$. Iterating like this it follows that $\xi = 0$.

□

Lemma 3.2.3 Any solution to (3.2.1a) $(u \in \overleftarrow{R_z^\ell})$ can uniquely be decomposed as $\xi = \xi' + \xi''$, where ξ' is a solution to $T(q)\xi' = 0$ and ξ'' is a solution to 3.2.1a with $\xi''(t) = 0$ for $t \geq t_0$, where t_0 is some integer.

Proof. Let ξ be a given solution to (3.2.1a). Since $u \in \overleftarrow{R_z^\ell}$ it is clear that there is a t_0 such that $T(q)\xi(t) = 0$ for $t \geq t_0$. Define ξ' as the unique solution in R_Z^r to $T(q)\xi' = 0$ such that $\xi'(t) = \xi(t)$, $t \geq t_0$. The uniqueness follows by Lemma 3.2.2. Put $\xi'' = \xi - \xi'$, then clearly $\xi''(t) = 0$, $t \geq t_0$.

□

Lemma 3.2.4 Let $y_+ \in \mathcal{Y}^+$. Then y_+ can be uniquely decomposed as $y_+ = y_+' + y_+''$, where $y_+' \in \mathcal{Y}_0^+$ and $y_+''(t) = 0$, $t \geq t_0$ for some $t_0 \in Z_+$.

Proof. Let $\xi \in R_Z^r$ be a solution to (3.2.1a) giving y. Decompose ξ according to lemma 3.2.3 as $\xi = \xi' + \xi''$ and define $y' = V(q)\xi'$ and $y'' = V(q)\xi'' + W(q)u$. Then $y = y' + y''$ and $y_+ = y_+' + y_+''$ is a decomposition of y_+ in the desired form.

Suppose there are two decompositions $y_+ = y_+' + y_+''$ and $y_+ = \bar{y}_+' + \bar{y}_+''$ of the desired type. Then

$$y_+' + y_+'' = \bar{y}_+' + \bar{y}_+''.$$

Define

$$\tilde{y}_+ \triangleq y_+' - \bar{y}_+' = \bar{y}_+'' - y_+''$$

where $\tilde{y}_+ \in \mathcal{Y}_0^+$ and $\tilde{y}_+(t) = 0$, $t \geq t_0$ for some $t_0 \in Z_{0+}$. We will show that $\tilde{y}_+ = 0$. Since $\tilde{y}_+ \in \mathcal{Y}_0^+$ there is a $\tilde{\xi}$ such that

$$T(q)\tilde{\xi} = 0 \qquad (3.2.6a)$$

$$\tilde{y} = V(q)\tilde{\xi} \qquad (3.2.6b)$$

where $\tilde{y}(t) = \tilde{y}_+(t)$ for $t \in Z_+$.

Let $X(q)$ be an $r \times r$ Laurent polynomial matrix such that

$$T(q) = T_0(q)\ X(q) \qquad (3.2.7a)$$

$$V(q) = V_0(q)\ X(q) \qquad (3.2.7b)$$

and $T_0(q)$ and $V_0(q)$ are relatively right prime.

By theorem 2.1.4 we have

$$Q(q)\, T_0(q) + P(q)\, V_0(q) = I. \tag{3.2.8}$$

Multiply (3.2.6b) from the left by $P(q)$ and use (3.2.7), (3.2.8), and (3.2.6a)

$$P(q)\tilde{y} = P(q)V_0(q)X(q)\tilde{\xi} = [I - Q(q)T_0(q)]X(q)\tilde{\xi} =$$

$$= X(q)\tilde{\xi} \triangleq z$$

Since $\tilde{y}(t) = 0$ for sufficiently large t, the same is true for $z(t)$. But z is by (3.2.6a) and (3.2.7a) a solution to

$$T_0(q)z = 0$$

By lemma 3.2.2 it follows that $z = 0$. This gives by (3.2.6b) and (3.2.7b) that $\tilde{y} = 0$ and therefore $\tilde{y}_+ = 0$. This means that the decomposition of y_+ is unique.

□

<u>Def. 3.2.6</u> Define the subset $\overline{\mathcal{Y}^{+}}$ of $\mathcal{Y}^+$ through $\overline{\mathcal{Y}^{+}} = \{y_+ \in \mathcal{Y}^+ \mid y_+(t) = 0,\ t \geq t_0 \text{ for some } t_0 \in Z_{0+}\}$.

<u>Remark 1</u>. Notice that $u \in \overline{R_z^{\ell}}$ in the definition of $\mathcal{Y}^+$.

<u>Remark 2</u>. It is easily shown that $\overline{\mathcal{Y}^{+}}$ is a vector space.

The following corollaries are direct consequences of lemma 3.2.4.

<u>Corollary 1</u>. The vector space $\mathcal{Y}^+$ can be decomposed as $\mathcal{Y}^+ = \mathcal{Y}_0^+ \oplus \overline{\mathcal{Y}^{+}}$, where $\oplus$ denotes the direct sum.

<u>Corollary 2</u>. $\dim \overline{\mathcal{Y}^{+}} = \dim \mathcal{Y}^+ - \dim \mathcal{Y}_0^+ = \dim(\mathcal{Y}^+/\mathcal{Y}_0^+) = n_D$.

Theorem 3.2.1 Let ξ be an arbitrary solution in R_z^r to (3.2.1a) with $u \in \overline{R_z^{\ell}}$ and let y be the corresponding output. To determine y_+ and ξ', defined through lemma 3.2.3, it is necessary and sufficient with $n = n_0 + n_D$ parameters.

Proof. Sufficiency. With n_0 parameters it is possible to determine ξ' since it belongs to the n_0-dimensional space X_0. y_+ can be decomposed as $y_+ = [V(q)\xi']_+ + [V(q)\xi'' + W(q)u]_+$, where ξ'' is given by lemma 3.2.3. The first term is known when ξ' is known. The second term belongs to $\overline{Y^+}$. Therefore it can be determined by $\dim \overline{Y^+} = n_D$ parameters.

Necessity. By lemma 3.2.4 y_+ can be decomposed as $y_+ = y_+' + y_+''$, where $y_+' \in Y_0^+$ and $y_+'' \in \overline{Y^+}$. It is sufficient to show that for any y_+'' in $\overline{Y^+}$ and ξ' in X_0 there is a solution ξ to (3.2.1) giving this y_+'' and ξ'. By definition of $\overline{Y^+}$ there is a ξ giving an arbitrary y_+'' in $\overline{Y^+}$. To this ξ can be added an element in X_0 so that the desired ξ' is obtained. This gives a contribution to y_+ in Y_0^+ but does not affect y_+'' since $\overline{Y^+}$ and Y_0^+ are independent subspaces.

□

Let ξ' and y_+'' be defined through lemma 3.2.3 and 3.2.4 respectively. We may define the state of the system at $t = 1$ as the n parameters needed to determine ξ' and y_+''. In that case we have to specify how ξ' and y_+'' can be determined from the state. Alternatively we can regard the pair (ξ', y_+'') as the state. We will do the latter thing here.

Def. 3.2.7 The state at time $t = 1$ of the system (3.2.1) is the pair (ξ', y_+''), where ξ' and y_+'' are defined through lemmas 3.2.3 and 3.2.4 respectively. The state space of the system is $X_0 \text{x} \overline{Y^+}$, where x denotes the set product.

Remark 1. If the state at $t = 1$ and u_+ are known then it is possible to determine y_+ and $\xi_+(t)$, $t \geq t_0$ for some $t_0 \in Z_{0+}$. In general it is not possible to determine ξ_+.

Remark 2. The state of any other time $t \in Z$ is defined analogously.

Example 3.2.2 Consider the same system as in example 3.2.1, namely

$$(2\, q^3 + q^2)\ \xi = (2 + q^{-1})\ u$$

$$y = \xi + (q^{-1} - q^{-3})\ u$$

The general solution is given by

$$\xi(t) = a(-\tfrac{1}{2})^t + u(t-3)$$

$$y(t) = a(-\tfrac{1}{2})^t + u(t-1)$$

Suppose $u \in \overline{R_Z}$. Then it is needed four parameters e.g. (a, $u(-2)$, $u(-1)$, and $u(0)$) to determine ξ_+, while it is needed only two parameters (a and $u(0)$) to determine y_+.

The order of the system was in example 3.2.2 shown to be $n = 2$.

With the two parameters a and $u(0)$ we can uniquely determine y_+ and $\xi(t)$ for $t \geq 3$. It is however not possible to determine ξ_+.

Define

$$g(t) = \begin{cases} u(0) & t = 1 \\ 0 & t \geq 2 \end{cases}$$

The state at $t = 1$, as defined in def. 3.2.7, is

$$(\xi', y''_+) = (a(-\tfrac{1}{2})^t, g(t))$$

□

We will now show how the order of the system (3.2.1) can be computed from the matrices $T(q)$, $U(q)$, $V(q)$, and $W(q)$.

Theorem 3.2.2 The number $n_0 = \dim X_0$ in def. 3.2.5 is given by $n_0 = \underline{\deg} \det T(q)$, where $\underline{\deg} \det T(q)$ is the degree of the Laurent polynomial $\det T(q)$.

Proof. By theorem 2.2.3 equation (3.2.1a) can be multiplied from the left by a unimodular Laurent polynomial matrix without influencing the solutions. Since $R(x]$ is a euclidean domain there is a unimodular $N(q) \in R^{r\times r}(q]$ such that $T_1(q) = N(q)T(q)$ is lower left triangular

$$T_1(q) = \begin{pmatrix} t_{11}(q) & & 0 \\ \vdots & \ddots & \\ t_{r1}(q) & \cdots & t_{rr}(q) \end{pmatrix}$$

and

$$\underline{\deg} \det T(q) = \underline{\deg} \det T_1(q) = \sum_{i=1}^{r} \underline{\deg}\, t_{ii}(q).$$

We want to determine how many independent parameters that are needed to uniquely determine ξ in $T_1(q)\xi = 0$. The system of equations can be solved by solving the equations one by one from the top. The first equation is $t_{11}(q)\xi_1 = 0$. The equation can be written

$$t_{11}^{k_2}\, \xi_1(t+k_2) + t_{11}^{k_2-1}\, \xi_1(t+k_2-1) + \ldots + t_{11}^{k_1}\, \xi_1(t+k_1) = 0,$$

where $k_2 \geq k_1$ and $t_{11}^{k_1} \neq 0$, $t_{11}^{k_2} \neq 0$. By standard theory for difference equations it is needed $k_2 - k_1 = \underline{\deg}\, t_{11}(q)$ parameters to uniquely determine ξ_1 in R_z. The second equation is

$$t_{22}(q)\, \xi_2 = -t_{21}(q)\, \xi_1.$$

Here ξ_1 and therefore $t_{21}(q)\xi_1$ is already determined.

A particular solution to the equation can then be computed. To completely determine ξ_2 one has to determine the solution to $t_{22}(q)\xi_2 = 0$. For this there is needed $\underline{\deg}\ t_{22}(q)$ independent parameters. Continuing in this way we find that the total number of independent parameters needed to determine $\xi_1, \xi_2, \ldots, \xi_r$ is

$$\sum_{i=1}^{r} \underline{\deg}\ t_{ii}(q).$$

□

We will now give a method to compute $n_D = \dim \overline{Y^+}$. Define the transfer function for the system (3.2.1) as

$$G(q) \triangleq V(q)\ T^{-1}(q)\ U(q) + W(q) \tag{3.2.9}$$

Decompose $G(q)$ as

$$G(q) = H(q) + D(q) \tag{3.2.10}$$

where $D(q)$ is a strictly proper rational matrix with all poles of its entries zero and no entries of $H(q)$ have poles that are zero. This decomposition of $G(q)$ is unique.

<u>Def. 3.2.8</u> The subspace $\overline{R_z^k}$ of R_z^k is defined as

$$\overline{R_z^k} = \left\{ v \in R_z^k \mid \exists\ t_0 \in Z \text{ such that } t > t_0 \Rightarrow v(t) = 0 \right\}$$

<u>Lemma 3.2.5</u> Let $T(q)$ be an $r \times r$ Laurent polynomial matrix with $\det T(q) \neq 0$. Then $T(q)$ is a bijective linear map from $\overline{R_z^r}$ to $\overline{R_z^r}$.

<u>Proof</u>. That $T(q)$ is a linear map from $\overline{R_z^r}$ to $\overline{R_z^r}$ is clear. It is injective by lemma 3.2.2. Let u be an arbitrary element in $\overline{R_z^r}$. We will show that there is a $v \in \overline{R_z^r}$ such that

$$T(q)\ v = u \tag{3.2.11}$$

As in the proof of lemma 3.2.2 there is a unimodular Laurent polynomial matrix $N(q)$ such that $Q(q) = N(q)T(q)$ is polynomial with $\det Q(0) \neq 0$. Therefore (3.2.11) is equivalent to

$$Q(q)\ v = N(q)\ u \tag{3.2.12}$$

With $Q_0 = Q(0)$ and $\overline{Q}(q) = Q(q) - Q_0$ we have

$$v(t) = -Q_0^{-1}[\overline{Q}(q)\ v(t) - N(q)\ u(t)] \qquad t \in Z \tag{3.2.13}$$

Here $Q_0^{-1}\overline{Q}(q)$ is polynomial with no constant term. Choose t_0 such that $N(q)u(t) = 0,\ t > t_0$ and define $v(t) = 0$ for $t > t_0$. Iterate (3.2.13) backwards in time starting at $t = t_0$. This defines v as a solution to (3.2.13) and therefore also to (3.2.11). Therefore $T(q)$ is subjective.

□

<u>Corollary</u>. Any $m\times\ell$ rational matrix $G(q)$ is a well defined linear map from $\overline{R_z^{\ell}}'$ to $\overline{R_z^{m}}'$.

<u>Proof</u>. $G(q)$ can be factored as $T^{-1}(q)U(q)$, where $T(q)$ is an $m\times m$ polynomial matrix with $\det T(q) \neq 0$ and $U(q)$ is an $m\times\ell$ polynomial matrix. $U(q)$ is clearly a well defined linear map from $\overline{R_z^{\ell}}'$ to $\overline{R_z^{m}}'$ and $T^{-1}(q)$ is by lemma 3.2.5 a well defined linear map from $\overline{R_z^{m}}'$ to $\overline{R_z^{m}}'$

□

<u>Lemma 3.2.6</u> Let $D(q)$ be defined through (3.2.10). Then $\overline{y^{+}}'$ is the range of the mapping $f(u) = [D(q)u]_+$, $u \in \overline{R_z^{\ell}}'$, where $D(q)$ is interpreted as a map from $\overline{R_z^{\ell}}'$ to $\overline{R_z^{m}}'$.

<u>Proof</u>. By the proof of lemma 3.2.4 any element y_+'' in $\overline{y^{+}}'$ can be written $y_+'' = [V(q)\xi'' + W(q)u]_+$ for some $u \in \overline{R_z^{\ell}}'$, where ξ'' is a solution to (3.2.1a) with $\xi'' \in \overline{R_z^{r}}'$. I.e.

$$T(q)\xi'' = U(q)u, \quad u \in \overline{R_z^{\ell}}', \quad \xi'' \in \overline{R_z^{r}}' \tag{3.2.14}$$

By lemma 3.2.5 it follows that ξ'' is uniquely determined by u and

$$\xi'' = T^{-1}(q)U(q)u \tag{3.2.15}$$

where $T^{-1}(q)U(q)$ is interpreted as a linear map from $\overline{R_z^{\ell}}$ to $\overline{R_z^{r}}$. This gives

$$y''_+ = [V(q)T^{-1}(q)U(q)u + W(q)u]_+ \tag{3.2.16}$$

or by (3.2.9) and (3.2.10)

$$y''_+ = [H(q)u]_+ + [D(q)u]_+ \tag{3.2.17}$$

It is shown e.g. in Rosenbrock (1970) that any $m\times\ell$ rational matrix can be written

$$H(q) = A^{-1}(q)B(q) \tag{3.2.18}$$

where $A(q)$ and $B(q)$ are $m\times m$ and $m\times\ell$ polynomial matrices respectively with $\det A(q) \neq 0$. Furthermore $A(q)$ and $B(q)$ are relatively left prime and every zero of $\det A(q)$ is a pole of an entry in $H(q)$ and vice versa. Since $H(q)$ by definition has no poles that are zero we have

$$\det A(0) \neq 0 \tag{3.2.19}$$

To evaluate (3.2.17) we have to find $z \in \overline{R_z^{m}}$ such that

$$z = H(q)u \tag{3.2.20}$$

This is equivalent to solving the equation

$$A(q)z = B(q)u, \qquad u \in \overline{R_z^{\ell}}, \qquad z \in \overline{R_z^{m}} \tag{3.2.21}$$

Introduce $A_0 = A(0)$ and $A_1(q) = A_0 - A(q)$. Then (3.2.21) gives

$$A_0 z - A_1(q)z = B(q)u \tag{3.2.22}$$

By (3.2.19) $\det A_0 \neq 0$. Therefore

$$z = A_0^{-1} A_1(q)\ z + A_0^{-1} B(q)\ u \tag{3.2.23}$$

Here $A_0^{-1}A_1(q)$ and $A_0^{-1}B(q)$ are polynomial matrices and

$A_0^{-1}A_1(q)$ has no constant term. Since $u \in \overline{R_z^\ell}$ we have that $A_0^{-1}B(q)u(t) = 0,\ t \geq 1$. Iterating (3.2.23) backwards in time, starting with $z(t) = 0,\ t \geq t_0$ for some $t_0 \in Z_+$ gives that $z(t) = 0,\ t \geq 1$. I.e.

$$[H(q)u]_+ = 0, \qquad u \in \overline{R_z^\ell} \tag{3.2.24}$$

It now follows from (3.2.17) that

$$y_+'' = [D(q)u]_+ \qquad u \in \overline{R_z^\ell} \tag{3.2.25}$$

We have shown that any y_+'' in $\overline{\mathcal{Y}^+}$ is given by (3.2.25) for some $u \in \overline{R_z^\ell}$.

Conversely any $u \in \overline{R_z^\ell}$ will by (3.2.15) give a unique ξ'' and by (3.2.16) a y_+'' in $\overline{\mathcal{Y}^+}$. But (3.2.16) is equivalent to (3.2.25). □

The following lemma is obvious.

<u>Lemma 3.2.7</u> Let $N(q)$ be a unimodular polynomial matrix. Then $N(q)$ is a bijective linear map from $\overline{R_z^\ell}$ to $\overline{R_z^\ell}$.

<u>Def. 3.2.9</u> For any rational matrix G define $\nu(G)$ as the sum of the degrees of the denominator polynomials in the McMillan form of G.

<u>Theorem 3.2.3</u> Let $D(q)$ be given by (3.2.10). Then $n_D = \nu(D)$.

<u>Proof</u>. Let $\mathcal{R}(f)$ denote the range space of the linear mapping f. We have to show that $\dim \mathcal{R}(f) = \nu(D)$, where $f : \overline{R_z^\ell} \to \mathcal{Y}^+$ is given by $f(u) = (D(q)u)_+,\ u \in \overline{R_z^\ell}$.

(i) Let $N(q)$ be a unimodular <u>polynomial</u> matrix. Then it follows by lemma 3.2.7 that if the mapping $f_1 : \overline{R_z^\ell} \to \mathcal{Y}^+$ is given by $f_1(u) = (D(q)N(q)u)_+$ then $\mathcal{R}(f_1) = \mathcal{R}(f)$.

(ii) Let $M(q)$ be a unimodular polynomial matrix. Define the linear map $f_2 : \mathcal{Y}^+ \to \mathcal{Y}^+$ through $f_2(y_1) = (M(q)y_1)_+$. This is clearly well defined. Define $f_3 : \overline{R_z^{\ell}} \to \mathcal{Y}^+$ as the composition map $f_3 = f_2 o f$, i.e. $f_3(u) = [M(q)(D(q)u)_+]_+ =$ $= [M(q)D(q)u]_+$. The last equality is true because $M(q)$ is a polynomial matrix. It is clear that

$$\dim \mathcal{R}(f_3) \le \dim \mathcal{R}(f). \tag{3.2.26}$$

Define $f_4 : \mathcal{Y}^+ \to \mathcal{Y}^+$ through $f_4(y_1) = (M^{-1}(q)y_1)_+$ and $f_5 : \overline{R_z^{\ell}} \to \mathcal{Y}^+$ as $f_5 = f_4 o f_3$. Then

$$\dim \mathcal{R}(f_5) \le \dim \mathcal{R}(f_3). \tag{3.2.27}$$

But f_5 is given by $f_5(u) = [M^{-1}(q)(M(q)D(q)u)_+]_+ =$ $= [M^{-1}(q)M(q)D(q)u]_+ = [D(q)u]_+$. I.e.

$$f_5 = f \tag{3.2.28}$$

It follows from (3.2.26)-(3.2.28) that

$$\dim \mathcal{R}(f_3) = \dim \mathcal{R}(f) \tag{3.2.29}$$

(iii) Let $f_6 : \overline{R_z^{\ell}} \to \mathcal{Y}^+$ be defined through $y = (M(q)D(q)N(q)u)_+$, where $M(q)$ and $N(q)$ are unimodular polynomial matrices. Then it follows by (i) and (ii) that

$$\dim \mathcal{R}(f_6) = \dim \mathcal{R}(f) \tag{3.2.30}$$

(iv) Choose $M(q)$ and $N(q)$ such that

$$\overline{D}_M(q) = M(q)D(q)N(q) \tag{3.2.31}$$

is the McMillan form of $D(q)$. Let $D_M(q)$ be the strictly proper part of $\overline{D}_M(q)$. Since $\nu(D_m) = \nu(\overline{D}_M) = \nu(D)$ and $[D_M(q)u]_+ = [\overline{D}_M(q)u]_+$, $u \in \overline{R_z^{\ell}}$, it is sufficient to show the theorem for $D_M(q)$.

(v) $D_M(q)$ has the form

$$D_M(q) = \begin{pmatrix} \frac{\varepsilon_1(q)}{q^{k_1}} & & & 0 \\ & \ddots & & \\ & & \frac{\varepsilon_p(q)}{q^{k_p}} & \\ 0 & & & 0 \end{pmatrix} \quad (3.2.32)$$

where $\varepsilon_i(q)$ are polynomials and $\varepsilon_i(0) \neq 0 \ \forall \ i$. Let $f_M : \overline{R_z^{\ell}} \to \mathcal{Y}^+$ be given by

$$y = (D_M(q)u)_+ \quad (3.2.33)$$

and let $u_{ij} \in \overline{R_z^{\ell}}$ be an input, which is zero except for the i:th component at time j. This component is equal to 1. Clearly $\{u_{ij}\}_{i \in [1,\ell], j \in Z_{0-}}$ span $\overline{R_z^{\ell}}$. Let y_{ij} be defined through $y_{ij} = (D_M(q)u_{ij})_+$. Then $\{y_{ij}\}$ span $\mathcal{R}(f_M)$. Observe that the only nonzero component of y_{ij} is the i:th because $D_M(q)$ is diagonal. Now $j \leq -k_i$ gives $y_{ij} = 0$. Therefore $\{y_{ij}\}$, $j = -k_i + 1, \ldots, 0$ and $i = 1, \ldots, p$ span $\mathcal{R}(f_M)$. Furthermore these y_{ij} are linearly independent since $y_{ij}(j + k_i) = \varepsilon_i(0)u_{ij}(j)$, where $\varepsilon_i(0) \neq 0$, and $y_{ij}(j + k) = 0$, $k > k_i$. Therefore

$$\dim \mathcal{R}(f_M) = \sum_{i=1}^{p} k_i = \nu(D_M) \quad (3.2.34)$$

□

We quote the following result from Rosenbrock (1970) (algorithm 5.1).

<u>Lemma 3.2.8</u> Let φ be the least common denominator of all minors of all orders of the rational matrix G. Then $\nu(G)$ is equal to the degree of the polynomial φ.

Let $D(q)$ be defined through (3.2.10) and define $D^*(x)$ as

$$D^*(x) = D(x^{-1}). \quad (3.2.35)$$

Then $D^*(x)$ is a polynomial matrix because $D(x)$ is proper and its entries have only zero poles.

Theorem 3.2.4 Compute all minors of all orders of D^*. The degree of the minor of highest degree is $\nu(D)$.

Proof. The degree of a minor of $D^*(x)$ is equal to the degree of the denominator of the corresponding minor of $D(x)$. Since the minors of $D(x)$ have denominators of the form x^k then the least common denominator is equal to the denominator of highest degree.

□

Corollary. Define G^* through $G^*(x) = G(x^{-1})$ and make the decomposition $G^*(x) = F^*(x) + E^*(x)$, where $F^*(x)$ is strictly proper and $E^*(x)$ is polynomial. Then n_D of def. 3.2.5 is equal to the degree of the minor of E^* that has highest degree.

Proof. The difference between $E^*(x)$ and $D^*(x)$ is independent of x. Therefore the degree of the minor of highest degree is the same. The result then follows from theorems 3.2.3 and 3.2.4.

Example 3.2.3 Consider the same system as in example 3.2.1 and 3.2.2

$$(2\,q^3 + q^2)\;\xi = (2 + q^{-1})\;u$$

$$y = \xi + (q^{-1} - q^{-3})\;u$$

Theorem 3.2.2 gives $n_0 = \underline{\deg}\ \det\ (2q^3 + q^2) = \underline{\deg}\ (2q^3 + q^2) = 1$. The transfer function is

$$G(q) = \frac{2 + q^{-1}}{2q^3 + q^2} + q^{-1} - q^{-3} = q^{-1} = \frac{1}{q}$$

$D(q)$, defined through (3.2.2), is given by

$$D(q) = \frac{1}{q}$$

and $n_D = \nu(D) = 1$. The order of the system is $n = 1 + 1 = 2$.

□

Example 3.2.4 Determine the order of the system

$$\begin{pmatrix} q^2-1 & q-1+q^{-1}-q^{-2} \\ q^2 & 1+q^{-1} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} q-1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1-q^{-1} & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}.$$

$$n_0 = \underline{\deg} \det \begin{pmatrix} q^2-1 & q-1+q^{-1}-q^{-2} \\ q^2 & 1+q^{-1} \end{pmatrix} = \underline{\deg} \left(-q^3+2q^2-q^{-1}\right) = 4$$

$$G(q) = \begin{pmatrix} 1 & 0 \\ 1-q^{-1} & 1 \end{pmatrix} \begin{pmatrix} q^2-1 & q-1+q^{-1}-q^{-2} \\ q^2 & 1+q^{-1} \end{pmatrix}^{-1} \begin{pmatrix} q-1 & 0 \\ 0 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} \dfrac{q+1}{-q^3+q^2+q+1} & \dfrac{-q^2-1}{-q^4+q^3+q^2+q} \\ \dfrac{-q^4+q^2-1}{-q^4+q^3+q^2+q} & \dfrac{q^4+q^2-q+1}{-q^5+q^4+q^3+q^2} \end{pmatrix} =$$

$$= \frac{1}{-q^3+q^2+q+1} \begin{pmatrix} q+1 & -q^2+1 \\ -q^3-q^2+2q+1 & -q^2+3q+2 \end{pmatrix} +$$

$$+ \begin{pmatrix} 0 & -\dfrac{1}{q} \\ -\dfrac{1}{q} & \dfrac{-2q+1}{q^2} \end{pmatrix}$$

By (2.2.2) $D(q)$ is defined as

$$D(q) = \begin{pmatrix} 0 & -\frac{1}{q} \\ -\frac{1}{q} & \frac{-2q+1}{q^2} \end{pmatrix}$$

Let $D_M(q)$ be the McMillan form of $D(q)$

$$D_M(q) = \begin{pmatrix} \frac{1}{q^2} & 0 \\ 0 & 1 \end{pmatrix}$$

and $n_D = \nu(D) = 2$ by theorem 3.2.3. We find that $n = n_0 + n_D = 4 + 2 = 6$.

Alternatively n_D can be computed using theorem 3.2.4.

$$D^*(x) = \begin{pmatrix} 0 & -x \\ -x & -2x + x^2 \end{pmatrix}$$

The highest degree of the 1×1 minors as $\deg(-2x+x^2) = 2$ and the degree of the 2×2 minor is $\deg(-x^2) = 2$. It follows that $n_D = 2$.

Another possibility is to use the corollary of theorem 3.2.4 to compute n_D.

$$G^*(x) = \begin{pmatrix} \frac{x^3 + x^2}{x^3 + x^2 + x - 1} & \frac{-x^4 - x^2}{x^3 + x^2 + x - 1} \\ \frac{-x^4 + x^2 - 1}{x^3 + x^2 + x - 1} & \frac{x^5 - x^4 + x^3 + x}{x^3 + x^2 + x - 1} \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{-x + 1}{x^3 + x^2 + x - 1} & \frac{-x^2 - 2x + 1}{x^3 + x^2 + x - 1} \\ \frac{x^2 - 2x}{x^3 + x^2 + x - 1} & \frac{x^2 - 3x + 2}{x^3 + x^2 + x - 1} \end{pmatrix} + \begin{pmatrix} 1 & -x + 1 \\ -x + 1 & x^2 - 2x + 2 \end{pmatrix}$$

and

$$E^*(x) = \begin{pmatrix} 1 & -x+1 \\ -x+1 & x^2-2x+2 \end{pmatrix}.$$

The highest occuring degree of any 1×1 or 2×2 minor is 2 and therefore $n_D = 2$.

□

3.3 Stability

We will use Lyapunov's definition of asymptotic stability. This can for the system

$$T(q)\xi = U(q)u \tag{3.3.1a}$$

$$y = V(q)\xi + W(q)u \tag{3.3.1b}$$

be formulated in the following way.

Def. 3.3.1 The system (3.3.1) is asymptotically stable if all solutions ξ to (3.3.1a) with $u \in \overbrace{R_z^\ell}$ satisfy $\lim_{t\to+\infty} \xi(t) = 0$.

Remark. $\lim_{t\to+\infty} \xi(t) = 0 \Rightarrow \lim_{t\to+\infty} y(t) = 0$ since $u \in \overbrace{R_z^\ell}$.

Theorem 3.3.1 The system (3.3.1) is asymptotically stable if and only if the zeros of the Laurent polynomial $\det T(q)$ all have a magnitude less than 1.

Proof. The right hand side of (3.3.1a) will be zero after some finite time T. Therefore it is sufficient to study the equation $T(q)\xi = 0$. As in the proof of theorem 3.2.2 there is a unimodular $N(q) \in R^{r\times r}(q]$ such that $T_1(q) = N(q)T(q)$ is lower left triangular.

$$T_1(q) = \begin{pmatrix} t_{11}(q) & & 0 \\ \vdots & \ddots & \\ t_{r1}(q) & \cdots & t_{rr}(q) \end{pmatrix}$$

Furthermore the equations $T_1(q)\xi = 0$ and $T(q)\xi = 0$ have the same solutions. Let λ_{ij}, $j = 1,2,\ldots,n_i$, where n_i is the degree of the Laurent polynomial t_{ii}, be the zeros of $t_{ii}(q)$. The solution to the first equation in $T_1(q)\xi = 0$ is

$$\xi_1(t) = \sum_{j=1}^{n_1} P_{1j}(t)\lambda_{1j}^t$$

where $P_{ij}(t)$ is an arbitrary polynomial of degree $m_{ij}-1$ and m_{ij} is the multiplicity of λ_{ij}. Now $\xi_1(t) \to 0$ if and only if $|\lambda_{2j}| < 1$, $j = 1,\ldots,n_2$. The second equation is $t_{22}(q)\xi_2 = t_{21}(q)\xi_1$. If $|\lambda_{1j}| < 1$, $j = 1,\ldots,n_1$, then $t_{21}(q)\xi_1$ gives a particular solution that tends to zero. Therefore $\xi_2(t) \to 0$ if and only if $|\lambda_{2j}| < 1$, $j = 1,\ldots,n_2$. Analogously it follows that the total solution $\xi(t) \to 0$ if and only if $|\lambda_{ij}| < 1$, $j = 1,\ldots,n_i$, $i = 1,\ldots,r$.

□

3.4 Causality

Consider system (3.3.1) and let $u \in \overline{R_z^\ell}$ be such that $u(t) = 0$, $t \leq 0$. The solution ξ to (3.3.1a) can as usual be decomposed into $\xi = \xi_1 + \xi_2$, where ξ_1 solves $T(q)\xi_1 = 0$ and $\xi_2 = T^{-1}(q)U(q)u$. Clearly ξ_1 does not depend on u. Now ξ_2 depends causally on u if and only if $\xi_2(t) = 0$, $t \leq 0$ for any u. If ξ represents physical variables then ξ_2 has to depend causally on u. However, to gain flexibility we will allow ξ to depend noncausally on u. We must then regard ξ as a variable in the mathematical model, a variable that is related to the physical variables, but not in a causal way.

The output y can be decomposed as $y = y_1 + y_2$, where $y_1 = V(q)\xi_1$ and $y_2 = V(q)\xi_2 + W(q)u$. In analogy with the previous case we allow a noncausal relationship between ξ and y. However, u and y are both physical variables, and we must demand that y depends causally on u. Since ξ_1 is independent on u we have no demand on y_1. Introducing the expression for ξ_2 into the expression for y_2 we get $y_2 = [V(q)T^{-1}(q)U(q) + W(q)]u = G(q)u$. The only causal relationship we have to demand is that between u and y_2. Decompose $G(q)$ as

$$G(q) = G_0(q) + G_1(q) \tag{3.4.1}$$

where $G_0(q)$ is strictly proper and $G_1(q)$ is polynomial. This decomposition is unique. Now it is clear that y_2 depends causally on u if and only if $G_1(q)$ is independent of q. We can therefore make the following definition.

<u>Def. 3.4.1</u> The system (3.3.1) with transfer function $G(q)$ is causal if $G_1(q)$, defined through (3.4.1), does not depend on q.

<u>Example 3.4.1</u> The system

$$\xi = qu$$
$$y = q^{-1}\xi$$

is causal, while the system

$$\xi = qu$$
$$y = \xi$$

is not causal.

□

3.5 Controllability

Consider the system

$$T(q)\xi = U(q)u \qquad (3.5.1a)$$

$$y = V(q)\xi + W(q)u \qquad (3.5.1b)$$

The state at time $t = 1$ of the system is defined in Def. 3.2.7. Clearly the state is uniquely given by the solution ξ to (3.5.1).

Def. 3.5.1 A state at $t = 1$ is called controllable if it is given by a solution ξ in $\overline{R}_z^r$ for some u in $\overline{R}_z^\ell$. The system is controllable if all states at $t = 1$ are controllable.

Remark. Observe that ξ must belong to $\overline{R}_z^r$ for the definition to be meaningful. The intuitive meaning of Def. 3.5.1 is the following. Suppose that the system initially is at rest i.e. $\xi(t)$ and $y(t)$ are both zero for large negative t. The controllable states are those that can be obtained using a finite input sequence.

The set of controllable states is a subspace of the state space called the controllable subspace. We will now give an algebraic condition for controllability. To do this we need the following two lemmas.

Lemma 3.5.1 Let $a(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0$ and $b(x) = b_{n-1}x^{n-1} + \ldots + b_0$ be two relatively prime polynomials and consider the system

$$a(q)\xi = b(q)u \qquad (3.5.2)$$

To any ξ' given by $a(q)\xi' = 0$ there is a $u \in \overline{R}_z$ such that the unique solution ξ in $\overline{R}_z$ to (3.5.2) has the property $\xi(t) = \xi'(t)$, $t \geq t_0$ for some finite t_0.

<u>Proof</u>. Explicitely equation (3.5.2) can be written

$$\xi(t+n) + a_{n-1}\xi(t+n-1) + \ldots + a_0\xi(t) =$$

$$= b_{n-1}u(t+n-1) + \ldots + b_0u(t) \tag{3.5.3}$$

Suppose $u(t) = 0$ for $t \leq -n$ and $t > 0$. The solution ξ in R_z can be found by iterating equation (3.5.3). We find that $\xi(t) = 0$ for $t \leq -n+1$. In the interval $-n+2 \leq$ $\leq t \leq n$ the solution $\xi(t)$ is given by

$$\left(\begin{array}{cccc|cccc}
1 & a_{n-1} & & & a_0 & 0 & & \\
0 & 1 & \ddots & & a_1 & a_0 & 0 & \\
 & \ddots & \ddots & & & & \ddots & 0 \\
 & & \ddots & \ddots & & & & a_0 \\
 & & 0 & 1 & a_{n-1} & \cdots & \cdots & a_1 \\
\hline
 & & & 0 & 1 & a_{n-1} & \cdots & a_2 \\
 & & & & \ddots & \ddots & \ddots & \vdots \\
 & & & & & \ddots & \ddots & \\
 & & & & & & 0 & 1
\end{array}\right)
\left(\begin{array}{c}
\xi(n) \\ \vdots \\ \vdots \\ \xi(1) \\ \hline \xi(0) \\ \vdots \\ \vdots \\ \xi(-n+2)
\end{array}\right) =$$

$$= \left(\begin{array}{cccc}
b_0 & 0 & & \\
b_1 & b_0 & 0 & \\
\vdots & & \ddots & 0 \\
\vdots & & \ddots & \\
b_{n-1} & \cdots & \cdots & b_0 \\
\hline
0 & b_{n-1} & \cdots & b_1 \\
 & \ddots & \ddots & \vdots \\
 & & \ddots & \vdots \\
 & & 0 & b_{n-1}
\end{array}\right)
\left(\begin{array}{c}
u(0) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ u(-n+1)
\end{array}\right) \tag{3.5.4}$$

For $t > n$, $\xi(t)$ is given by

$$\xi(t) = -a_{n-1}\xi(t-1) - \ldots - a_0\xi(t-n) \tag{3.5.5}$$

i.e. ξ is equal to a solution ξ' to $a(q)\xi' = 0$ for $t > n$. It has to be shown that it can be made equal to <u>any</u> solution ξ' to $a(q)\xi' = 0$ for $t > n$ by a suitable choice of u. This is true if the vector $(\xi(n) \ldots \xi(1))^T$ can be given an arbitrary value by chosing u. Equation (3.5.4) can be written

$$\left(\begin{array}{cccc|cccc} a_0 & 0 & & & b_0 & 0 & & \\ a_1 & a_0 & 0 & & b_1 & b_0 & 0 & \\ \vdots & & & 0 & \vdots & & \ddots & \\ \vdots & & & a_0 & \vdots & & & 0 \\ a_{n-1} & \cdots & \cdots & a_1 & b_{n-1} & \cdots & \cdots & b_0 \\ \hline 1 & a_{n-1} & & a_2 & 0 & b_{n-1} & & b_1 \\ & \ddots & & & & \ddots & \ddots & \vdots \\ & & \ddots & & & & \ddots & \vdots \\ & & & 1 & & & 0 & b_{n-1} \end{array}\right) \left(\begin{array}{c} -\xi(0) \\ \vdots \\ -\xi(-n+2) \\ \hline u(0) \\ \vdots \\ u(-n+1) \end{array}\right) =$$

$$= \left(\begin{array}{cccc} 1 & a_{n-1} & \cdots & a_1 \\ 0 & 1 & & \\ & & \ddots & \\ & & & 1 \\ \hline 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{array}\right) \begin{pmatrix} \xi(n) \\ \vdots \\ \xi(1) \end{pmatrix} \tag{3.5.6}$$

The square matrix on the left hand side is invertible since the polynomials $a(x)$ and $b(x)$ are relatively prime (see e.g. Rosenbrock (1970)). Therefore $(u(0) \ldots u(-n+1))^T$ can be solved for any $(\xi(n) \ldots \xi(1))^T$.

□

The following lemma is formulated and proven for polynomial matrices in Rosenbrock, Hayton (1974).

Lemma 3.5.2 Let $T(x)$ and $U(x)$ be respectively $r \times r$ and $r \times \ell$ relatively left prime Laurent polynomial matrices with $\det T(x) \neq 0$. Let $T_1(x)$ and $U_1(x)$ have the same properties. Suppose that

$$T_1^{-1}(x)U_1(x) = T^{-1}(x)U(x) \tag{3.5.9}$$

Then there is a unimodular Laurent polynomial matrix $R(x)$ such that

$$T_1(x) = R(x)T(x) \tag{3.5.10}$$

$$U_1(x) = R(x)U(x) \tag{3.5.11}$$

Proof. The proof is analogous to the one of Theorem 2 in Rosenbrock, Hayton (1974).

□

Theorem 3.5.1 The system (3.5.1) is controllable if and only if the Laurent polynomial matrices $T(q)$ and $U(q)$ are relatively left prime.

Proof. Suppose that there is an $r \times r$ Laurent polynomial matrix $L(q)$ with $\underline{\deg} \det L(q) = n_L > 0$ such that $T(q) = L(q)T_1(q)$ and $U(q) = L(q)U_1(q)$. If $\underline{\deg} \det T(q) = n_0$ then $\underline{\deg} \det T_1(q) = n_0 - n_L < n_0$. The solutions ξ in $\overline{R_z^r}$ to (3.5.1a) and to

$$T_1(q)\xi = U_1(q)u \qquad (3.5.12)$$

are identical by Theorem 2.2.4 since $\det L(q) \not\equiv 0$. Any solution ξ to (3.5.12) with $u \in \overline{R_Z^{\ell}}$ is equal to a solution ξ' to $T_1(q)\xi' = 0$, when $t > t_0$ for some $t_0 \in Z$. The solutions to $T_1(q)\xi' = 0$ span an $n_0 - n_L$ dimensional space by Theorem 3.2.2. Since the solutions to $T(q)\xi' = 0$ span an n_0 dimensional space there is a solution ξ' to $T(q)\xi' = 0$ which is not a solution to $T_1(q)\xi' = 0$. This ξ' will not be equal to any solution ξ in $\overline{R_Z^r}$ to (3.5.12) and therefore neither to (3.5.1a) for large t. Thus, there exist states (ξ', y_+'') that are not controllable and the system is not controllable.

To prove the converse suppose that $T(q)$ and $U(q)$ are relatively left prime Laurent polynomial matrices. Let $\Psi^{-1}(q)\varepsilon(q)$ be the McMillan form of $T^{-1}(q)U(q)$. Here $\Psi(q)$ and $\varepsilon(q)$ are diagonal relatively left prime polynomial matrices. This means that there are unimodular polynomial matrices $N(q)$, $M(q)$ such that

$$[T(q)N(q)]^{-1}U(q)M(q) = \Psi^{-1}(q)\varepsilon(q) \qquad (3.5.13)$$

Since $T(q)N(q)$ and $U(q)M(q)$ are relatively left prime Laurent polynomial matrices it follows from Lemma 3.5.2 that there is a unimodular Laurent polynomial matrix $R(q)$ such that

$$R(q)T(q)N(q) = \Psi(q) \qquad (3.5.14a)$$

$$R(q)U(q)M(q) = \varepsilon(q) \qquad (3.5.14b)$$

Define $f_c : \overline{R_Z^{\ell}} \to X_0$ (Def. 3.2.2) in the following way. To any $u \in \overline{R_Z^{\ell}}$ there is a unique solution $\xi \in \overline{R_Z^r}$ to

$$T(q)\xi = U(q)u \qquad (3.5.1a)$$

This ξ uniquely determines a $\xi' \in X_0$ through $\xi'(t) = = \xi(t)$, $t > t_0$ for some $t_0 \in Z$.

We will show that f_c is surjective by showing that $\dim R(f_c) = \dim X_0$.

Multiplying (3.5.1a) from the left by $R(q)$, defined through (3.5.14) doesn't change the solutions by Theorem 2.2.3. By Lemma 3.2.7 u can, in (3.5.1a), be substituted by $M(q)u$, where $M(q)$ is defined in (3.5.13), without changing $R(f_c)$. Furthermore ξ in (3.5.1a) can be substituted by $N(q)\xi$ without changing $\dim R(f_c)$ because $N(q)$ is by Theorem 2.2.3 a bijective map on R_z^r. As a consequence equation (3.5.1a) can be substituted by

$$\Psi(q)\xi = \varepsilon(q)u \tag{3.5.15}$$

in the definition of f_c without changing $\dim R(f_c)$.

It follows by Lemma 3.5.1 that $\dim R(f_c) = \dim X_0^{\Psi}$, where X_0^{Ψ} is the space of solutions in R_z^r to $\Psi(q)\xi' = 0$. Theorem 3.2.2 now gives $\dim X_0^{\Psi} = \underline{\deg} \det \Psi(q) = \underline{\deg} \det T(q) = = \dim X_0$ and f_c is surjective.

The state space is $X_0 \times \overline{Y^{+}}$ and we have shown that any $\xi' \in X_0$ can be obtained by a suitable choice of $u \in \overline{R_z^{\ell}}$. This u will, however, also give a component of the state in $\overline{Y^{+}}$ unless $u(t) = 0$, $t > \tau$, where τ is some negative integer.

Take a fix $\xi' \in X_0$. Then $q^{\tau}\xi' \in X_0$ because $T(q)\xi' = 0 \Leftrightarrow \Leftrightarrow q^{\tau}T(q)\xi' = 0 \Leftrightarrow T(q)q^{\tau}\xi' = 0$. Let $u \in \overline{R_z^{\ell}}$ be such that $f_c(u) = q^{\tau}\xi'$. This implies that $f_c(q^{-\tau}u) = \xi'$ because f_c is time invariant. Clearly $q^{-\tau}u(t) = 0$ for $t > \tau$. Therefore any $\xi' \in X_0$ can be obtained with the component of the state in $\overline{Y^{+}}$ equal to zero. This means that X_0 is included in the controllable subspace.

By definition of $\overline{y^+}$ any component of $\overline{y^+}$ can be obtained from an $u \in \overline{R_z^\ell}$. In general this u also gives a component in X_0, but this component can be deleted by another superposed u because X_0 belongs to the controllable subspace. Therefore also $\overline{y^+}$ belongs to the controllable subspace and the system is controllable.

□

Remark. Controllability was defined for states at time t = 1 but this theorem shows that the property of controllability is indpendent of t.

Corollary. Suppose that T(q) and U(q) in (3.5.1) have a common left divisor L(q) so that

$$T(q) = L(q)T_1(q) \tag{3.5.16a}$$

$$U(q) = L(q)U_1(q) \tag{3.5.16b}$$

and $T_1(q)$ and $U_1(q)$ are relatively left prime. Let $X_0^{(c)}$ be the space of all $\xi' \in R_z^r$ such that

$$T_1(q)\xi' = 0 \tag{3.5.16c}$$

Then $X_0^{(c)} \times \overline{y^+}$ is the controllable subspace.

Proof. In the definition of controllable states (Def. 3.5.2) only solutions ξ in $\overline{R_z^r}$ are regarded. But (3.5.1a) and

$$T_1(q)\xi = U_1(q)u \tag{3.5.17}$$

have the same solutions in $\overline{R_z^r}$ by Theorem 2.2.4 because det $L(q) \neq 0$. The system (3.5.17), (3.5.1b) is by Theorem 3.5.1 controllable and have a controllable state space $X_0^{(c)} \times \overline{y^+}$.

□

From the corollary it follows that the zeros of $\det T_1(q)$ in (3.5.16) are those zeros of $\det T(q)$ that are associated with the controllable states. The zeros of $\det L(q)$ correspond to solutions in X_0 that cannot be excited by the input, they are decoupled from the input. Following Rosenbrock (1970) we will call them input decoupling zeros.

Let $S_L(q)$ be the Smith form of $L(q)$. Then there are unimodular Laurent polynomial matrices $M_1(q)$ and $N_1(q)$ such that $L(q) = M_1(q)S_L(q)N_1(q)$ and the zeros of the invariant factors of $S_L(q)$ are the input decoupling zeros. By Theorem 2.1.4 there are unimodular Laurent polynomial matrices $M_2(q)$ and $N_2(q)$ such that $\begin{pmatrix}T_1(q) & U_1(q)\end{pmatrix} = M_2(q)\begin{pmatrix}I & 0\end{pmatrix} N_2(q)$. Therefore

$$\begin{pmatrix}T(q) & U(q)\end{pmatrix} = M_1(q)S_L(q)N_1(q)M_2(q)\begin{pmatrix}I & 0\end{pmatrix}N_2(q) =$$

$$= M_1(q)\begin{pmatrix}S_L(q) & 0\end{pmatrix}\begin{pmatrix}N_1(q)M_2(q) & 0 \\ 0 & I\end{pmatrix}N_2(q)$$

By the uniqueness of the Smith form it follows that $\begin{pmatrix}S_L(q) & 0\end{pmatrix}$ is the Smith form of $\begin{pmatrix}T(q) & U(q)\end{pmatrix}$. We can thus formally define the input decoupling zeros in the following way.

Def. 3.5.2. The input decoupling (i.d.) zeros of a system are the zeros of the invariant factors of the Smith form of $\begin{pmatrix}T(q) & U(q)\end{pmatrix}$.

Example 3.5.1. Consider the system

$$(q-2)(q-1)\xi = (q-2)u \qquad (3.5.18a)$$

$$y = \xi \qquad (3.5.18b)$$

The system is of second order and the state space is spanned by the two states

$$(\xi', y_+'') = (1,0) \tag{3.5.19}$$

$$(\xi', y_+'') = (2^t, 0) \tag{3.5.20}$$

For this system $\overline{y^{+'}} = 0$, by Lemma 3.2.6, since the transfer function has no poles that are zero.

By the corollary of Theorem 3.5.1 the controllable subspace is spanned by the state (3.5.19). We will show this directly.

Choose $u(0) = a$ and $u(t) = 0$, $t \in Z \setminus \{0\}$. Equation (3.5.18a) can be written

$$\xi(t) = 3\xi(t-1) - 2\xi(t-2) + u(t-1) - 2u(t-2) \tag{3.5.21}$$

The solution in $\overline{R_z}$ is found by having $\xi(t) = 0$, $t \leq t_1$ for some $t_1 \in Z_-$. Iterating (3.5.21) then gives $\xi(t) = 0$ $t \leq 0$. Continuing the iteration gives $\xi(t) = a$, $t > 0$. It follows that the subspace spanned by the state $(\xi', y_+'') =$ $= (1,0)$ is included in the controllable subspace.

Introduce $z = (q-1)\xi$ into (3.5.18a) which gives

$$(q-2)z = (q-2)u \tag{3.5.22}$$

This can be written

$$z(t) = u(t) + 2[z(t-1) - u(t-1)] \tag{3.5.23}$$

Suppose $u \in \overline{R_z}$, i.e. $u(t) = 0$, $t < t_0$ for some t_0. The solution to (3.5.23) in $\overline{R_z}$ is found by iteration this equation starting with $z(t) = 0$ for $t < t_0$.

We find that $z(t_0) = u(t_0)$. Inserting $t = t_0 + 1$ in (3.5.23) then gives $z(t_0+1) = u(t_0+1)$ since the quantity inside the brackets is zero. Continuing like this we find

$$z(t) = u(t) \quad \forall t \in Z \tag{3.5.24}$$

The state of $t = 1$ is found by letting $u_+ = 0$. This gives by (3.5.24) $z(t) = 0$, $t > 0$ and therefore $\xi(t) =$ $=$ constant for $t > 0$. This means that it is not possible to reach any states that are not of the form $(\xi', y_+'') =$ $= (a,0)$, where $a \in R$. Therefore the controllable subspace is included in the space spanned by the state $(\xi', y_+'') =$ $= (1,0)$.

Since the reversed inclusion was shown earlier the controllable subspace is equal to the space spanned by the state (3.5.19).

□

3.6. Observalibity

The system

$$T(q)\xi = U(q)u \tag{3.6.1a}$$

$$y = V(q)\xi + W(q)u \tag{3.6.1b}$$

has a state space $X_0 \times \overline{Y^{+'}}$ according to Def. 3.2.7. Define the map f_0: $X_0 \times \overline{Y^{+'}} \to Y^+$ as

$$f_0(\xi', y_+'') = [V(q)\xi']_+ + y_+'' \tag{3.6.2}$$

f_0 is called the observability map and $y_+ = f_0(\xi', y_+'')$ is the output from the system (3.6.1) with (ξ', y_+'') as state at $t = 1$ and $u_+ = 0$ (see the first part of the proof of Theorem 3.2.1).

Def. 3.6.1. The system (3.6.1) is observable if f_0 is injective. If f_0 is not injective then the nullspace of f_0 is called the unobservable subspace.

Theorem 3.6.1. The system (3.6.1) is observable if and only if $T(q)$ and $V(q)$ are relatively right prime.

Proof. The system is observable if and only if the nullspace of f_0 is zero. The nullspace of f_0 is given by the states (ξ', y_+'') satisfying

$$[V(q)\xi']_+ + y_+'' = 0 \qquad (3.6.3)$$

Since $[V(q)\xi']_+ \in \mathcal{Y}_0^+$, $y_+'' \in \overline{\mathcal{Y}^+}'$ and $\mathcal{Y}_0^+$ and $\overline{\mathcal{Y}^+}'$ are, by Corollary 1 of Lemma 3.2.4, linearly independent it follows that (3.6.3) is equivalent to

$$[V(q)\xi']_+ = 0 \qquad (3.6.4)$$

$$y_+'' = 0 \qquad (3.6.5)$$

Let the $r \times r$ Laurent polynomial matrix $R(q)$ be the greatest common right divisor of $T(q)$ and $V(q)$. $R(q)$ is unique up to multiplication from the left by unimodular Laurent polynomial matrices. We have

$$T(q) = T_1(q)R(q) \qquad (3.6.6)$$

$$V(q) = V_1(q)R(q) \qquad (3.6.7)$$

where $T_1(q)$ and $V_1(q)$ are relatively right prime. Insert (3.6.7) into (3.6.4)

$$[V_1(q)R(q)\xi']_+ = 0 \qquad (3.6.8)$$

We will show that (3.6.8), and therefore (3.6.4), is equivalent to

$$R(q)\xi' = 0 \qquad (3.6.9)$$

If (3.6.9) is true then clearly (3.6.8) is true.

Since $T_1(q)$ and $V_1(q)$ are relatively right prime there are by Theorem 2.1.4 Laurent polynomial matrices $X(q)$ and $Y(q)$, respectively $r \times r$ and $r \times m$, such that

$$X(q)T_1(q) + Y(q)V_1(q) = I \tag{3.6.10}$$

Define y' as

$$y' = V_1(q)R(q)\xi' \tag{3.6.11}$$

Then we have

$$\begin{aligned} Y(q)y' &= Y(q)V_1(q)R(q)\xi' = \\ &= [I - X(q)T_1(q)]R(q)\xi' = \\ &= R(q)\xi' - X(q)T(q)\xi' = R(q)\xi' \end{aligned} \tag{3.6.12}$$

where the first equality follows from (3.6.11), the second from (3.6.10), the third from (3.6.6) and the fourth because $\xi' \in X_0$.

If (3.6.8) is true then $y'_+ = 0$. From (3.6.12) it follows that $R(q)\xi'$ is zero for $t \geq t_0$, some t_0. Lemma 3.2.2 then shows that (3.6.9) is true since $R(q)\xi'$ satisfies the equation $T_1(q)R(q)\xi' = 0$.

We have shown that (3.6.3) is equivalent to

$$R(q)\xi' = 0 \tag{3.6.13}$$

$$y''_+ = 0 \tag{3.6.14}$$

i.e. the nullspace of f_0 is given by the solutions to (3.6.13) and (3.6.14). By Theorem 2.2.3 $\xi' = 0$ is the only solution to (3.6.13) if and only if $R(q)$ is a unimodular Laurent polynomial matrix, i.e. if and only if $T(q)$ and

V(q) are relatively right prime.

□

<u>Remark</u>. Observability was defined for states at $t = 1$ but the theorem shows that the property of observability is independent of t.

<u>Corollary</u>. Let $R(q)$ be the greatest common right divisor of $T(q)$ and $V(q)$. Let $X_0^{(0)}$ be the space of all solutions $\xi' \in R_z^r$ to

$$R(q)\xi' = 0 \tag{3.6.15}$$

Then $X_0^{(0)} \times 0$, where 0 is the zero element in Y^+, is the unobservable subspace of the system (3.6.1).

<u>Proof</u>. The corollary follows from the fact that the null-space of f_0 is given by the solutions (ξ', y_+'') to (3.6.13) and (3.6.14).

□

<u>Remark</u>. The greatest common right divisor $R(q)$ is unique up to multiplication from the left by unimodular Laurent polynomial matrices. The solutions in R_z^r to (3.6.15) are not affected when (3.6.15) is multiplied from the left by a unimodular Laurent polynomial matrix. This follows from Theorem 2.2.3. Therefore $R(q)$, in (3.6.15), can be substituted by any greatest common right divisor of $T(q)$ and $V(q)$.

From the corollary it follows that the zeros of $\det R(q)$ are those zeros of $\det T(q)$ that correspond to solutions in X_0 that do not influence the output, they are decoupled from the output. As in Section 3.5 it follows that if $S_R(q)$ is the Smith form of $R(q)$ then $(S_R^T(q) \;\; 0)^T$ is the Smith form of $(T^T(q) \;\; V^T(q))^T$. Following Rosenbrock (1970) we make the definition.

Def. 3.6.2. The output decoupling (o.d.) zeros of a system are the zeros of the invariant factors of the Smith form of $\left(T^T(q) \quad V^T(q)\right)^T$.

Let $\{\beta_i\}$ be the set of i.d. zeros of the system (3.6.1) and let $\{\theta_i\}$ be the set of zeros to the invariant factors of $\left(T_1(q) \quad U(q)\right)$ where $T_1(q)$ is given by (3.6.6). As in Rosenbrock (1970) it is shown that $\{\theta_i\} \subset \{\beta_i\}$.

Def. 3.6.3. The set $\{\beta_i\} \setminus \{\theta_i\}$ is called the set of input-output decoupling (i.o.d.) zeros of the system (3.6.1).

Remark. Rosenbrock (1970) considers systems (3.6.1) with $T(q)$, $U(q)$, $V(q)$ and $W(q)$ restricted to polynomial matrices. Computation of the decoupling zeros according to Def. 3.5.2, 3.6.2 and 3.6.3 for such a system will give the same decoupling zeros as the ones obtained from Rosenbrock's definitions except for the ones that are zero. Our definition will never give any decoupling zeros that are zero.

Example 3.6.1. Consider the system

$$(q-1)(q-2)\xi = u \tag{3.6.14a}$$

$$y = (q-2)\xi \tag{3.6.14b}$$

(Compare with the system in Example 3.5.1) The system is of second order and the state space is spanned by

$$(\xi', y_+'') = (1,0) \tag{3.6.15}$$

$$(\xi', y_+'') = (2^t,0) \tag{3.6.16}$$

By the corollary of Theorem 3.6.1 the unobservable subspace is spanned by the state (3.6.16). We will show this directly.

The output is, for $t \geq 1$ and $u_+ = 0$, given by the state in the following way

$$y_+ = [V(q)\xi']_+ + y''_+ \tag{3.6.17}$$

An arbitrary state is given by a linear combination of the states (3.6.15) and (3.6.16).

$$(\xi', y''_+) = a(1,0) + b(2^t,0) \qquad a,b \in R \tag{3.6.18}$$

Introducing (3.6.18) into (3.6.17) gives

$$y_+ = [(q-2)(a+b2^t)]_+ + 0 = [-a]_+ \tag{3.6.19}$$

It follows that $y_+ = 0$ if and only if $a = 0$. Inserting $a = 0$ into (3.6.18) gives the unobservable subspace

$$(\xi', y''_+) = b(2^t,0) \qquad b \in R \tag{3.6.20}$$

□

4. AN EQUIVALENCE RELATION

An equivalence relation on the class of systems considered in Chapter 3 is defined in Section 4.1. In Section 4.2 it is shown that a system can be described by its system matrix and an equivalence relation on the set of system matrices is defined. It is shown that two systems are equivalent if and only if their system matrices are equivalent. In Section 4.3 it is shown that concepts like the order of a system, controllability, observability, stability and causality, defined in Chapter 3, are invariant under equivalence. In Section 4.4 it is examined when a system matrix has equivalent system matrices, which are in certain simple forms. In Section 4.5 we specialize to causal systems and give methods for computing the order of a system when the system matrix is in one of a few special forms. In Section 4.6 and 4.7 controllability indices and observability indices are defined from properties of the solutions to the differential equations. It is shown how these indices can be computed from the system matrix.

4.1 Definition of equivalence

As in Chapter 3 consider the system

$$T(q)\xi = U(q)u \tag{4.1.1a}$$

$$y = V(q)\xi + W(q)u \tag{4.1.1b}$$

where $T(q) \in R^{r\times r}(q]$, $U(q) \in R^{r\times \ell}(q]$, $V(q) \in R^{m\times r}(q]$, $W(q) \in R^{m\times \ell}(q]$, $u \in R_z^{\ell}$, $\xi \in R_z^{r}$, $y \in R_z^{m}$ and $\det T(q) \neq 0$.

Recall the definition (Def. 3.2.2) of the set of solutions X_u and the set of outputs Y_u to the system (4.1.1) for the fixed input $u \in \overline{R}^{\ell}_z$. Consider another system of the same type

$$T_1(q)\xi_1 = U_1(q)u_1 \qquad (4.1.2a)$$

$$y_1 = V_1(q)\xi_1 + W_1(q)u_1 \qquad (4.1.2b)$$

where $T_1(q) \in R^{r1 \times r1}(q]$, $U_1(q) \in R^{r1 \times \ell}(q]$, $V_1(q) \in R^{m \times r1}(q]$, $W_1(q) \in R^{m \times \ell}(q]$, $u_1 \in \overline{R}^{\ell}_z$, $\xi_1 \in R^{r1}_z$, $y \in R^m_z$ and $\det T_1(q) \neq 0$. Let the system (4.1.2) have the set of solutions and outputs $\mathcal{X}^1_{u1}$ and $\mathcal{Y}^1_{u1}$ respectively.

Def. 4.1.1. The systems (4.1.1) and (4.1.2) are equivalent if for any $u \in \overline{R}^{\ell}_z$ and $u_1 = u$

(i) there are $Z(q) \in R^{r1 \times r}(q]$ and $Y(q) \in R^{r1 \times \ell}(q]$ such that the mapping

$$\xi_1 = Z(q)\xi + Y(q)u \qquad (4.1.3)$$

is a bijective mapping from $\mathcal{X}_u$ to $\mathcal{X}^1_u$ and

(ii) $\mathcal{Y}^1_u = \mathcal{Y}_u$ and the diagram

$$\begin{array}{ccc} \mathcal{X}_u & \xrightarrow{(4.1.1b)} & \\ {\scriptstyle (4.1.3)}\downarrow & & \mathcal{Y}_u \\ \mathcal{X}^1_u & \xrightarrow{(4.1.2b)} & \end{array} \qquad (4.1.4)$$

commutes.

□

Remark 1. Note that the matrices $Z(q)$ and $Y(q)$ must be independent of u.

Remark 2. It is easy to show that this is an equivalence relation.

We will immediately give an algebraic condition for testing injectivity of the mapping (4.1.3).

Theorem 4.1.1. The mapping (4.1.3) is injective if and only if the Laurent polynomial matrices $T(q)$ and $Z(q)$ are relatively right prime. The inverse, defined on the range of the mapping (4.1.3), is of the same type, i.e.

$$\xi = Q(q)\xi_1 + L(q)u \tag{4.1.5}$$

Proof. Suppose $T(q)$ and $Z(q)$ are not relatively right prime. Then there is a nonunimodular matrix $R(q)$ such that

$$T_1(q)R(q)\xi = U(q)u \tag{4.1.6}$$

$$\xi_1 = Z_1(q)R(q)\xi + Y(q)u \tag{4.1.7}$$

Since $R(q)$ is not unimodular there is a nonzero $\xi' \in R_z^r$ such that

$$R(q)\xi' = 0 \tag{4.1.8}$$

Let ξ be an arbitrary element in X_u. Then $\xi + \xi' \in X_u$. The two solutions ξ and $\xi + \xi'$ give the same ξ_1 and therefore (4.1.3) is not injective.

If $T(q)$ and $Z(q)$ are relatively right prime then there are, by Theorem 2.1.4, $Q(q)$ and $X(q)$ such that

$$X(q)T(q) + Q(q)Z(q) = I \tag{4.1.9}$$

Therefore

$$\begin{aligned} Q(q)\xi_1 &= Q(q)Z(q)\xi + Q(q)Y(q)u = \\ &= \xi - X(q)T(q)\xi + Q(q)Y(q)u = \\ &= \xi + [Q(q)Y(q) - X(q)U(q)]\,u \end{aligned} \tag{4.1.10}$$

where the first equality follows from (4.1.3), the second from (4.1.9) and the third from (4.1.1a). Introducing $L(q) = - Q(q)Y(q) + X(q)U(q)$ gives (4.1.5).

□

Comparing with Theorem 3.6.1 we see that the mapping (4.1.3) is injective if and only if the system

$$T(q)\xi = U(q)u$$

$$\xi_1 = Z(q)\xi + Y(q)u$$

is observable.

We will briefly discuss the different conditions of Definition 4.1.1. The behaviour of system (4.1.1) for a given input u is completely described by a solution ξ, which is in X_u. It is therefore natural to demand from the equivalent system (4.1.2) that for $u_1 = u$ there is a unique solution ξ_1 in X_u^1 to every ξ in X_u and vice versa, i.e. that there exist a bijective transformation between X_u and X_u^1. Furthermore, it is natural to demand that two corresponding solutions ξ and ξ_1 shall give the same output y. In other words it is natural to demand that there exist a bijective mapping (*) such that the diagram

$$(*) \quad \begin{array}{ccc} X_u & \xrightarrow{(4.1.1b)} & \\ \downarrow & & y_u \\ X_u^1 & \xrightarrow{(4.1.2b)} & \end{array} \qquad (4.1.11)$$

commutes. It is, however, not clear that the mapping (*) shall be of the form (4.1.3).

In one special case, namely if the two systems (4.1.1) and (4.1.2) are observable, then the mapping (*) must be of the form (4.1.3). This can be shown in the following way.

Since the system (4.1.2) is observable the matrices $T_1(q)$ and $V_1(q)$ are by Theorem 3.6.1 relatively right prime. The mapping (4.1.2b) is therefore by Theorem 4.1.1 invertible and the inverse is given by

$$\xi_1 = Q(q)y_1 + L(q)u_1 \tag{4.1.12}$$

where $Q(q)$ and $L(q)$ are Laurent polynomial matrices. If the systems (4.1.1) and (4.1.2) are equivalent then $u_1 = u$ implies that $y_1 = y$ by (4.1.4). Using this and inserting (4.1.1b) into (4.1.12) gives

$$\xi_1 = Q(q)[V(q)\xi + W(q)u] + L(q)u$$

$$\Leftrightarrow \xi_1 = Q(q)V(q)\xi + [Q(q)W(q) + L(q)]u \tag{4.1.13}$$

which is of the form (4.1.3).

For nonobservable systems one has to introduce further conditions on the mapping (*) to get the form (4.1.3) as the only possibility. We will not do that, but instead directly demand the structure (4.1.3) even for nonobservable systems. This structure has the pleasing feature that the restriction to X_0 has the form

$$\xi_1' = Z(q)\xi' \tag{4.1.14}$$

i.e. it is a linear, time invariant mapping between X_0 and X_0^1. Since the state spaces for the two systems are $X_0 \times \overline{Y^{+1}}$ and $X_0^1 \times \overline{Y^{+1}}$ respectively, then there is a bijective, linear, time invariant mapping between the state spaces of equivalent systems.

Example 4.1.1. Consider the two systems

$$(q-1)(q-2)\xi = q^{-1}u \tag{4.1.15a}$$

$$y = (1+q^{-1})\xi \tag{4.1.15b}$$

and

$$q^2(q-1)(q-2)\xi_1 = (q+1)u_1 \tag{4.1.16a}$$

$$y_1 = \xi_1 \tag{4.1.16b}$$

We will show that the two systems are equivalent. Put $u_1 = u$ and let u be an arbitrary element in $\overline{R}_Z$. The elements in X_u are given by

$$\xi = \xi^p + a + b2^t \qquad a,b \in R \tag{4.1.17}$$

where ξ^p is a particular solution to (4.1.15a). Analogously the elements in X_u^1 are given by

$$\xi_1 = \xi_1^p + c + d2^t \qquad c,d \in R \tag{4.1.18}$$

Let the linear mapping

$$f(\xi) = (1+q^{-1})\xi \tag{4.1.19}$$

be defined on X_u. The range space of f is obtained by inserting (4.1.17) into (4.1.19).

$$f(\xi) = (1+q^{-1})\xi^p + (1+q^{-1})(a+b2^t) =$$

$$= (1+q^{-1})\xi^p + 2a + \frac{3}{2}b2^t \tag{4.1.20}$$

Inserting $\xi_1 = (1+q^{-1})\xi^p$ into (4.1.16a) and using (4.1.15a) shows that $(1+q^{-1})\xi^p$ is a particular solution to (4.1.16a). Putting

$$\xi_1^p = (1+q^{-1})\xi^p \tag{4.1.20}$$

$$c = 2a \tag{4.1.21}$$

$$d = \frac{3}{2}b \tag{4.1.22}$$

we see that any $f(\xi)$, given by (4.1.20), belongs to X_u^1 and vice versa. Therefore $f(\xi)$, $\xi \in X_u$ is a surjective mapping from X_u to X_u^1. It is injective by Theorem 4.1.1 since the Laurent polynomials $(q-1)(q-2)$ and $(1+q^{-1})$ are relatively prime. The injectivity can also easily be shown directly. We have shown that (i) of Def. 4.1.1 is fulfilled.

Inserting $\xi_1 = (1+q^{-1})\xi$ into (4.1.16b) and using (4.1.15b) gives

$$y_1 = \xi_1 = (1+q^{-1})\xi = y \tag{4.1.23}$$

This shows that (ii) of Def. 4.1.1 is fulfilled.

□

4.2 The system matrix

The system equations

$$T(q)\xi = U(q)u \tag{4.2.1a}$$

$$y = V(q)\xi + W(q)u \tag{4.2.1b}$$

can be written

$$\left(\begin{array}{c|c} T(q) & U(q) \\ \hline -V(q) & W(q) \end{array}\right)\left(\begin{array}{c} \xi \\ \hline -u \end{array}\right) = \left(\begin{array}{c} 0 \\ \hline -y \end{array}\right) \tag{4.2.2}$$

Here $T(q) \in R^{r\times r}[q]$, $U(q) \in R^{r\times \ell}[q]$, $V(q) \in R^{m\times r}[q]$, $W(q) \in R^{m\times \ell}[q]$ and $\det T(q) \neq 0$.

Def. 4.2.1 The Laurent polynomial matrix

$$P(q) \triangleq \left(\begin{array}{c|c} T(q) & U(q) \\ \hline -V(q) & W(q) \end{array}\right) \tag{4.2.3}$$

is called the system matrix for the system (4.2.1).

Remark. Notice the two differences from the system matrix defined by Rosenbrock (1970). Rosenbrock considers polynomial matrices and demands that $r \geq n$, where r is the dimension of $T(q)$ and n the order of the system. We consider Laurent polynomial matrices and have no restrictions on r.

We can now define an equivalence relation on the set of system matrices. Let

$$P_1(q) = \left(\begin{array}{c|c} T_1(q) & U_1(q) \\ \hline -V_1(q) & W_1(q) \end{array}\right) \tag{4.2.4}$$

be another system matrix.

Def. 4.2.2. The system matrices $P(q)$ and $P_1(q)$ are equivalent if there is a nonnegative integer k and Laurent polynomial matrices $M(q)$, $N(q)$, $X(q)$ and $Y(q)$, where $M(q)$ and $N(q)$ are unimodular, such that

$$\left(\begin{array}{cc|c} I_{k-r_1} & 0 & 0 \\ 0 & T_1(q) & U_1(q) \\ \hline 0 & -V_1(q) & W_1(q) \end{array}\right) = \left(\begin{array}{c|c} M(q) & 0 \\ \hline X(q) & I \end{array}\right) \left(\begin{array}{cc|c} I_{k-r} & 0 & 0 \\ 0 & T(q) & U(q) \\ \hline 0 & -V(q) & W(q) \end{array}\right) \cdot$$

$$\cdot \left(\begin{array}{c|c} N(q) & Y(q) \\ \hline 0 & I \end{array}\right) \tag{4.2.5}$$

Here I_a is an identity matrix of dimension a and r and r_1 are the dimensions of the matrices $T(q)$ and $T_1(q)$ respectively.

Remark. It is easy to show that this is an equivalence relation.

This definition of equivalence is equal to Rosenbrock's strict system equivalence except that we allow Laurent polynomial matrices while Rosenbrock demands polynomial matrices.

We will eventually show that two systems are equivalent if and only if their system matrices are equivalent. One half of this result is relatively easy to prove.

Theorem 4.2.1. Two systems are equivalent if their system matrices are equivalent.

Proof. The equations corresponding to the system matrix can be written

$$\left(\begin{array}{cc|c} I_{k-r} & 0 & 0 \\ 0 & T(q) & U(q) \\ \hline 0 & -V(q) & W(q) \end{array}\right)\left(\begin{array}{c} \tilde{\xi} \\ \xi \\ \hline -u \end{array}\right) = \left(\begin{array}{c} 0 \\ 0 \\ \hline -y \end{array}\right) \qquad (4.2.6)$$

It follows from (4.2.6) that $\tilde{\xi} = 0$. Multiply (4.2.6) from the left by

$$\left(\begin{array}{c|c} M(q) & 0 \\ \hline X(q) & I \end{array}\right) \qquad (4.2.7)$$

and introduce

$$P_2(q) = \left(\begin{array}{c|c} M(q) & 0 \\ \hline X(q) & I \end{array}\right) \left(\begin{array}{cc|c} I_{k-r} & 0 & 0 \\ 0 & T(q) & U(q) \\ \hline 0 & -V(q) & W(q) \end{array}\right) \tag{4.2.8}$$

This gives

$$P_2(q) \left(\begin{array}{c} \tilde{\xi} \\ \xi \\ \hline -u \end{array}\right) = \left(\begin{array}{c} 0 \\ 0 \\ \hline -y \end{array}\right) \tag{4.2.9}$$

It follows that y is not influenced by this operation. The equations determining $(\tilde{\xi}^T \ \xi^T)^T$ are multiplied from the left by the unimodular matrix $M(q)$ and therefore $(\tilde{\xi}^T \ \xi^T)^T$ is not influenced. Equation (4.2.9) can be written

$$P_2(q) \left(\begin{array}{c|c} N(q) & Y(q) \\ \hline 0 & I \end{array}\right) \left(\begin{array}{c|c} N^{-1}(q) & -N^{-1}(q)\,Y(q) \\ \hline 0 & I \end{array}\right) \left(\begin{array}{c} \tilde{\xi} \\ \xi \\ \hline -u \end{array}\right) =$$

$$= \left(\begin{array}{c} 0 \\ 0 \\ \hline -y \end{array}\right) \tag{4.2.10}$$

Introducing

$$\begin{pmatrix} \tilde{\xi}_1 \\ \xi_1 \end{pmatrix} = N^{-1}(q) \begin{pmatrix} \tilde{\xi} \\ \xi \end{pmatrix} - N^{-1}(q)Y(q)u \tag{4.2.11}$$

and using (4.2.5) gives

$$\left(\begin{array}{cc|c} I_{k-r_1} & 0 & 0 \\ 0 & T_1(q) & U_1(q) \\ \hline 0 & -V_1(q) & W_1(q) \end{array}\right) \left(\begin{array}{c} \tilde{\xi}_1 \\ \xi_1 \\ \hline -u \end{array}\right) = \left(\begin{array}{c} 0 \\ 0 \\ \hline -y \end{array}\right) \tag{4.2.12}$$

It follows that $\tilde{\xi}_1 = 0$ and (4.2.11) can be written

$$\begin{pmatrix} 0 \\ \xi_1 \end{pmatrix} = N^{-1}(q)\begin{pmatrix} 0 \\ \xi \end{pmatrix} - N^{-1}(q)Y(q)u \qquad (4.2.13)$$

It follows from (4.2.12) that this is a mapping from X_u to X_u^1 such that the diagram

$$\begin{array}{ccc} X_u & \xrightarrow{(4.2.1b)} & Y_u \\ {\scriptstyle (4.2.13)}\downarrow & \nearrow_{(*)} & \\ X_u^1 & & \end{array} \qquad (4.2.14)$$

commutes. Here the mapping (*) is given by the last row of (4.2.12). It remains to be shown that (4.2.13) is bijective.

Suppose that $\xi \in X_u$ and $\xi + \Delta\xi \in X_u$ gives the same ξ_1 through (4.2.13). This implies that

$$N^{-1}(q)\begin{pmatrix} 0 \\ \Delta\xi \end{pmatrix} = 0 \qquad (4.2.15)$$

Since $N^{-1}(q)$ is unimodular it follows by Theorem 2.2.3 that $\Delta\xi = 0$ and (4.2.13) is injective.

Let ξ_1 be an arbitrary element in X_u^1. By an argument analogous to the one above it follows that the mapping

$$\begin{pmatrix} 0 \\ \xi \end{pmatrix} = N(q)\begin{pmatrix} 0 \\ \xi_1 \end{pmatrix} + Y(q)u \qquad (4.2.16)$$

gives a ξ in X_u. This ξ is by (4.2.13) mapped onto ξ_1. This means that (4.2.13) is surjective.

Partition $N^{-1}(q)$ and $-N^{-1}(q)Y(q)$ as

$$N^{-1}(q) = \begin{pmatrix} Z_{11}(q) & Z_{12}(q) \\ Z_{21}(q) & Z_{22}(q) \end{pmatrix} \tag{4.2.17}$$

$$N^{-1}(q)Y(q) = \begin{pmatrix} Y_1(q) \\ Y_2(q) \end{pmatrix} \tag{4.2.18}$$

Then (4.2.13) implies

$$\xi_1 = Z_{22}(q)\xi + Y_2(q)u \tag{4.2.19}$$

This is the desired bijective mapping between X_u and X_u^1 making the diagram (4.2.14) commute.

□

Before we can prove the converse of Theorem 4.2.1 we need a few other results.

Theorem 4.2.2. Two equivalent systems have the same transfer function.

Proof. We use the notation introduced in Section 4.1. Equation (4.1.1a) has a unique solution $\xi = T^{-1}(q)U(q)u$ in ΓR_z^r. By (4.1.3) this ξ is mapped to a ξ_1 in $\Gamma R_z^{r_1}$. By uniqueness this ξ_1 is equal to $\xi_1 = T_1^{-1}(q)U_1(q)u$. Diagram (4.1.4) says that ξ and ξ_1 are mapped to the same y by (4.1.1b) and (4.1.2b) respectively, i.e.

$$V(q)T^{-1}(q)U(q)u + W(q)u = V_1(q)T_1^{-1}(q)U_1(q)u + W_1(q)u \tag{4.2.20}$$

Introducing the transfer functions $G(q)$ and $G_1(q)$

$$\big(G(q) - G_1(q)\big)u = 0 \qquad \forall\, u \in \Gamma R_z^{\ell} \tag{4.2.21}$$

Letting all components of u be zero except the j:th gives

$$\left(g(q)^{ij} - g_1(q)^{ij}\right) u_j = 0 \qquad i = 1,\dots,m \tag{4.2.22}$$

where $g(q)^{ij}$ is the i,j:th element of $G(q)$. Since (4.2.22) is true for all u_j it follows by Theorem 2.2.2 that $g(q)^{ij} - g_1(q)^{ij} = 0$. This is true for all i and j. Therefore

$$G(q) = G_1(q) \tag{4.2.23}$$

□

Theorem 4.2.3. Consider two systems in state space form

$$(qI-A)x = Bu \tag{4.2.24a}$$

$$y = Cx + D(q)u \tag{4.2.24b}$$

and

$$(qI-A_1)x_1 = B_1 u \tag{4.2.25a}$$

$$y_1 = C_1 x_1 + D_1(q)u \tag{4.2.25b}$$

Suppose that the matrices $(A \quad B)$, $(A^T \quad C^T)$, $(A_1 \quad B_1)$ and $(A_1^T \quad C_1^T)$ all have linearly independent rows and that the systems (4.2.24) and (4.2.25) are equivalent. Then the two systems are system similar (s.s.), i.e. there is a non-singular matrix T such that

$$A_1 = TAT^{-1}, \quad B_1 = TB, \quad C_1 = CT^{-1} \quad \text{and}$$

$$D_1(q) = D(q) \tag{4.2.26}$$

Remark 1. The relations (4.2.26) clearly imply that the system matrices corresponding to (4.2.24) and (4.2.25) are equivalent.

Remark 2. The condition that the rows of $(A \quad B)$ are linearly independent means in terms of Rosenbrock (1970) that the system (4.2.24) has no input decoupling zero that is equal to zero. The condition on $(A^T \quad C^T)$ means the same thing for output decoupling zeros.

Proof. The systems (4.2.24) and (4.2.25) can by s.s. be brought to the form

$$\begin{pmatrix} qI-\bar{A} & 0 \\ 0 & qI-\tilde{A} \end{pmatrix} \begin{pmatrix} \bar{x} \\ \tilde{x} \end{pmatrix} = \begin{pmatrix} \bar{B} \\ \tilde{B} \end{pmatrix} u \tag{4.2.27a}$$

$$y = (\bar{C} \quad \tilde{C}) \begin{pmatrix} \bar{x} \\ \tilde{x} \end{pmatrix} + D(q)u \tag{4.2.27b}$$

and

$$\begin{pmatrix} qI-\bar{A}_1 & 0 \\ 0 & qI-\tilde{A}_1 \end{pmatrix} \begin{pmatrix} \bar{x}_1 \\ \tilde{x}_1 \end{pmatrix} = \begin{pmatrix} \bar{B}_1 \\ \tilde{B}_1 \end{pmatrix} u \tag{4.2.28a}$$

$$y_1 = (\bar{C}_1 \quad \tilde{C}_1) \begin{pmatrix} \bar{x}_1 \\ \tilde{x}_1 \end{pmatrix} + D_1(q)u \tag{4.2.28b}$$

where all eigenvalues of $\tilde{A}$ and $\tilde{A}_1$ are zero while no eigenvalues of $\bar{A}$ and $\bar{A}_1$ are zero. The transformations from (4.2.24) and (4.2.25) to (4.2.27) and (4.2.28) can be achieved with real transformation matrices.

It is therefore equivalent to show the theorem for the systems (4.2.27) and (4.2.28).

Theorem 4.2.2 gives that

$$\bar{C}(qI-\bar{A})^{-1}\bar{B} + \tilde{C}(qI-\tilde{A})^{-1}\tilde{B} + D(q) =$$

$$= \bar{C}_1(qI-\bar{A}_1)^{-1}\bar{B}_1 + \tilde{C}_1(qI-\tilde{A}_1)^{-1}\tilde{B}_1 + D_1(q) \tag{4.2.29}$$

The strictly proper part with poles that are zero is uniquely determined. Therefore

$$\tilde{C}(qI-\tilde{A})^{-1}\tilde{B} = \tilde{C}_1(qI-\tilde{A}_1)\tilde{B}_1 \tag{4.2.30}$$

The fact that $(A \quad B)$ has independent rows implies that $(\tilde{A} \quad \tilde{B})$ have independent rows. Since the eigenvalues of $\tilde{A}$ all are zero this means that the rows of $(xI-\tilde{A} \quad \tilde{B})$ are independent for all complex x and the pair $(\tilde{A},\tilde{B})$ is a controllable pair. In the same way it is shown that $(\tilde{A}_1,\tilde{B}_1)$ is a controllable pair and that $(\tilde{A},\tilde{C})$ and $(\tilde{A}_1,\tilde{C}_1)$ are observable pairs. It therefore follows from (4.2.30) that there is a nonsingular $\tilde{T}$ such that

$$\tilde{A}_1 = \tilde{T}\tilde{A}\tilde{T}^{-1} \qquad \tilde{B}_1 = \tilde{T}\tilde{B} \qquad \tilde{C}_1 = \tilde{C}\tilde{T}^{-1} \tag{4.2.31}$$

This is a standard result in system theory shown i.e. by Rosenbrock (1970).

Since the systems (4.2.27) and (4.2.28) are equivalent there are Laurent polynomial matrices $Z(q)$ and $Y(q)$ such that

$$\begin{pmatrix} \bar{x}_1 \\ \tilde{x}_1 \end{pmatrix} = Z(q)\begin{pmatrix} \bar{x} \\ \tilde{x} \end{pmatrix} + Y(q)u \tag{4.2.32}$$

is a bijective mapping between X_u and X_u^1. Its inverse is by Theorem 4.1.1 given by

$$\begin{pmatrix} \bar{x} \\ \tilde{x} \end{pmatrix} = Q(q)\begin{pmatrix} \bar{x}_1 \\ \tilde{x}_1 \end{pmatrix} + L(q)u \tag{4.2.33}$$

Partitioning $Z(q)$ and $Y(q)$ gives

$$\begin{pmatrix} \bar{x}_1 \\ \tilde{x}_1 \end{pmatrix} = \begin{pmatrix} Z_1(q) & Z_2(q) \\ Z_3(q) & Z_4(q) \end{pmatrix}\begin{pmatrix} \bar{x} \\ \tilde{x} \end{pmatrix} + \begin{pmatrix} Y_1(q) \\ Y_2(q) \end{pmatrix}u \tag{4.2.34}$$

which implies

$$\bar{x}_1 = Z_1(q)\bar{x} + Z_2(q)\tilde{x} + Y_1(q)u \qquad (4.2.35)$$

Equation (4.2.27a) implies

$$(qI-\tilde{A})\tilde{x} = \tilde{B}u \qquad (4.2.36)$$

But $(qI-\tilde{A})$ is a unimodular Laurent polynomial matrix because it is polynomial and $\det(qI-\tilde{A}) = q^k$, where k is the dimension of $\tilde{A}$. Therefore by Theorem 2.2.3

$$\tilde{x} = (qI-\tilde{A})^{-1}\tilde{B}u \qquad (4.2.37)$$

where $(qI-\tilde{A})^{-1}\tilde{B}$ is a Laurent polynomial matrix. Introducing (4.2.37) into (4.2.35) gives

$$\bar{x}_1 = Z_1(q)\bar{x} + Y_3(q)u \qquad (4.2.38)$$

where $Y_3(q) = Z_2(q)(qI-\tilde{A})^{-1}\tilde{B} + Y_1(q)$.

Write $Z_1(q) = Z_5(q) + Z_6(q)$, where $Z_5(q)$ is a polynomial matrix and $Z_6(q)$ contains only negative powers of q. Using the division algorithm for polynomial matrices, Mac Duffee (1946), gives

$$Z_5(q) = Z_7(q)(qI-\bar{A}) + Z_8 \qquad (4.2.39)$$

where $Z_7(q)$ is a polynomial matrix and Z_8 is independent of q. Using the invertibility of $\bar{A}$ it can analogously be shown that $Z_6(q)$ can be written

$$Z_6(q) = Z_9(q)(qI-\bar{A}) + Z_{10} \qquad (4.2.40)$$

where $Z_9(q)$ is a polynomial matrix in q^{-1} and Z_{10} is independent of q. Introduce (4.2.39) and (4.2.40) into (4.2.38).

$$\bar{x}_1 = Z_7(q)(qI-\bar{A})\bar{x} + Z_8\bar{x}+Z_9(q)(qI-\bar{A})\bar{x} + Z_{10}\bar{x} + Y_3(q)u =$$

$$= Z_7(q)\bar{B}u + Z_8\bar{x} + Z_9(q)\bar{B}u + Z_{10}\bar{x} + Y_3(q)u =$$

$$= H\bar{x} + Y_4(q)u \tag{4.2.41}$$

where $H = Z_8 + Z_{10}$ and $Y_4(q) = Z_7(q)\bar{B} + Z_9(q)\bar{B} + Y_3(q)$. The second equality in (4.2.41) follows from (4.2.27a). Analogously it follows from (4.2.33) that

$$\bar{x} = S\bar{x}_1 + L_1(q)u \tag{4.2.42}$$

where S is independent of q. Introducing (4.2.41) into (4.2.42) gives

$$\bar{x} = SH\bar{x} + [SY_n(q) + L_1(q)]u \tag{4.2.42}$$

Putting $u = 0$

$$\bar{x} = SH\bar{x}, \quad \forall\ \bar{x} \text{ solution to } q\bar{x} = \bar{A}\bar{x} \tag{4.2.43}$$

Or since $\bar{A}$ is nonsingular

$$SH = I \tag{4.2.44}$$

Analogously

$$HS = I \tag{4.2.45}$$

Therefore H is nonsingular.

Introducing (4.2.41) into (4.2.28a) gives

$$(qI-\bar{A}_1)\left(H\bar{x}+Y_4(q)u\right) = \bar{B}_1u$$

$$\Leftrightarrow (qI-H^{-1}\bar{A}_1H)\bar{x} = H^{-1}[\bar{B}_1 - (qI-\bar{A}_1)Y_4(q)]u \tag{4.2.46}$$

Compare this with (4.2.27a)

$$(qI-\bar{A})\bar{x} = \bar{B}u \tag{4.2.47}$$

Put $u = 0$ and subtract (4.2.47) from (4.2.46)

$$(\bar{A} - H^{-1}\bar{A}_1 H)\bar{x} = 0, \quad \forall\ \bar{x} \text{ solution to } q\bar{x} = \bar{A}\bar{x} \tag{4.2.48}$$

Or since $\bar{A}$ is nonsingular

$$\bar{A} = H^{-1}\bar{A}_1 H$$

$$\Leftrightarrow \bar{A}_1 = H\bar{A}H^{-1} \tag{4.2.49}$$

Introducing (4.2.49) into (4.2.46) and subtracting (4.2.47)

$$[H^{-1}\bar{B}_1 - \bar{B} - H^{-1}(qI-\bar{A}_1)Y_4(q)]u = 0$$

Or since this is true for all u

$$H^{-1}\bar{B}_1 - \bar{B} - H^{-1}(qI-\bar{A}_1)Y_4(q) = 0 \tag{4.2.50}$$

It follows that $(qI-\bar{A}_1)Y_4(q)$ is independent of q. But this implies that $Y_4(q) = 0$ because $\bar{A}_1$ is nonsingular. Therefore (4.2.50) gives

$$H^{-1}\bar{B}_1 = \bar{B}$$

$$\Leftrightarrow \bar{B}_1 = HB \tag{4.2.51}$$

and (4.2.41) gives

$$\bar{x}_1 = H\bar{x} \tag{4.2.52}$$

Introducing (4.2.52) into (4.2.28b) and using the fact that $\tilde{x}_1 = \tilde{T}\tilde{x}$

$$y_1 = \bar{C}_1 H\bar{x} + \tilde{C}_1\tilde{T}\tilde{x} + D_1(q)u \tag{4.2.53}$$

Now $y_1 = y$ because the systems are equivalent. Subtracting (4.2.27b) from (4.2.53) and using (4.2.31) gives

$$(\bar{C}_1 H - \bar{C})\bar{x} + \left(D_1(q) - D(q)\right)u = 0 \qquad (4.2.54)$$

Put $u = 0$. Then (4.2.54) is true for all solutions to $q\bar{x} = \bar{A}\bar{x}$. Because $\bar{A}$ is nonsingular this implies

$$\bar{C}_1 H - \bar{C} = 0$$

$$\Leftrightarrow \bar{C}_1 = \bar{C}H^{-1} \qquad (4.2.55)$$

Introducing (4.2.55) into (4.2.54) gives

$$D_1(q) = D(q) \qquad (4.2.56)$$

Thus taking

$$T = \begin{pmatrix} H & 0 \\ 0 & \tilde{T} \end{pmatrix}$$

with H and $\tilde{T}$ defined through (4.2.41) and (4.2.31) resp. gives the desired result.

□

For future purposes we need the following corollary.

<u>Corollary</u>. Consider the two systems

$$(qI-A)x = Bu$$

$$y = Cx + E(q)u$$

and

$$(qI-A_1)x_1 = B_1 u$$

$$y_1 = C_1 x_1 + E_1(q)u$$

where A and A_1 are nonsingular matrices and $E(q)$ and $E_1(q)$ are Laurent polynomial matrices. If the systems are equivalent then the corresponding two system matrices are related through the following equalities.

$$A_1 = TAT^{-1} \qquad B_1 = TB \qquad C_1 = CT^{-1} \qquad E_1(q) = E(q)$$

Proof. The proof follows the lines of the proof of the theorem. The following modifications are made:

In (4.2.27) and (4.2.28) $\tilde{A}$ and $\tilde{A}_1$ vanish because A and A_1 are nonsingular.

The part from (4.2.29) to (4.2.37) can be deleted.

The rest of the proof will remain invariant.

□

Theorem 4.2.4. Let

$$P(q) = \left(\begin{array}{c|c} T(q) & U(q) \\ \hline -V(q) & W(q) \end{array}\right)$$

be an arbitrary system matrix and let r be the dimension of $T(q)$ and n the order of the system. Then there is a state space representation $(A,B,C,D(q))$ with $D(q)$ a polynomial matrix, A an $n \times n$ matrix and $\text{rank}(A \quad B) = \text{rank}(A^T \quad C^T) = n$ such that

$$\left(\begin{array}{cc|c} I_{k-n} & 0 & 0 \\ 0 & qI-A & B \\ \hline 0 & -C & D(q) \end{array}\right) = \left(\begin{array}{c|c} M(q) & 0 \\ \hline X(q) & I \end{array}\right) \left(\begin{array}{cc|c} I_{k-r} & 0 & 0 \\ 0 & T(q) & U(q) \\ \hline 0 & -V(q) & W(q) \end{array}\right) \left(\begin{array}{c|c} N(q) & Y(q) \\ \hline 0 & I \end{array}\right) \tag{4.2.57}$$

for some $M(q)$, $N(q)$, $X(q)$ and $Y(q)$ as in Def. 4.2.2. Furthermore k can be chosen $k = \max(n,r)$.

Proof. We will prove the theorem for the case $r \le n$. To prove the theorem in the case $r > n$ only minor changes have to be made in this proof.

By Theorem 2.1.4 there are unimodular Laurent polynomial matrices $M_1(q)$ and $N_1(q)$ such that

$$S(q) = M_1(q)T(q)N_1(q) \tag{4.2.58}$$

where $S(q)$ is the Smith form of $T(q)$. Here $S(q)$ is by definition a polynomial matrix with deg det $S(q)$ = deg det $T(q)$ = $= n_0$. By a standard result in matrix theory there is an $n_0 \times n_0$ matrix A_1 and unimodular polynomial matrices $M_2(q)$ and $N_2(q)$ such that

$$\begin{pmatrix} I_{n-n_0} & 0 \\ 0 & qI-A_1 \end{pmatrix} = M_2(q)\begin{pmatrix} I_{n-r} & 0 \\ 0 & S(q) \end{pmatrix} N_2(q) \tag{4.2.59}$$

Therefore

$$\begin{pmatrix} I_{n-n_0} & 0 \\ 0 & qI-A_1 \end{pmatrix} = M(q)\begin{pmatrix} I_{n-r} & 0 \\ 0 & T(q) \end{pmatrix} N(q) \tag{4.2.60}$$

where

$$M(q) = M_2(q)\begin{pmatrix} I_{n-r} & 0 \\ 0 & M_1(q) \end{pmatrix}$$

and

$$N(q) = \begin{pmatrix} I_{n-r} & 0 \\ 0 & N_1(q) \end{pmatrix} N_2(q)$$

It follows that

$$P_2(q) \triangleq \left(\begin{array}{cc|c} I_{n-n_0} & 0 & U_1(q) \\ 0 & qI-A & U_2(q) \\ \hline -V_1(q) & -V_2(q) & W(q) \end{array}\right) \triangleq$$

$$\triangleq \left(\begin{array}{c|c} M(q) & 0 \\ \hline 0 & I \end{array}\right) \left(\begin{array}{cc|c} I_{n-r} & 0 & 0 \\ 0 & T(q) & U(q) \\ \hline 0 & -V(q) & W(q) \end{array}\right) \left(\begin{array}{c|c} N(q) & 0 \\ \hline 0 & I \end{array}\right) \tag{4.2.61}$$

Define $P_3(q)$ and $W_2(q)$ through

$$P_3(q) \triangleq \left(\begin{array}{cc|c} I_{n-n_0} & 0 & 0 \\ 0 & qI-A_1 & U_2(q) \\ \hline 0 & -V_2(q) & W_2(q) \end{array}\right) \triangleq$$

$$\triangleq \left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & I & 0 \\ \hline V_1(q) & 0 & I \end{array}\right) P_2(q) \left(\begin{array}{cc|c} I & 0 & -U_1(q) \\ 0 & I & 0 \\ \hline 0 & 0 & I \end{array}\right) \tag{4.2.62}$$

Observe that A_1 is nonsingular because the <u>polynomial</u> $\det S(q)$ has no zeros that are zero. Therefore as in the proof of Theorem 4.2.3 see (4.2.39-40) there is a Laurent polynomial matrix $U_3(q)$ and a constant matrix B_1 such that

$$U_2(q) = (qI-A_1)U_3(q) + B_1 \tag{4.2.63}$$

In the same way

$$V_2(q) = V_3(q)(qI-A_1) + C_1 \tag{4.2.64}$$

Define $P_4(q)$ and $W_3(q)$ through

$$P_4(q) \triangleq \left(\begin{array}{cc|c} I_{n-n_0} & 0 & 0 \\ 0 & qI-A_1 & B_1 \\ \hline 0 & -C_1 & W_3(q) \end{array}\right) \triangleq$$

$$\triangleq \left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & I & 0 \\ \hline 0 & V_3(q) & I \end{array}\right) P_3(q) \left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & I & -U_3(q) \\ \hline 0 & 0 & I \end{array}\right) \tag{4.2.65}$$

The transfer function for $P_4(q)$ is

$$G(q) = C_1(qI-A_1)^{-1}B_1 + W_3(q) \tag{4.2.66}$$

The first part has no poles that are zero while all the poles of the Laurent polynomial matrix $W_3(q)$ are zero. $W_3(q)$ can be uniquely decomposed as $W_3(q) = W_4(q) + D(q)$, where $W_4(q)$ is strictly proper and $D(q)$ is a polynomial matrix. By Theorem 3.2.3 we have

$$\nu(W_4) = n - n_0 \tag{4.2.67}$$

By standard theory for linear systems i.e. Rosenbrock (1970) there is a least order state space representation $\{A_0, B_0, C_0, D(q)\}$ of $W_3(q)$ with the dimension of A_0 equal to $n-n_0$. Therefore

$$W_3(q) = C_0(qI-A_0)B_0 + D(q) \tag{4.2.68}$$

Observe that $(qI-A_0)$ is a unimodular Laurent polynomial matrix since A_0 has all its eigenvalues equal to zero. Define $P_5(q)$ through

$$P_5(q) \triangleq \left(\begin{array}{cc|c} qI-A_0 & 0 & 0 \\ 0 & qI-A_1 & B_1 \\ \hline 0 & -C_1 & W_3(q) \end{array}\right) = P_4(q) \left(\begin{array}{cc|c} qI-A_0 & 0 & 0 \\ 0 & I & 0 \\ \hline 0 & 0 & I \end{array}\right) \tag{4.2.69}$$

and define $P_6(q)$ through

$$P_6(q) = \left(\begin{array}{cc|c} qI-A_0 & 0 & B_0 \\ 0 & qI-A_1 & B_1 \\ \hline -C_0 & -C_1 & D(q) \end{array}\right) =$$

$$= \left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & I & 0 \\ \hline -C_0(qI-A_0)^{-1} & 0 & I \end{array}\right) P_5(q) \left(\begin{array}{cc|c} I & 0 & (qI-A_0)^{-1}B_0 \\ 0 & I & 0 \\ \hline 0 & 0 & I \end{array}\right) \qquad (4.2.70)$$

We have shown that the system matrices $P(q)$, $P_2(q)$, $P_3(q)$, $P_4(q)$, $P_5(q)$ and $P_6(q)$ are equivalent. This means that (4.2.57) has been proven. Since $(A_0,B_0,C_0,D(q))$ is of least order we have that $(A_0 \ \ B_0)$ has linearly independent rows. Furthermore A_1 has linearly independent rows because A_1 has no zero eigenvalues. Therefore

$$\begin{pmatrix} A_0 & 0 & B_0 \\ 0 & A_1 & B_1 \end{pmatrix} \qquad (4.2.71)$$

has full rank equal to n. Analogously

$$\begin{pmatrix} A_0^T & 0 & C_0^T \\ 0 & A_1^T & C_1^T \end{pmatrix} \qquad (4.2.72)$$

has full rank equal to n.

□

For future purposes we need the following corollary.

Corollary. Let

$$P(q) = \left(\begin{array}{c|c} T(q) & U(q) \\ \hline -V(q) & W(q) \end{array}\right)$$

be an arbitrary system matrix and let r be the dimension of $T(q)$ and $n_0 = \underline{\deg} \det T(q)$. Then there are Laurent polynomial matrices $M(q)$, $N(q)$, $X(q)$ and $Y(q)$, with $M(q)$ and $N(q)$ unimodular, such that

$$\left(\begin{array}{cc|c} I_{k-n_0} & 0 & 0 \\ 0 & qI-A & B \\ \hline 0 & -C & E(q) \end{array}\right) = \left(\begin{array}{c|c} M(q) & 0 \\ \hline X(q) & I \end{array}\right) \left(\begin{array}{cc|c} I_{k-r} & 0 & 0 \\ 0 & T(q) & U(q) \\ \hline 0 & -V(q) & W(q) \end{array}\right) \left(\begin{array}{c|c} N(q) & Y(q) \\ \hline 0 & I \end{array}\right)$$

Here A is an $n_0 \times n_0$ nonsingular matrix and $E(q)$ is a Laurent polynomial matrix. Furthermore k can be chosen $k = \max(n_0, r)$.

Proof. The proof follows the lines of the theorem. The following modifications are made.

We prove the corollary for $r \leq n_0$.

n_0 is substituted for n throughout the proof.

The step (4.2.62) is not necessary since the identity matrix on the left side vanishes.

The proof is finished after (4.2.65).

□

Theorem 4.2.5. Two equivalent systems have the same order.

Proof. The order n is given by $n = n_0 + n_D$, where n_0 is the dimension of X_0 (Def. 3.2.4) and n_D is determined by the transfer function (see Theorem 3.2.3). It follows by Theorem 4.2.2 that n_D is invariant under equivalence.

Using the definition of equivalence (Def. 4.1.1) and putting $u = 0$ it follows that there is a bijective mapping

$$\xi_1 = Z(q)\xi \tag{4.2.73}$$

from X_0 to X_0^1. The sets X_0 and X_0^1 are vector spaces and the mapping (4.2.73) is linear. Therefore (4.2.73) is a vector space isomorfism between X_0 and X_0^1. It follows that $\dim X_0 = \dim X_0^1$. This means that n_0 is invariant under equivalence.

□

Corollary. The numbers n_0 and n_D, defined in Def. 3.2.5, are equal for equivalent systems.

We are now in a position to prove the converse of Theorem 4.2.1.

Theorem 4.2.6. If two systems are equivalent then the corresponding two system matrices are equivalent. Furthermore, in Definition 4.2.2, k can be chosen as

$$k = k_0 \triangleq \max(r, r_1, n_0) \tag{4.2.74}$$

where n_0 is given by Def. 3.2.5.

Proof. Let

$$T(q)\xi = U(q)u \tag{4.2.75a}$$

$$y = V(q)\xi + W(q)u \tag{4.2.75b}$$

and

$$T_1(q)\xi_1 = U(q)u_1 \tag{4.2.76a}$$

$$y_1 = V_1(q)\xi_1 + W_1(q)u_1 \tag{4.2.76b}$$

be the two given systems and let $P(q)$ and $P_1(q)$ be their system matrices.

By the corollary of Theorem 4.2.4 there is a system

$$qx = Ax + Bu \tag{4.2.77a}$$

$$y = Cx + E(q)u \tag{4.2.77b}$$

where A is an $n_0 \times n_0$ nonsingular matrix and $E(q)$ is a Laurent polynomial matrix, such that its system matrix $\bar{P}(q)$ is equivalent to $P(q)$ with $k = \max(r, n_0)$ and clearly also with $k = k_0$.

Similarly there is a system

$$qx_1 = A_1 x_1 + Bu_1 \tag{4.2.78a}$$

$$y_1 = C_1 x_1 + E_1(q)u_1 \tag{4.2.78b}$$

with a system matrix $\bar{P}_1(q)$, which is equivalent to $P_1(q)$ with $k = \max(r_1, n_0)$ and therefore also with $k = k_0$ and A_1 is an $n_0 \times n_0$ nonsingular matrix and $E_1(q)$ is a Laurent polynomial matrix.

By Theorem 4.2.1 the systems (4.2.75) and (4.2.77) are equivalent. Analogously the systems (4.2.76) and (4.2.78) are equivalent. Consequently the systems (4.2.77) and (4.2.78) are equivalent. By the corollary of Theorem 4.2.3 they are system similar. It follows by Remark 1 of Theorem 4.2.3 that $\bar{P}(q)$ and $\bar{P}_1(q)$ are equivalent with $k = n_0$ and therefore with $k = k_0$. Consequently $P(q)$ and $P_1(q)$ are equivalent with $k = k_0$.

□

<u>Theorem 4.2.7</u>. Two system matrices are equivalent with $k \geq k_0 \triangleq \max(r, r_1, n_0)$ if and only if they are equivalent with $k = k_0$.

<u>Proof</u>. If the system matrices are equivalent for $k = k_0$ then clearly they are equivalent with $k \geq k_0$.

Conversely if the system matrices are equivalent with $k \geq k_0$ then by Theorem 4.2.1 the corresponding systems are equivalent and by Theorem 4.2.6 the system matrices are equivalent with $k = k_0$.

□

The corresponding result for strict system equivalence, see Rosenbrock (1970), is shown in Pernebo (1977) and Rosenbrock (1977).

Example 4.2.1. Consider the two system matrices

$$\left(\begin{array}{c|c} (q-1)(q-2) & q^{-1} \\ \hline -(1+q^{-1}) & 0 \end{array}\right) \tag{4.2.79}$$

and

$$\left(\begin{array}{c|c} q^2(q-1)(q-2) & (q+1) \\ \hline -1 & 0 \end{array}\right) \tag{4.2.80}$$

The two corresponding systems were shown to be equivalent in Example 4.1.1. By Theorem 4.2.6 it follows that the two system matrices are equivalent. Furthermore it follows that k in Def. 4.2.2 can be chosen as $k = \max(r, r_1, n_0) = 2$ since $n_0 = 2$. We find that

$$\left(\begin{array}{cc|c} -q^2+4q & 1 & 0 \\ -q^4+3q^3-2q^2 & q^2+q & 0 \\ \hline 1 & 0 & 1 \end{array}\right)\left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & (q-1)(q-2) & q^{-1} \\ \hline 0 & -(1+q^{-1}) & 0 \end{array}\right) \cdot$$

$$\cdot \left(\begin{array}{cc|c} \frac{1}{6}(1+q^{-1}) & \frac{1}{6}(-q^2+3q-2) & \frac{1}{6}(-q^{-1}-q^{-2}) \\ \frac{1}{6} & \frac{1}{6}(-q^2+4q) & -\frac{1}{6}q^{-1} \\ \hline 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & q^2(q-1)(q-2) & q+1 \\ \hline 0 & -1 & 0 \end{array}\right)$$

which verifies that the two system matrices are equivalent and k can be chosen $k = 2$.

Suppose we can choose $k = 1$. This means that there are Laurent polynomials $m(q)$, $n(q)$, $x(q)$ and $y(q)$ with $m(q) = aq^i$ and $n(q) = bq^j$, where $a,b \in R\setminus\{0\}$ and $i,j \in Z$, such that

$$\left(\begin{array}{c|c} m(q) & 0 \\ \hline x(q) & 1 \end{array}\right)\left(\begin{array}{c|c} (q-1)(q-2) & q^{-1} \\ \hline -(1+q^{-1}) & 0 \end{array}\right)\left(\begin{array}{c|c} n(q) & y(q) \\ \hline 0 & 1 \end{array}\right) = \left(\begin{array}{c|c} q^2(q-1)(q-2) & q+1 \\ \hline -1 & 0 \end{array}\right) \qquad (4.2.81)$$

Taking the $(1,2)$ element of (4.2.81) gives

$$m(q)[(q^2-3q+2)y(q) + q^{-1}] = q + 1 \qquad (4.2.82)$$

which cannot be satisfied. Therefore it is not possible to choose $k = 1$.

□

4.3 Invariants under equivalence

In the previous section we showed that equivalence between systems of the type

$$T(q)\xi = U(q)u \qquad (4.3.1a)$$

$$y = V(q)\xi + W(q)u \qquad (4.3.1b)$$

is the same as equivalence between the corresponding system matrices with k in Def. 4.2.2 equal to $\max(r,r_1,n_o)$. In practice it is often more convenient to work with system matrices, while the concept of equivalence for systems (Def. 4.1.1) is intuitively more satisfactory. It was shown in Section 4.2 that the transfer function and the order of the system are invariant under equivalence. This is very easily verified using equivalence for system matrices. In this section we will show that all the other concepts defined in Chapter 3 such as stability, causality, controllability and observability are invariant under equivalence.

Theorem 4.3.1. Stability is invariant under equivalence.

Proof. Suppose that the system matrices

$$P(q) = \left(\begin{array}{c|c} T(q) & U(q) \\ \hline -V(q) & W(q) \end{array}\right) \tag{4.3.2}$$

and

$$P_1(q) = \left(\begin{array}{c|c} T_1(q) & U_1(q) \\ \hline -V_1(q) & W_1(q) \end{array}\right) \tag{4.3.3}$$

are equivalent. Then by (4.2.5)

$$\begin{pmatrix} I_{k-r_1} & 0 \\ 0 & T_1(q) \end{pmatrix} = M(q) \begin{pmatrix} I_{k-r} & 0 \\ 0 & T(q) \end{pmatrix} N(q) \tag{4.3.4}$$

It follows that

$$\det T_1(q) = \det M(q) \det T(q) \det N(q) \tag{4.3.5}$$

and the zeros of the Laurent polynomials $\det T(q)$ and $\det T_1(q)$ are the same. The theorem now follows from Theorem 3.3.1.

□

Theorem 4.3.2. Causality is invariant under equivalence.

Proof. Causality is by Def. 3.4.1 defined from the transfer function and this is by Theorem 4.2.2 invariant under equivalence.

□

Theorem 4.3.3. The following are invariant under equivalence.

(i) the set of i.d. zeros
(ii) the set of o.d. zeros
(iii) the set of i.o.d. zeros.

Proof. The theorem is proved for polynomial system matrices in Rosenbrock (1970). Only minor changes need to be made.

□

Corollary 1. The property of controllability is invariant under equivalence.

Proof. A system is controllable if and only if it has no i.d. zeros.

□

Corollary 2. The property of observability is invariant under equivalence.

4.4 Special forms of the system matrix

In this section we will show that a system matrix can by equivalence transformations be brought into a form, which is more easy to deal with than a general system matrix. In section 4.2 we have already shown that any system matrix can be brought to an equivalent state space form.

We will consider a system

$$T(q)\xi = U(q)u \tag{4.4.1a}$$

$$y = V(q)\xi + W(q)u \tag{4.4.1b}$$

with system matrix $P(q)$.

Theorem 4.4.1. Let $P(q)$ be an arbitrary system matrix. Then there is an equivalent system matrix

$$P_1(q) = \left(\begin{array}{c|c} T_1(q) & U_1(q) \\ \hline -V_1(q) & W_1(q) \end{array}\right) \tag{4.4.2}$$

in polynomial form with the matrices $(T_1(0) \quad U_1(0))$ and $(T_1^T(0) \quad V_1^T(0))^T$ having full rank.

Proof. The state space form of Theorem 4.2.4 has the desired properties.

□

Theorem 4.4.2. Let $P(q)$ be an arbitrary system matrix. Then there is an equivalent system matrix

$$P^*(q^{-1}) = \left(\begin{array}{c|c} T^*(q^{-1}) & U^*(q^{-1}) \\ \hline -V^*(q^{-1}) & W^*(q^{-1}) \end{array}\right) \tag{4.4.3}$$

which is a <u>polynomial matrix in q^{-1}</u>. Furthermore the matrices $\left(T^*(0) \quad U^*(0)\right)$ and $\left(T^{*T}(0) \quad V^{*T}(0)\right)^T$ both have full rank.

<u>Proof</u>. Regard the system matrix in Theorem 4.2.4 $P(q)$ as a Laurent polynomial matrix in q^{-1}. Going through the proof this will lead to a final system matrix of the form

$$P_1(q) = \left(\begin{array}{c|c} q^{-1}I - A & B \\ \hline -C & D(q^{-1}) \end{array}\right)$$

where $D(q^{-1})$ is a polynomial matrix in q^{-1} and the matrices $(A \quad B)$ and $(A^T \quad C^T)^T$ have both full rank.

<u>Def. 4.4.1</u>. We will say that the system matrix $P(x)$ is in <u>standard polynomial form in x</u> if $P(x)$ is a polynomial matrix and the matrices $\left(T(0) \quad U(0)\right)$ and $\left(T^T(0) \quad V^T(0)\right)^T$ both have full rank. Here x can be q or q^{-1}.

<u>Example 4.4.1</u>. Consider the system matrix

$$P(q) = \left(\begin{array}{cc|c} 1 & q^{-1} & q^{-1} \\ q & q^{-1} & q^{-1} \\ \hline -1 & 0 & 0 \\ 0 & -1 & 0 \end{array}\right)$$

Multiplying from the left by $\mathrm{diag}(q,q,1,1)$ gives

$$P_1(q) = \left(\begin{array}{cc|c} q & 1 & 1 \\ q^2 & 1 & 1 \\ \hline -1 & 0 & 0 \\ 0 & -1 & 0 \end{array}\right)$$

$P_1(q)$ is not in standard polynomial form in q since

$$\begin{pmatrix} T(0) & U(0) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

has linearly dependent rows. It can be brought to standard polynomial form in q by subtracting row 1 of $P_1(q)$ from row 2 and multiplying the resulting row by q^{-1}.

$$P_2(q) \triangleq \left(\begin{array}{cc|c} q & 1 & 1 \\ q-1 & 0 & 0 \\ \hline -1 & 0 & 0 \\ 0 & -1 & 0 \end{array}\right) = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ -q^{-1} & q^{-1} & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) P_1(q)$$

$P_2(q)$ is in standard polynomial form in q.

□

The transformation of strict system equivalence (s.s.e) is defined by Rosenbrock (1970). It is a special case of our transformations of equivalence for Laurent polynomial system matrices. Therefore, and because of Theorem 4.4.1, the standard forms under s.s.e. are also standard forms under equivalence for system matrices.

<u>Theorem 4.4.3</u>. Let $P(q)$ be the system matrix of an observable system. Then there is an equivalent system matrix $P_1(q)$ in standard polynomial form in q of the form

$$P_1(q) = \left(\begin{array}{c|c} T_1(q) & U_1(q) \\ \hline -I & 0 \end{array}\right) \tag{4.4.4}$$

<u>Proof</u>. If the system is observable then it has no o.d. zeros. By Theorem 4.4.1 it is equivalent to a system matrix $P_2(q)$ in standard polynomial form in q. This system matrix has no o.d. zeros by Theorem 4.3.3. This, together with

the fact that $\left(T_2^T(0) \quad V_2^T(0)\right)$ has full rank ensures that the system has no o.d. zeros in the sense of Rosenbrock (see the remark after Def. 3.6.3). By Rosenbrock (1970), Corollary 1 of Theorem 3.2.1, it is s.s.e. to

$$P_3(q) = \left(\begin{array}{c|c} T_3(q) & U_3(q) \\ \hline -I & W_3(q) \end{array}\right)$$

but

$$P_1(q) = \left(\begin{array}{c|c} T_1(q) & U_1(q) \\ \hline -I & 0 \end{array}\right) \triangleq P_3(q)\left(\begin{array}{c|c} I & W_3(q) \\ \hline 0 & I \end{array}\right)$$

□

The following theorem was shown for polynomial system matrices and s.s.e. by Rowe (1971). Our proof will follow the lines of Rowe.

Theorem 4.4.4. Let $P(q)$ and $P_1(q)$ be Laurent polynomial system matrices of the form

$$P(q) = \left(\begin{array}{c|c} T(q) & U(q) \\ \hline -I & 0 \end{array}\right)$$

and

$$P_1(q) = \left(\begin{array}{c|c} T_1(q) & U_1(q) \\ \hline -I & 0 \end{array}\right)$$

$P(q)$ and $P_1(q)$ are equivalent if and only if there is a unimodular Laurent polynomial matrix $Q(q)$ such that

$$T_1(q) = Q(q)T(q) \tag{4.4.5a}$$

$$U_1(q) = Q(q)U(q) \tag{4.4.5b}$$

Proof. If (4.4.5) is true then $P(q)$ and $P_1(q)$ clearly are equivalent.

To prove the converse let $L(q)$ and $L_1(q)$ be such that

$$\left(T(q) \quad U(q)\right) = L(q)\left(\bar{T}(q) \quad \bar{U}(q)\right) \tag{4.4.6}$$

and

$$\left(T_1(q) \quad U_1(q)\right) = L_1(q)\left(\bar{T}_1(q) \quad \bar{U}_1(q)\right) \tag{4.4.7}$$

where $\bar{T}(q)$ and $\bar{U}(q)$ are relatively left prime and so are $\bar{T}_1(q)$ and $\bar{U}_1(q)$. $P(q)$ and $P_1(q)$ give the same transfer function

$$\bar{T}^{-1}\bar{U}(q) = T^{-1}(q)U(q) = T_1^{-1}(q)U_1(q) = \bar{T}_1^{-1}(q)\bar{U}_1(q) \tag{4.4.8}$$

Lemma 3.5.2 shows that there is a unimodular $R(q)$ such that

$$\bar{T}_1(q) = R(q)\bar{T}(q) \tag{4.4.9}$$

$$\bar{U}_1(q) = R(q)\bar{U}(q) \tag{4.4.10}$$

Now by (4.4.6), (4.4.7), (4.4.9) and (4.4.10)

$$(T_1 \quad U_1) = L_1(\bar{T}_1 \quad \bar{U}_1) = L_1R(\bar{T} \quad \bar{U}) = L_1RL^{-1}(T \quad U)$$

We have to show that L_1RL^{-1} is a unimodular Laurent polynomial matrix.

Since $P(q)$ and $P_1(q)$ are equivalent there are $M(q)$, $N(q)$, $X(q)$ and $Y(q)$ such that

$$\left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & L_1R\bar{T} & L_1R\bar{U} \\ \hline 0 & -I & 0 \end{array}\right) = \begin{pmatrix} M & 0 \\ X & I \end{pmatrix}\left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & L\bar{T} & L\bar{U} \\ \hline 0 & -I & 0 \end{array}\right)\begin{pmatrix} N & Y \\ 0 & I \end{pmatrix} \tag{4.4.11}$$

where we have omitted the argument q. This gives

$$\begin{pmatrix} I & 0 \\ 0 & L_1R\bar{T} \end{pmatrix} = M\begin{pmatrix} I & 0 \\ 0 & L\bar{T} \end{pmatrix}N \tag{4.4.12}$$

and

$$\left\{X\begin{pmatrix} I & 0 \\ 0 & L\bar{T} \end{pmatrix} + (0 \quad -I)\right\}N = (0 \quad -I) \tag{4.4.13}$$

Solving (4.4.12) for N and substituting into (4.4.13) gives

$$\left\{X\begin{pmatrix} I & 0 \\ 0 & L\bar{T} \end{pmatrix} + (0 \quad -I)\right\}\begin{pmatrix} I & 0 \\ 0 & \bar{T}^{-1}L^{-1} \end{pmatrix}M^{-1}\begin{pmatrix} I & 0 \\ 0 & L_1R\bar{T} \end{pmatrix} = (0 \quad -I)$$

$$\Leftrightarrow \quad [X + (0 \quad -\bar{T}^{-1}L^{-1})] = (0 \quad -I)\begin{pmatrix} I & 0 \\ 0 & \bar{T}^{-1}R^{-1}L_1^{-1} \end{pmatrix}M$$

$$\Leftrightarrow \quad [X + (0 \quad -\bar{T}^{-1}L^{-1})] = \bar{T}^{-1}R^{-1}L_1^{-1}(0 \quad -I)M$$

$$\Leftrightarrow \quad L_1R\bar{T}X + (0 \quad -L_1RL^{-1}) = (0 \quad -I)M \tag{4.4.14}$$

It follows that L_1RL^{-1} is a Laurent polynomial matrix. Taking determinants of (4.4.12) gives

$$\det L_1R\bar{T} = \det M \det L\bar{T} \det N$$

$$\Leftrightarrow \quad \det L_1R = cq^k \det L$$

$$\Leftrightarrow \quad \det L_1RL^{-1} = cq^k \tag{4.4.15}$$

Therefore L_1RL^{-1} is unimodular and the theorem is proved.

□

Theorems 4.4.3 and 4.4.4 can of course be formulated for controllable systems and system matrices of the type

$$P(q) = \left(\begin{array}{c|c} T(q) & I \\ \hline -V(q) & 0 \end{array}\right)$$

as well.

<u>Theorem 4.4.5</u>. Let $P(q)$ and $P_1(q)$ be two system matrices in standard polynomial form in q. Then $P(q)$ and $P_1(q)$ are equivalent if and only if they are s.s.e.

<u>Proof</u>. If $P(q)$ and $P_1(q)$ are s.s.e then clearly they are equivalent.

Suppose that $P(q)$ and $P_1(q)$ are equivalent.

It is shown by Rosenbrock (1970) that $P(q)$ and $P_1(q)$ are s.s.e to some system matrices

$$\bar{P}(q) = \left(\begin{array}{c|c} qI-A & B \\ \hline -C & D(q) \end{array}\right)$$

and

$$\bar{P}_1(q) = \left(\begin{array}{c|c} qI-A_1 & B_1 \\ \hline -C & D_1(q) \end{array}\right)$$

and that the Smith form of the <u>polynomial</u> matrices $(T(q)\ \ U(q))$ and $(qI-A\ \ B)$ are the same. Analogously the Smith form of the polynomial matrices $(T^T(q)\ \ V^T(q))$ and $(qI-A^T\ \ C^T)$ are the same. Therefore the conditions that $(T(0)\ \ U(0))$ and $(T^T(0)\ \ V^T(0))$ have full rank imply that $(A\ \ B)$ and $(A^T\ \ C^T)$ have full rank. Furthermore

$\bar{P}(q)$ and $\bar{P}_1(q)$ are equivalent because $P(q)$ and $P_1(q)$ are equivalent. By Theorem 4.2.3 $\bar{P}(q)$ and $\bar{P}_1(q)$ are s.s. and therefore s.s.e. It follows that $P(q)$ and $P_1(q)$ are s.s.e.

□

Theorem 4.4.6. Let

$$P(q) = \left(\begin{array}{c|c} T(q) & U(q) \\ \hline -I & 0 \end{array}\right)$$

and

$$P_1(q) = \left(\begin{array}{c|c} T_1(q) & U_1(q) \\ \hline -I & 0 \end{array}\right)$$

be system matrices in standard polynomial form in q. They are equivalent if and only if there is a unimodular polynomial matrix $Q(q)$ such that

$$T_1(q) = Q(q)T(q) \tag{4.4.16a}$$

$$U_1(q) = Q(q)U(q) \tag{4.4.16b}$$

Proof. If (4.4.16) is true then clearly $P(q)$ and $P_1(q)$ are equivalent.

If $P(q)$ and $P_1(q)$ are equivalent then they are s.s.e. by Theorem 4.4.5. The result then follows from Rowe (1971).

□

Theorem 4.4.7. Let $P(q)$ and $P_1(q)$ be two system matrices without decoupling zeros. Then $P(q)$ and $P_1(q)$ are equivalent if and only if they have the same transfer function.

Proof. If $P(q)$ and $P_1(q)$ are equivalent then they have the same transfer function by Theorem 4.2.2.

Conversely suppose that $P(q)$ and $P_1(q)$ have the same transfer function. By Theorem 4.4.1 there are system matrices $\bar{P}(q)$ and $\bar{P}_1(q)$ in standard polynomial form in q equivalent to $P(q)$ and $P_1(q)$ respectively. $\bar{P}(q)$ and $\bar{P}_1(q)$ have by Theorem 4.2.2 the same transfer function. Furthermore they have no decoupling zeros by Theorem 4.3.3. This together with the fact that they are in standard polynomial form in q implies that they have no decoupling zeros in the sense of Rosenbrock (see the remark after Def. 3.6.3). It is then shown in Rosenbrock (1970) that they are s.s.e. Therefore $P(q)$ and $P_1(q)$ are equivalent.

□

Example 4.4.2 (Rowe (1971)). Let

$$P(q) = \left(\begin{array}{cc|cc} q+1 & 0 & 1 & 1 \\ 0 & q+2 & 1 & 2 \\ \hline -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right)$$

and

$$P_1(q) = \left(\begin{array}{cc|cc} \frac{1}{2}(q+1)(q+3) & -\frac{1}{2}(q+1)(q+2) & 1 & -\frac{1}{2}(q-1) \\ -\frac{1}{2}(q+1)(q+2) & \frac{1}{2}q(q+2) & 1 & \frac{1}{2}(q-2) \\ \hline -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right)$$

If

$$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad N = \frac{1}{2}\begin{pmatrix} q+3 & -(q+2) \\ q+1 & -q \end{pmatrix}$$

$$X = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \qquad Y = \begin{pmatrix} 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} \end{pmatrix}$$

then

$$\left(\begin{array}{c|c} M & 0 \\ \hline X & I \end{array}\right) P(q) \left(\begin{array}{c|c} N & Y \\ \hline 0 & I \end{array}\right) = P_1(q)$$

By Theorem 4.4.6 there is a unimodular polynomial matrix $Q(q)$ such that (4.4.16) is satisfied. In this case $Q(q)$ can be chosen as

$$Q(q) = \frac{1}{2}\begin{pmatrix} q+3 & -(q+1) \\ -(q+2) & q \end{pmatrix}$$

□

We have shown that if the system matrices take some special forms then also the equivalence transformations can without loss of generality be taken from a smaller class of transformations. This is summarized below.

Theorem 4.4.5 shows that for system matrices in standard polynomial form in q the systems are equivalent if and only if the system matrices are s.s.e. If we further specialize the system matrices to be in a state space form which also is in standard polynomial form in q then the systems are equivalent if and only if the system matrices are s.s. by Theorem 4.2.3. A specialization in another direction is given by Theorem 4.4.4. However, unlike in the two previous cases, not every system matrix can be brought to the desired form, but only observable ones. If the system matrices in Theorem 4.4.4 are in standard polynomial form in q then $Q(q)$ can be chosen as a unimodular polynomial matrix as shown by Theorem 4.4.6.

4.5. Causal systems in standard polynomial form

In this section we will specialize to systems in standard polynomial form in q or q^{-1}. First we will give criteria for causality for these systems. Then we will further specialize to causal systems and present alternative methods for computing the order of such a system.

Consider the system matrix

$$P(q) = \left(\begin{array}{c|c} T(q) & U(q) \\ \hline -V(q) & W(q) \end{array}\right)$$

where $T(q)$ has dimension r. Let $P\binom{i_1 \ldots i_p}{j_1 \ldots j_p}$ denote the minor of $P(q)$ formed by row $1,2,\ldots,r,r+i_1,\ldots,r+i_p$ and column $1,2,\ldots,r,r+j_1,\ldots,r+j_p$, where $0 \leq p \leq \min(\ell,m)$. This notation is introduced in Rosenbrock (1970) and the following definition can be found in Rosenbrock (1974).

Definition 4.5.1. Let $P(x)$ be a polynomial system matrix in x, where x is q or q^{-1}. The degree of $P(x)$ is $\max\left(\deg P\binom{i_1 \ldots i_p}{j_1 \ldots j_p}\right)$ where the maximum is taken over all nonzero $P\binom{i_1 \ldots i_p}{j_1 \ldots j_p}$, $0 \leq p \leq \min(\ell,m)$.

Theorem 4.5.1. Let $P(q)$ be a polynomial system matrix in q. Then the system is causal if and only if the degree of $P(q)$ is equal to $\deg \det T(q)$.

Proof. It is easy to see that

$$G^{i_1 \ldots i_p}_{j_1 \ldots j_p} = \frac{P^{i_1 \ldots i_p}_{j_1 \ldots j_p})}{\det T} \tag{4.5.1}$$

(see Rosenbrock (1970)). Here $G^{i_1 \ldots i_p}_{j_1 \ldots j_p}$ denotes the minor of the transfer function $G(q)$ that is formed by rows $i_1, \ldots, i_p$ and columns $j_1, \ldots, j_p$.

Suppose that the system is causal. Every element of $G(q)$ is then by definition proper. This implies that every minor of $G(q)$ is proper. By (4.5.1) it follows that the degree of $P(q)$ is less than or equal to $\deg \det T(q)$.

Conversely suppose that the degree of $P(q)$ is less than or equal to $\deg \det T(q)$. In particular it is true that $\deg P^{i_1}_{j_1}) \leq \deg \det T$. By (4.5.1) $G^{i_1}_{j_1}$ is proper for all $1 \leq i_1 \leq m$, $1 \leq j_1 \leq \ell$. But $G^{i_1}_{j_1}$ is element (i_1, j_1) of $G(q)$ and the theorem is proved.

□

Lemma 4.5.1. Let $P(x)$ be a system matrix in standard polynomial form in x. Then there is a minor $P^{i_1 \ldots i_p}_{j_1 \ldots j_p})$ which is nonzero for $x = 0$.

Proof. Let $T(x)$ be of dimension r and put $k = \text{rank } T(0)$. By assumption $\text{rank}\big(T(0) \;\; U(0)\big) = r$. Therefore there are $r-k$ columns $U_{j_1}, \ldots, U_{j_{r-k}}$ of $U(0)$ such that $\text{rank}\big(T(0) \;\; U_{j_1} \ldots U_{j_{r-k}}\big) = r$. In the same way there are $r-k$ rows $V_{i_1}, \ldots, V_{i_{r-k}}$ of $-V(0)$ such that $\text{rank}\big(T^T(0) \;\; V^T_{i_1} \ldots V^T_{i_{r-k}}\big)^T = r$. It follows that $P^{i_1 \ldots i_{r-k}}_{j_1 \ldots j_{r-k}}) \neq 0$ for $x = 0$.

□

Theorem 4.5.2. Let $P^*(q^{-1})$ be a system matrix in standard polynomial form in q^{-1}. Then the system is causal if and only if $\det T^*(0) \neq 0$.

Proof. Let $P(q)$ be an equivalent system matrix in standard polynomial form in q. Then there are Laurent polynomial matrices $M(q)$, $N(q)$, $X(q)$ and $Y(q)$, where $M(q)$ and $N(q)$ are unimodular, such that

$$\left(\begin{array}{cc|c} I_{k_0-r^*} & 0 & 0 \\ 0 & T^*(q^{-1}) & U^*(q^{-1}) \\ \hline 0 & -V^*(q^{-1}) & W^*(q^{-1}) \end{array}\right) = \left(\begin{array}{c|c} M(q) & 0 \\ \hline X(q) & I \end{array}\right) \left(\begin{array}{cc|c} I_{k_0-r} & 0 & 0 \\ 0 & T(q) & U(q) \\ \hline 0 & -V(q) & W(q) \end{array}\right) \left(\begin{array}{c|c} N(q) & Y(q) \\ \hline 0 & I \end{array}\right) \tag{4.5.2}$$

Consider a minor $P\binom{i_1 \ldots i_p}{j_1 \ldots j_p}$ of $P(q)$. If

$$\bar{P}(q) = \begin{pmatrix} I_{k_0-r} & 0 \\ 0 & P(q) \end{pmatrix}$$

then $\bar{P}\binom{i_1 \ldots i_p}{j_1 \ldots j_p} = P\binom{i_1 \ldots i_p}{j_1 \ldots j_p} \cdot \det I_{k_0-r} = P\binom{i_1 \ldots i_p}{j_1 \ldots j_p}$. The same is true for $P^*(q^{-1})$. Furthermore it is easily shown that $P^*\binom{i_1 \ldots i_p}{j_1 \ldots j_p}$ is independent of $X(q)$ and $Y(q)$. Define k_M and k_N through $\det M(q) = aq^{-k_M}$ and $\det N(q) = bq^{-k_N}$. Then we have

$$P^*\binom{i_1 \ldots i_p}{j_1 \ldots j_p}(q^{-1}) = abq^{-(k_M+k_N)} P\binom{i_1 \ldots i_p}{j_1 \ldots j_p}(q) \tag{4.5.3}$$

Put $n_1 = \deg \det T(q)$ and n_2 as the degree of $P(q)$. Then $k_M + k_N \geq n_2$ because every minor $P^*\binom{i_1 \ldots i_p}{j_1 \ldots j_p}(q^{-1})$ is a polynomial in q^{-1}. Suppose $k_M + k_N > n_2$. Then every minor $P^*\binom{i_1 \ldots i_p}{j_1 \ldots j_p}(q^{-1})$ has a factor q^{-1}. By Lemma 4.5.1 this is not possible. Therefore $k_M + k_N = n_2$. Now from (4.5.3) with index $p = 0$ we have the equivalence

$$\det T^*(0) \neq 0 \iff k_M + k_N = n_1$$

or since $k_M + k_N = n_2$

$$\det T^*(0) \neq 0 \iff n_1 = n_2$$

But $n_1 = n_2$ if and only if the system is causal by Theorem 4.5.1.

□

We will now turn to the problem of calculating the order of a system (Def. 3.2.5).

Theorem 4.5.3. Let $P(q)$ be a system matrix in standard polynomial form in q. The order n of the corresponding system is given by $n = \deg \det T(q)$.

Proof. It follows from Theorem 4.2.4 that there is an equivalent system matrix $\bar{P}(q)$ in state space form in standard polynomial form in q. Furthermore the A-matrix of this state space form is an $n \times n$ matrix, where n is the order of the system. Therefore $n = \deg \det(qI-A)$.

By Theorem 4.4.5 $P(q)$ and $\bar{P}(q)$ are s.s.e. Therefore there are unimodular polynomial matrices $M(q)$ and $N(q)$ such that $(qI-A) = M(q)T(q)N(q)$. Taking determinants gives $\det(qI-A) = c \det T(q)$ and $\deg \det T(q) =$ $= \deg \det(qI-A) = n$.

□

Theorem 4.5.4. Let $P^*(q^{-1})$ be a causal system matrix in standard polynomial form in q^{-1}. The order n of the system is equal to the degree of $P^*(q^{-1})$.

Proof. Let $P(q)$ be an equivalent system matrix in standard polynomial form in q. Then (4.5.2) is true and it follows as in the proof of Theorem 4.5.2 that

$$P^*\binom{i_1 \dots i_p}{j_1 \dots j_p}(q^{-1}) = abq^{-(k_M+k_N)} P\binom{i_1 \dots i_p}{j_1 \dots j_p}(q) \tag{4.5.4}$$

Here $k_M + k_N = \deg \det T(q) = n$, where the first equality follows because the system is causal and the second by Theorem 4.5.3. Lemma 4.5.1 applied to $P(q)$ shows that there is a minor $P^*\binom{i_1 \dots i_p}{j_1 \dots j_p}(q^{-1})$ of $P^*(q^{-1})$ of degree n. No minor is of higher degree. Therefore the degree of the system matrix $P^*(q^{-1})$ is n.

□

Consider a system matrix of the form

$$P(x) = \left(\begin{array}{c|c} T(x) & U(x) \\ \hline -I & 0 \end{array}\right) \tag{4.5.5}$$

where $T(x)$ is of dimension r. The minor $P\binom{i_1 \dots i_p}{j_1 \dots j_p}(x)$ is here, after possibly a sign change, equal to the determinant of the matrix formed by columns $j_1, \dots, j_p$ of $U(x)$ and the columns of $T(x)$ that are left when columns $i_1, \dots, i_p$ are deleted. Therefore the degree of $P(x)$ can be computed in the following way in this case. Compute the determinants of all matrices that can be obtained by choosing r columns from the matrix $(T(x) \;\; U(x))$. The highest degree of these determinants is equal to the degree of $P(x)$.

Theorem 4.5.4 can be reformulated for systems of the type (4.5.5).

<u>Theorem 4.5.5</u>. Consider the causal system

$$A(q^{-1})y = B(q^{-1})u \tag{4.5.6}$$

where $A(q^{-1})$ and $B(q^{-1})$ are polynomial matrices in q^{-1}.

The dimension of $A(q^{-1})$ is $m \times m$, $\det A(q^{-1}) \neq 0$ and the rows of $(A(0) \quad B(0))$ are linearly independent. Compute the determinants of all matrices that are obtained by choosing m columns of the matrix $(A(q^{-1}) \quad B(q^{-1}))$. These determinants are polynomials in q^{-1}. The order of the system is equal to the highest degree of these determinants.

Example 4.5.1. Consider the system of Example 3.2.4. The corresponding system matrix is

$$P(q) = \left(\begin{array}{cc|cc} q^2-1 & q-1+q^{-1}-q^{-2} & q-1 & 0 \\ q^2 & 1+q^{-1} & 0 & 1 \\ \hline -1 & 0 & 0 & 0 \\ -1+q^{-1} & -1 & 0 & 0 \end{array}\right)$$

It can be brought to standard polynomial form in q by multiplying the first column by q and the second by q^2. This gives

$$P_1(q) = \left(\begin{array}{cc|cc} q^3-1 & q^3-q^2+q-1 & q-1 & 0 \\ q^3 & q^2+q & 0 & 1 \\ \hline -q & 0 & 0 & 0 \\ -q+1 & -q^2 & 0 & 0 \end{array}\right) \triangleq \left(\begin{array}{c|c} T_1(q) & U_1(q) \\ \hline -V_1(q) & 0 \end{array}\right)$$

Multiplications of the first and second rows and the first column of $P(q)$ by q^{-1} brings it to standard polynomial form in q^{-1}. The result is

$$P_2(q^{-1}) = \left(\begin{array}{cc|cc} 1-q^{-2} & 1-q^{-1}+q^{-2}-q^{-3} & 1-q^{-1} & 0 \\ 1 & q^{-1}+q^{-2} & 0 & q^{-1} \\ \hline -q^{-1} & 0 & 0 & 0 \\ -q^{-1}+q^{-2} & -1 & 0 & 0 \end{array}\right) \triangleq \left(\begin{array}{c|c} T_2(q^{-1}) & U_2(q^{-1}) \\ \hline -V_2(q^{-1}) & 0 \end{array}\right)$$

Theorem 4.5.1 can be used on $P_1(q)$ to determine if the system is causal or not. It is, however, tedious to check all relevant minors. It is easier to use Theorem 4.5.2 on $P_2(q)$. We find that

$$\det T_2(0) = \det \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = -1 \neq 0$$

and the system is causal.

Causality can also be determined from the transfer function. In Example 3.2.4 the transfer function $G(q)$ is calculated. Since all entries of $G(q)$ are proper the system is causal by Def. 3.4.1.

The order of the system can be calculated using Theorem 4.5.3.

$$n = \deg \det T_1(q) = 6$$

This result was also obtained in Example 3.2.4.

To use Theorem 4.5.4 to calculate n is tedious. We see, however, that it is not sufficient to check $\deg \det T_2(x) = 4$. We also see that $\deg \det P_2(x) = 3$, but

$$\deg P_2 \, {}^{2}_{2})(x) = \deg \det \begin{pmatrix} 1-x^2 & 1-x+x^2-x^3 & 0 \\ 1 & x+x^2 & x \\ -x+x^2 & -1 & 0 \end{pmatrix} = 6$$

□

4.6 Controllability indices

Suppose we have a controllable but not necessarily observable system of the type

$$T(q)\xi = U(q)u \tag{4.6.1a}$$

$$y = V(q)\xi + W(q)u \tag{4.6.1b}$$

In Def. 3.2.7 the state and the state space is defined for this system. Let $u(t) = 0$ for $t < 0$ and $u(0) = u_1 \neq 0$. Furthermore let the state at $t = 0$ be zero. In general the state will be nonzero for $t = 1$. Because the system is controllable it is possible to determine a sequence $\{u(t)\}_{t=1}^{\tau}$ such that the state is zero at $t = \tau + 1$ for some τ. Let τ_0 be the smallest τ-value that can be obtained with some input sequence for the given u_1. Now minimize τ_0 with respect to u_1 and call the minimal value λ_1. If a non-zero u_1 gives a state that is zero at $t = 1$ then $\lambda_1 = 0$. It is easy to see that all u_1 that minimize τ_0 form a linear subspace Λ_1 of R^{ℓ}. Define $d_1 = \dim \Lambda_1$ and put $\lambda_i = \lambda_1$ for $i = 2,\ldots,d_1$.

Now choose $u(0) = u_2 \notin \Lambda_1$ and minimize τ_0 with respect to u_2 as before. Call the minimal value λ_{d_1+1}. Clearly $\lambda_{d_1+1} > \lambda_1$. It is easy to see that all u_2' giving a τ_0 less than or equal to λ_{d_1+1} form a linear subspace $\Lambda_2 \supset \Lambda_1$. Put $d_2 = \dim \Lambda_2 > d_1$ and define $\lambda_i = \lambda_{d_1+1}$ for $i = d_1+2, \ldots, d_2$.

Continue in this way until $\Lambda_k = R^{\ell}$ for some k. This is possible because $\Lambda_{i+1} \supset \Lambda_i$ and $\dim \Lambda_{i+1} > \dim \Lambda_i$. We have now defined ℓ numbers $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{\ell}$.

Def. 4.6.1. The numbers $\lambda_1,\ldots,\lambda_{\ell}$ are the controllability indices of the controllable system (4.6.1)

Remark. Observe that $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_\ell$.

As we have seen the space R^ℓ, in which the input $u(t)$ takes its values, has a subspace structure $\Lambda_1 \subset \Lambda_2 \subset \dots \subset \subset \Lambda_k = R^\ell$ for some $k \leq \ell$. Suppose that a disturbance in form of a pulse in the input enters in Λ_1. Then the effect of this disturbance can be made to vanish in λ_1 time steps, but not in fewer. If the disturbance enters in Λ_2 but not in Λ_1, then the effect of it can be made to vanish in λ_{d_1+1} time steps but not in fewer, and so on. The effect of any disturbance in the input can be made to vanish in λ_ℓ time steps.

Theorem 4.6.1. The controllability indices are invariant under equivalence.

Proof. Let the system

$$T_1(q)\xi_1 = U_1(q)u \qquad (4.6.2a)$$

$$y_1 = V_1(q)\xi_1 + W_1(q)u \qquad (4.6.2b)$$

be equivalent to (4.6.1). Then by definition there is a bijective mapping.

$$\xi_1 = Z(q)\xi + Y(q)u \qquad (4.6.3)$$

which gives $y_1 = y$ for any input u.

When defining the controllability indices we suppose that $u(t) = 0$ for $t < 0$ and that the state is zero at $t = 0$. This implies that $\xi \in \overline{R_z^r}$ and ξ is uniquely determined by u. The mapping (4.6.3) then gives a ξ_1 in $\overline{R_z^{r1}}$. This ξ_1 is, by uniqueness, the solution to (4.6.2a) in $\overline{R_z^{r1}}$.

Consequently the solutions in $\overline{R_z^r}$ and $\overline{R_z^{r_1}}$ are related via (4.6.3).

We will show that if the solutions to the systems (4.6.1) and (4.6.2) are related via (4.6.3) then the state of (4.6.1) is zero if and only if the state of (4.6.2) is zero. We will show it for the states at $t = 1$. Because the systems are time invariant it will then be true for any t.

In order to determine the states at time $t = 1$ for the systems (4.6.1) and (4.6.2) we put $u_+ = 0$. Since $y_1 = y$ and y''_+, defined through Lemma 3.2.4, is uniquely give by y, the states at $t = 1$ for the two systems are (ξ', y''_+) and (ξ'_1, y''_+) respectively. Here ξ' and ξ'_1 are defined through Lemma 3.2.3. ξ' and ξ'_1 coincide, by definition, for large t with the solutions to (4.6.1a) and (4.6.2a) with $u_+ = 0$. Therefore for large t we have by (4.6.3)

$$\xi'_1 = Z(q)\xi' \tag{4.6.4}$$

Since the mapping (4.6.4) is bijective and the inverse is a mapping of the same type we have for large t

$$\xi'_1 = 0 \iff \xi' = 0 \tag{4.6.5}$$

By Lemma 3.2.2 it follows that (4.6.5) is true for all t. Therefore

$$(\xi', y''_+) = 0 \iff (\xi'_1, y''_+) = 0$$

It follows from the definition of controllability indices that they are the same for the two systems.

□

We will use a state space representation in standard polynomial form in q to calculate the controllability indices of a system. In the definition of controllability indices we have used $X_0 \times \overline{Y^{+}}$ as a state space. We will show that if the system is in state space form in standard polynomial form in q then this state space is isomorfic to what is usually regarded as the state space, i.e. the vector space R^n of n-vectors x(1).

Let

$$qx = Ax + Bu \tag{4.6.6a}$$

$$y = Cx + D(q)u \tag{4.6.6b}$$

be in standard polynomial form in q. By Theorem 4.2.4 the dimension of A is n, the order of the system. The system (4.6.6) is s.s. to a system

$$q\begin{pmatrix} \bar{x} \\ \tilde{x} \end{pmatrix} = \begin{pmatrix} \bar{A} & 0 \\ 0 & \tilde{A} \end{pmatrix}\begin{pmatrix} \bar{x} \\ \tilde{x} \end{pmatrix} + \begin{pmatrix} \bar{B} \\ \tilde{B} \end{pmatrix}u \tag{4.6.7a}$$

$$y = (\bar{C} \quad \tilde{C})\begin{pmatrix} \bar{x} \\ \tilde{x} \end{pmatrix} + D(q)u \tag{4.6.7b}$$

where all the eigenvalues of $\tilde{A}$ are zero and no eigenvalues of $\bar{A}$ are zero. Suppose $u_+ = 0$, then $\bar{x}$ is given by

$$\bar{x}(t) = \bar{A}^{t-1}\bar{x}(1) \qquad t \geq 1 \tag{4.6.8}$$

We see that $\bar{x}(1) \neq 0 \Rightarrow \bar{x}(t) \neq 0 \quad t \geq 1.$

$$\tilde{x}(t) = \tilde{A}^{t-1}\tilde{x}(1) \qquad t \geq 1 \tag{4.6.9}$$

and $\tilde{x}(t) = 0$ for $t > \dim \tilde{A}$.

It follows that the state $(x', y_+'') \in X_0 \times \overline{Y^{+}}$ is in this case given by

$$x'(t) = \begin{pmatrix} \bar{A}^{t-1}\bar{x}(1) \\ 0 \end{pmatrix} \qquad \forall t \qquad (4.6.10)$$

and

$$y_+''(t) = \tilde{C}\tilde{x}(t) \qquad t \geq 1 \qquad (4.6.11)$$

Therefore (4.6.8) - (4.6.11) define a mapping f from R^n to $X_0 \times \overline{Y^{+1}}$. This mapping is linear and the dimension of the range space is by definition equal to n. The dimension of the domain is also n, as pointed out earlier. Hence f is an isomorfism.

Observe that this is not true if the system is not in standard polynomial form in q. Then A will be of larger dimension than n.

For the system (4.6.6) form the matrix

$$[B \quad AB \quad A^2B \quad \ldots \quad A^{n-1}B] \qquad (4.6.12)$$

Choose columns from the left in the following way. Accept every column linearly independent of the previously accepted columns. Reject the others. Because the system is controllable and in standard polynomial form in q it is possible to choose n independent columns in this way. The chosen columns will be of the form

$$[b_1 \quad Ab_1 \ldots A^{\lambda_1'-1}b_1 \quad b_2 \quad Ab_2 \ldots A^{\lambda_2'-1}b_2 \ldots$$
$$\ldots \lambda_\ell'^{-1}b_\ell] \qquad (4.6.13)$$

for some integers $\lambda_1', \ldots, \lambda_\ell'$. Here $b_1, \ldots, b_\ell$ are the columns of B.

It is well known (see e.g. Rosenbrock (1970)) that the system (4.6.6) is s.s. to a system

$$
q\begin{pmatrix} x_1 \\ \vdots \\ x_{p_1} \\ \vdots \\ x_{p_2} \\ \vdots \\ \vdots \\ x_{p_\ell} \end{pmatrix}
=
\begin{pmatrix}
0 & 1 & & & & & & & \\
 & \ddots & \ddots & & & & & & \\
 & & 0 & 1 & & & & & \\
a_{11} & \cdots & a_{1p1} & \cdots & \cdots & \cdots & \cdots & \cdots & a_{1n} \\
 & & & 0 & 1 & & & & \\
 & & & & \ddots & \ddots & & & \\
 & & & & 0 & 1 & & & \\
a_{21} & \cdots & \cdots & a_{2p_1+1} & \cdots & a_{2p_2} & \cdots & \cdots & a_{2n} \\
 & & & & & & \ddots & & \\
 & & & & & & & 0 & 1 \\
 & & & & & & & \ddots & \ddots \\
 & & & & & & & 0 & 1 \\
a_{\ell 1} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & a_{\ell p_{\ell-1}+1} \cdots & a_{\ell p_\ell}
\end{pmatrix}
\begin{pmatrix} x_1 \\ \vdots \\ x_{p_1} \\ x_{p_1+1} \\ \vdots \\ x_{p2} \\ \vdots \\ x_{p_{\ell-1}+1} \\ \vdots \\ x_{p_\ell} \end{pmatrix}
+
$$

$$
+\begin{pmatrix}
1 & b_{12} & \cdots & b_{1\ell} \\
0 & 1 & b_{23} \cdots & b_{2\ell} \\
 & & & \\
0 & \cdots & 0 & 1
\end{pmatrix}
\begin{pmatrix} u_1 \\ \\ u_\ell \end{pmatrix}
\qquad (4.6.14)
$$

where

$$p_i = \sum_{j=1}^{i} \lambda_j'$$

All unmarked entries of the matrices are zero. The C-matrix has no special form.

Theorem 4.6.2. The set of controllability indices $\{\lambda_1 \ldots \lambda_\ell\}$ are equal to the set of numbers $\{\lambda_1' \ldots \lambda_\ell'\}$ not taken in order.

Proof. By Theorem 4.6.1 it is sufficient to show the theorem for the system (4.6.14) with $x(t)$ as the state at time t. Let the input be written $u(t) = H\bar{u}(t)$, where H is non-singular and choose H such that BH becomes as B in (4.6.14) but with all b_{ij} equal to zero.

Suppose that $x(0) = 0$ and that $\bar{u}_i(0)$, the i:th component of $\bar{u}(0)$, is equal to $a \neq 0$. This gives $x_{p_i}(1) = a$. Integrating the equations (4.1.14) we see that $x_{p_i-1}(2) = x_{p_i-2}(3) = \ldots = x_{p_{i-1}+1}(\lambda_i') = a$ independent of how $\{\bar{u}(t)\}_{t=1}^{\lambda_i'}$ is chosen.

It is possible to choose $\{\bar{u}(t)\}_{t=1}^{\lambda_i'}$ such that $x(\lambda_i'+1) = 0$. This is done by choosing $\bar{u}(t) = -Lx(t)$ for $t = 1,2,\ldots,\lambda_i'$, where L is given by

$$L = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{\ell 1} & \cdots & a_{\ell n} \end{pmatrix} \qquad (4.6.15)$$

This gives

$$A - BHL = \mathrm{diag}\left(\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}\right) \qquad (4.6.16)$$

and clearly $x(\lambda_i'+1) = 0$.

To determine the controllability index λ_1 we have to find the $\bar{u}(0) \neq 0$ that can give $x(\tau) = 0$ for the smallest possible $\tau > 0$. Let λ_{i_1}' be the smallest λ_i' and suppose there are d λ_i':s that are equal to λ_{i_1}'. Call the corresponding subscripts $i_2 \ldots i_d$. Now any $\bar{u}(0)$ with arbitrary values of $\bar{u}_{i_1}(0),\ldots,\bar{u}_{i_d}(0)$ will give $x(\lambda_{i_1}'+1) = 0$ if $\{\bar{u}(t)\}_{t=1}^{\lambda i_1}$ is chosen as $\bar{u}(t) = Lx(t)$, where L is given by (4.6.15). Furthermore it was shown above that it is not possible to make $x(\tau) = 0$ for $0 < \tau < \lambda_{i_1}'$ if $\bar{u}(0) \neq 0$. Therefore $\lambda_1 = \lambda_{i_1}'$. Since $\left(\bar{u}_{i_1}(0)\ldots\bar{u}_{i_d}(0)\right)^T$ span a d-dimensional space it follows that $\lambda_j = \lambda_{i_1}'$ for $j = 2, \ldots, d$.

Let λ_{k_1}' be the smallest λ_i' not equal to $\lambda_{i_1}' \ldots \lambda_{i_d}'$ and suppose $\lambda_{k_j}' = \lambda_{k_1}'$ for $j = 2,\ldots,e$. Then $\left(\bar{u}_{i_1}(0)\ldots\bar{u}_{i_d}(0), \bar{u}_{k_1}(0)\ldots\bar{u}_{k_e}(0)\right)^T \neq 0$ will give $x(\lambda_{k_1}') = 0$. Therefore $\lambda_{d+i} = \lambda_{k_1}'$ for $i = 1,\ldots,e$.

Continuing like this it follows that $\{\lambda_i'\}$ are the controllability indices.

□

The following corollary follows from Theorem 4.6.2.

<u>Corollary</u>

$$\sum_{i=1}^{\ell} \lambda_i = n$$

where n is the order of the system.

Def. 4.6.1. Let $A(x)$ be a polynomial matrix in x. The column degree of column i in $A(x)$ is equal to the highest power of x occuring in column i.

Def. 4.6.2. Let k_i be the column degree of column i in $A(x)$. $A(x)$ is column proper if the matrix

$$\lim_{x\to+\infty} A(x)\operatorname{diag}(x^{-k_1},\dots,x^{-k_p})$$

has linearly independent columns.

Theorem 4.6.3. Any controllable system is equivalent to a system

$$P_1(q) = \begin{pmatrix} T_1(q) & I \\ \hline -V_1(q) & 0 \end{pmatrix} \tag{4.6.17}$$

in standard polynomial form in q. Furthermore $T_1(q)$ is column proper with column degrees equal to the controllability indices (not in order).

Proof. This is shown as Theorem 3.2.1 in Rosenbrock (1970) by making operations of s.s.e. on the system (4.6.14). The details are omitted.

□

Lemma 4.6.1. Let $A(x)$ be an $m \times \ell$ polynomial matrix in x with $m \geq \ell$. Suppose $A(x)$ has full rank i.e. there is a nonzero $\ell \times \ell$ minor of $A(x)$. Then there is a unimodular polynomial matrix $R(x)$ such that $A(x)R(x)$ is column proper.

Proof. Let P be a permutation matrix such that the column degrees of $B(x) = A(x)P$ increase from left to right. Let δ_i be the column degree of column i in $B(x)$ and

$$[b_1 \ \dots \ b_\ell] = \lim_{x\to\infty} B(x)\,\mathrm{diag}(x^{-\delta_1} \ \dots \ x^{-\delta_\ell})$$

If the columns $b_1 \dots b_\ell$ are linearly independent then $B(x)$, and therefore $A(x)$, is column proper.

Suppose that the columns $b_1 \dots b_\ell$ are linearly dependent. Examine the columns from the left and suppose that b_k is the first one linearly dependent on the previous ones. Then it is possible to find constants $c_1 \dots c_{k-1}$ such that the k:th column of the matrix

$$[b_1 \ \dots \ b_\ell]\begin{pmatrix} 1 & & & c_1 & & \\ & \ddots & & \vdots & & \\ & & \ddots & c_{k-1} & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \qquad (4.6.18)$$

is zero. Here $c_1 \dots c_{k-1}$ are in the k:th column of the second matrix. The matrix

$$B_1(x) = B(x)\begin{pmatrix} 1 & & & c_1 x^{\delta_k-\delta_1} & & \\ & \ddots & & & & \\ & & \ddots & c_{k-1}x^{\delta_k-\delta_{k-1}} & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \qquad (4.6.18)$$

has the same column degrees as $B(x)$ except for the k:th column, which has lower degree. Since $\delta_i \geq \delta_j$ for $i > j$ the second matrix in (4.6.18) is a unimodular polynomial matrix.

If $B_1(x)$ is column proper then we are ready. If $B_1(x)$ is not column proper the procedure can be repeated with $B_1(x)$ substituted for $A(x)$.

Continuing like this we will eventually get a column proper matrix. That this is true follows from the fact that each iteration decreases the sum of the column degrees. If we have not got a column proper matrix before, we will eventually get a matrix with all column degrees equal to zero. Call this matrix H. Then

$$H = A(x)R(x) \tag{4.6.19}$$

for some unimodular matrix $R(x)$. By assumption there is a nonzero $\ell \times \ell$ minor of $A(x)$. It follows from (4.6.19) that the corresponding minor of H is nonzero. Therefore H is column proper.

□

Theorem 4.6.3 shows that to any controllable system there exists an equivalent system with system matrix (4.6.17) such that the controllability indices are the column degrees of $T(q)$. The next theorem shows that the column degrees of any such system matrix are the controllability indices of the system.

Theorem 4.6.4. Let the system matrix

$$P(q) = \left(\begin{array}{c|c} T(q) & I \\ \hline -V(q) & 0 \end{array}\right) \tag{4.6.20}$$

be in standard polynomial form in q with $T(q)$ column proper. Then the column degrees of $T(q)$ are the controllability indices (not in order).

Proof. Since $P(q)$ is controllable it is by Theorem 4.6.3 equivalent to $P_1(q)$ in (4.6.17). By Theorem 4.4.6 there is a unimodular polynomial matrix $Q(q)$ such that

$$T(q) = T_1(q)Q(q) \tag{4.6.21a}$$

$$V(q) = V_1(q)Q(q) \tag{4.6.21b}$$

By Theorem 4.6.3 the column degrees of $T_1(q)$ are equal to the controllability indices $\lambda_1 \ldots \lambda_\ell$ of $P_1(q)$ and therefore of $P(q)$ by Theorem 4.6.1. Let the column degrees of $T(q)$ be $\lambda'_1 \ldots \lambda'_\ell$ where $\lambda'_1 \leq \lambda'_2 \leq \ldots \leq \lambda'_\ell$. Because $T(q)$ and $T_1(q)$ are both column proper we have

$$\sum_{i=1}^{\ell} \lambda_i = \deg \det T_1(q) \tag{4.6.22}$$

$$\sum_{i=1}^{\ell} \lambda'_i = \deg \det T(q) \tag{4.6.23}$$

Since the orders of the systems corresponding to $P(q)$ and $P_1(q)$ are equal it follows by Theorem 4.5.3 that

$$\sum_{i=1}^{\ell} \lambda'_i = \sum_{i=1}^{\ell} \lambda_i \tag{4.6.24}$$

For the simplicity of notation suppose that the column degrees of $T(q)$ and $T_1(q)$ are ascending from left to right. This is no restriction. Then column i in $T_1(q)$ has degree λ_i and similarly for $T(q)$. Define

$$\bar{T}_1 = \lim_{x \to +\infty} T_1(x) \operatorname{diag}(x^{-\lambda_1}, \ldots, x^{-\lambda_\ell})$$

We see that $\lambda'_1 \geq \lambda_1$ because column 1 in $T(q)$ is by

(4.6.21a) a linear combination (by polynomials) of the columns of $T_1(q)$ and $\bar{T}_1$ has full rank.

Furthermore, $\lambda_2' \geq \lambda_2$ because otherwise $\lambda_1' \leq \lambda_2' < \lambda_2 \leq \leq \dots \leq \lambda_\ell$. This means that the linear combinations giving columns 1 and 2 in $T(q)$ cannot include columns 2, 3, ..., ℓ of $T_1(q)$ because $\bar{T}_1$ has full rank. So columns 1 and 2 in $T(q)$ must be multiples by polynomials of column 1 in $T_1(q)$. But this implies that $\det T(q) = 0$. Therefore $\lambda_2' \geq \lambda_2$.

Analogously $\lambda_i' \geq \lambda_i$ for $i = 3,4,\dots,\ell$. This together with (4.6.24) gives $\lambda_i' = \lambda_i$ for $i = 1,2,\dots,\ell$.

□

Theorem 4.6.4 can be used together with Lemma 4.6.1 and Theorem 4.4.4 to calculate the controllability indices for a controllable system in the following way. Transform the system to the form

$$P(q) = \left(\begin{array}{c|c} T(q) & I \\ \hline -V(q) & 0 \end{array}\right) \qquad (4.6.25)$$

which is in standard polynomial form in q. Use the method of the proof of Lemma 4.6.1 to make $T(q)$ column proper. The controllability indices are then the column degrees of $T(q)$.

Another way to determine the controllability indices is of course to transform the system to a state space form in standard polynomial form in q and then determine $\lambda_1 \dots \lambda_\ell$ from (4.6.12).

The controllability indices can also be determined from a system matrix in standard polynomial form in q^{-1}.

Theorem 4.6.5. Let the system matrix

$$P^*(q^{-1}) = \left(\begin{array}{c|c} T^*(q^{-1}) & I \\ \hline -V^*(q^{-1}) & 0 \end{array}\right) \tag{4.6.26}$$

be in standard polynomial form in q^{-1}. Suppose that the system is causal and that the matrix

$$\begin{pmatrix} T^*(x) \\ -V^*(x) \end{pmatrix}$$

is column proper. Then the column degrees of

$$\begin{pmatrix} T^*(x) \\ -V^*(x) \end{pmatrix}$$

are the controllability indices (not in order).

Proof. Let the column degrees of

$$\begin{pmatrix} T^*(x) \\ -V^*(x) \end{pmatrix}$$

be $\lambda_1', \ldots, \lambda_\ell'$. Because

$$\begin{pmatrix} T^*(x) \\ -V^*(x) \end{pmatrix}$$

is column proper there is an $\ell \times \ell$ minor of degree

$$\sum_{i=1}^{\ell} \lambda_i'$$

and no minor of higher degree. It follows from Theorem 4.5.5 that

$$n = \sum_{i=1}^{\ell} \lambda_i' \tag{4.6.26}$$

Define $Q(q) = \mathrm{diag}(q^{\lambda_1'}, \ldots, q^{\lambda_\ell'})$ and

$$P(q) \triangleq \left(\begin{array}{c|c} T(q) & I \\ \hline -V(q) & 0 \end{array}\right) = \left(\begin{array}{c|c} T^*(q^{-1}) & I \\ \hline -V^*(q^{-1}) & 0 \end{array}\right) \left(\begin{array}{c|c} Q(q) & 0 \\ \hline 0 & I \end{array}\right) \tag{4.6.27}$$

Clearly $P(q)$ is a polynomial matrix.

$$\begin{pmatrix} T(0) \\ -V(0) \end{pmatrix} = \lim_{x\to+\infty} \begin{pmatrix} T^*(x) \\ -V^*(x) \end{pmatrix} Q(x^{-1}) \tag{4.6.28}$$

Because

$$\begin{pmatrix} T^*(x) \\ -V^*(x) \end{pmatrix}$$

is column proper it follows that

$$\begin{pmatrix} T(0) \\ -V(0) \end{pmatrix}$$

has full rank and therefore $P(q)$ is in standard polynomial form in q. By Theorem 4.5.3

$$n = \deg \det T(q) \tag{4.6.29}$$

Let $\lambda_1'', \ldots, \lambda_\ell''$ be the column degrees of $T(q)$. It is always true that

$$\sum_{i=1}^{\ell} \lambda_i'' \geq \deg \det T(q) \tag{4.6.30}$$

This together with (4.6.29) and (4.6.26) gives

$$\sum_{i=1}^{\ell} \lambda_i'' \geq \sum_{i=1}^{\ell} \lambda_i' \tag{4.6.30}$$

But it follows from (4.6.27) that $\lambda_i'' \leq \lambda_i'$, $i = 1,\ldots,\ell$. Therefore

$$\lambda_i'' = \lambda_i' \quad \text{for} \quad i = 1, \ldots, \ell \tag{4.6.31}$$

Now (4.6.31), (4.6.26) and (4.6.29) give

$$\sum_{i=1}^{\ell} \lambda_i'' = \deg \det T(q) \tag{4.6.32}$$

This means that $T(q)$ is column proper. By Theorem 4.6.4 the numbers $\lambda_1'',\ldots,\lambda_\ell''$ are the controllability indices (not in order). Using (4.6.31) gives the desired result.

□

<u>Example 4.6.1</u>. Consider the system matrix

$$P(q^{-1}) = \left(\begin{array}{cc|cc} 1+q^{-1} & q^{-1} & 1 & 0 \\ q^{-2} & 1+q^{-2} & 0 & 1 \\ \hline -q^{-1} & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right) \triangleq \left(\begin{array}{c|c} T(q^{-1}) & I \\ \hline -V(q^{-1}) & 0 \end{array}\right)$$

It is easy to see that $P(q^{-1})$ is causal, controllable, and in standard polynomial form in q^{-1}. The matrix

$$\begin{pmatrix} T(x) \\ -V(x) \end{pmatrix}$$

is, however, not column proper since the matrix

$$\lim_{x\to\infty}\begin{pmatrix} 1+x & x \\ x^2 & 1+x^2 \\ -x & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x^{-2} & 0 \\ 0 & x^{-2} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

has linearly dependent columns.

Subtract the first column of $P(q^{-1})$ from the second. This gives the equivalent system matrix

$$P_1(q^{-1}) = \left(\begin{array}{cc|cc} 1+q^{-1} & -1 & 1 & 0 \\ q^{-2} & 1 & 0 & 1 \\ \hline -q^{-1} & q^{-1} & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right) \triangleq \left(\begin{array}{c|c} T_1(q^{-1}) & I \\ \hline -V_1(q^{-1}) & 0 \end{array}\right)$$

$P_1(q^{-1})$ is in standard polynomial form in q^{-1} and

$$\begin{pmatrix} T_1(x) \\ -V_1(x) \end{pmatrix}$$

is column proper.

By Theorem 4.6.5 the controllability indices are $\lambda_1 = 1$ and $\lambda_2 = 2$. By the corollary of Theorem 4.6.2 the order of the system is $n = \lambda_1 + \lambda_2 = 3$.

Multiplication of column 1 in $P_1(q^{-1})$ by q^2 and column 2 by q gives

$$P_2(q) = \left(\begin{array}{cc|cc} q^2+q & -q & 1 & 0 \\ 1 & q & 0 & 1 \\ \hline -q & 1 & 0 & 0 \\ 0 & -q & 0 & 0 \end{array}\right) \triangleq \left(\begin{array}{c|c} T_2(q) & I \\ \hline -V_2(q) & 0 \end{array}\right)$$

$P_2(q)$ is in standard polynomial form in q and $T_2(x)$ is column proper.

By Theorem 4.6.4 the controllability indices are $\lambda_1 = 1$ and $\lambda_2 = 2$.

$P_2(q)$ can be brought to state space form in the following way. The sign $\sim$ means "is equivalent to".

$$P_2(q) \sim \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 0 \\ 0 & q^2+q & -q & 1 & 0 \\ 0 & 1 & q & 0 & 1 \\ \hline 0 & -q & 1 & 0 & 0 \\ 0 & 0 & -q & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 0 \\ 0 & q^2+q+1 & 0 & 1 & 1 \\ 0 & 1 & q & 0 & 1 \\ \hline 0 & -q & 1 & 0 & 0 \\ 0 & 0 & -q & 0 & 0 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|cc} 1 & q & 0 & 0 & 0 \\ 0 & q^2+q+1 & 0 & 1 & 1 \\ 0 & 1 & q & 0 & 1 \\ \hline 0 & -q & 1 & 0 & 0 \\ 0 & 0 & -q & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|cc} 1 & q & 0 & 0 & 0 \\ -q-1 & 1 & 0 & 1 & 1 \\ 0 & 1 & q & 0 & 1 \\ \hline 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|cc} q & -1 & 0 & 0 & 0 \\ 1 & q+1 & 0 & 1 & 1 \\ 1 & 0 & q & 0 & 1 \\ \hline 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

Therefore (A,B,C,D) is an equivalent system in state space form, where

$$A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 0 \end{pmatrix} \qquad D = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

The controllability indices can be found by selecting linearly independent columns of the matrix.

$$[B \quad AB \quad A^2B] = \begin{pmatrix} 0 & 0 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & -1 \end{pmatrix}$$

as explained after (4.6.12). The first three columns are independent. Therefore $\lambda_1 = 1$ and $\lambda_2 = 2$.

□

4.7 Observability indices

Suppose we have an observable but not necessarily controllable system of the type

$$T(q)\xi = U(q)u \tag{4.7.1a}$$

$$y = V(q)\xi + W(q)u \tag{4.7.1b}$$

For any $y_+ \neq 0$ in Y^+ (see Def. 3.2.4) define τ so that $y_+(\tau) \neq 0$ and $y_+(t) = 0$ for $t = 1,2,\ldots,\tau-1$ and define $\eta \in R^m$ through $\eta = y_+(\tau)$.

Let $u \in \overline{R_z^\ell}$ and suppose that the state of the system at time $t = 1$ is nonzero. Because the system is observable this state gives a nonzero y_+ and therefore a τ and an η. All nonzero states in the same onedimensional subspace of the state space give by linearity the same τ. Because the state space is finite dimensional and the system is observable there is a finite largest value of τ that can be obtained from the states at $t = 1$. Call this largest value $\bar{\mu}_1$. It is easy to see that all η that can be obtained from states with $\tau = \bar{\mu}_1$ form a linear subspace $M_1 \subset R^m$. Let the dimension of M_1 be d_1 and define $\bar{\mu}_i = \bar{\mu}_1$, $i = 2,\ldots,d_1$.

Now find the largest τ that can be obtained from states giving $\eta \notin M_1$. Call this largest value $\bar{\mu}_{d_1+1}$. Clearly $\bar{\mu}_{d_1+1} < \bar{\mu}_1$. All η that can be obtained from states giving $\tau \geq \bar{\mu}_{d_1+1}$ form a linear subspace $M_2 \subset R^m$. Furthermore, $M_2 \supseteq M_1$. Let the dimension of M_2 be d_2 then $d_2 > d_1$.

Define $\bar{\mu}_i = \bar{\mu}_{d_1+1}$, $i = d_1+2, \ldots, d_2$.

Continue in this way until there are no states giving $\eta \notin M_{k-1}$ for some k. If $M_{k-1} \neq R^m$ define $M_k = R^m$ and $\bar{\mu}_i = 0$, $i = d_{k-1}+1, \ldots, m$. We have now defined m numbers $\bar{\mu}_1 \geq \bar{\mu}_2 \geq \ldots \geq \bar{\mu}_m$.

Def. 4.7.1. Define the observability indices $\{\mu_1, \ldots, \mu_m\}$ of the observable system (4.7.1) as $\mu_i = \bar{\mu}_{m-i+1}$, $i = 1, \ldots, m$.

Remark. Observe that $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_m$.

We have seen that the space R^m, in which the output $y_+(t)$ takes its values, has a subspace structure $M_1 \subset M_2 \subset \ldots \subset \subset M_k = R^m$ for some k. Introduce a basis $\{v_i\}_{i=1}^m$ in R^m in the following way. Let $v_1 \ldots v_{d_1}$ be a basis of M_1. Determine $v_{d_1+1} \ldots v_{d_2}$ so that $v_1 \ldots v_{d_2}$ becomes a basis of M_2 and so on. This choice of basis is of course not unique.

Because the system is observable it is possible to determine the state at $t = 1$ from y_+. It is, however, not necessary to know y_+ completely. It is sufficient to know $y_+(t)$ for $t = 1, \ldots, \bar{\mu}_m$, the projection of $y_+(t)$ on M_{k-1} along $\{v_i\}_{i=d_{k-1}+1}^m$ for $t = \bar{\mu}_m+1, \ldots, \bar{\mu}_{d_{k-1}}$, the projection of $y_+(t)$ on M_{k-2} along $\{v_i\}_{i=d_{k-2}+1}^m$ for $t = \bar{\mu}_{d_{k-1}}+ +1, \ldots, \bar{\mu}_{d_{k-2}}$ etc. This can be seen in the following way.

Suppose that two states $x_1(1)$ and $x_2(1)$ give the same projections defined above. Then the state $x(1) = x_1(1) - - x_2(1)$ will give zero projections. We will show that $x(1) = = 0$. Since $x(1)$ gives $y_+(t) = 0$ for $1 \leq t \leq \bar{\mu}_m$ it follows by definition of $\bar{\mu}_m$ that $x(1)$ gives an η in M_{k-1}. But the projection of $y_+(t)$ on M_{k-1} is zero for $1 \leq t \leq \leq \bar{\mu}_{d_{k-1}}$. It follows, by definition of $\bar{\mu}_{d_{k-1}}$ that η belongs to M_{k-2}. Continuing like this it follows eventually that $\eta \in M_1$. But the projection of $y_+(t)$ on M_1 is zero for $1 \leq t \leq \bar{\mu}_{d_1} = \bar{\mu}_1$. Therefore $y_+(t) = 0$ for $1 \leq t \leq \bar{\mu}_1$. By definition of $\bar{\mu}_1$ it follows that $y_+ = 0$ and therefore by observability that $x(1) = 0$.

Theorem 4.7.1. The observability indices are invariant under equivalence.

Proof. Let the system

$$T_1(q)\xi_1 = U_1(q)u \tag{4.7.2a}$$

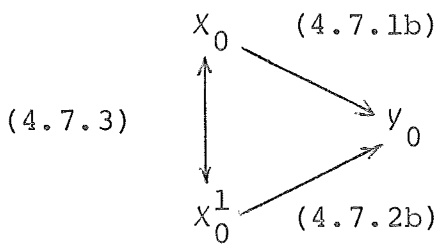
$$y_1 = V_1(q)\xi_1 + W_1(q)u \tag{4.7.2b}$$

where $u \in \overline{R_z^\ell}$, be equivalent to (4.7.1).

By the definition of equivalence (Def. 4.1.1) there is a bijective transformation

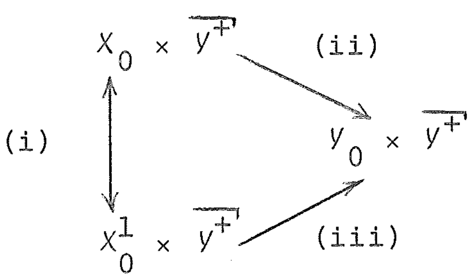
$$\xi_1' = Z(q)\xi' \tag{4.7.3}$$

where $\xi' \in X_0$ and $\xi_1' \in X_0^1$, between X_0 and X_0^1 such that the diagram



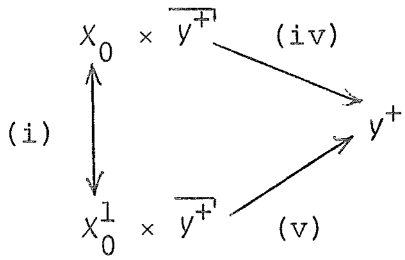
(4.7.4)

commutes. It follows that the diagram



(4.7.5)

where the mappings (i), (ii), and (iii) are defined from (4.7.4) in the obvious way, commutes and that (i) is bijective. Here $\overline{Y^{+}}$ is defined in Def. 3.2.6. Clearly diagram (4.7.5) is true if Y_0^+ is substituted for Y_0 and the mappings (ii) and (iii) modified accordingly. (For Y_0^+ see Def. 3.2.4) By Lemma 3.2.4 there is a bijection between $Y_0^+ \times \overline{Y^{+}}$ and Y^+. Therefore there are mappings (iv) and (v) such that the diagram



(4.7.7)

commutes. Now $X_0 \times \overline{Y^{+}}$ and $X_0^1 \times \overline{Y^{+}}$ are the state spaces for (4.7.1) and (4.7.2) respectively. The theorem now follows from the definition of observability indices because (i) is linear and bijective.

□

We will now find methods to compute the observability indices. It has been shown that there is an equivalent system in state space form and in standard polynomial form in q.

$$qx = Ax + Bu \tag{4.7.8a}$$

$$y = Cx + D(q)u \tag{4.7.8b}$$

Form the matrix

$$\left[C^T \quad A^T C^T \quad {A^2}^T C^T \quad \ldots \quad {A^{n-1}}^T C^T \right] \tag{4.7.9}$$

choose linearly independent columns as in (4.6.12) - (4.6.13). The n chosen columns will be of the form.

$$\left[c_1 \quad A^T c_1 \ldots (A^{\mu_1-1})^T c_1 \quad c_2 \quad A^T c_2 \ldots (A^{\mu_2-1})^T c_2 \ldots (A^{\mu_m-1})^T c_m \right]$$

(4.7.10)

for some integers $\mu_1' \ldots \mu_m'$. Here $C_1, \ldots, C_m$ are the columns of C^T.

It is well known that the system (4.7.8) is s.s. to a system of the form

$$
q\begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ \vdots \\ x_{p_m}\end{pmatrix} = \begin{pmatrix}
0 & & & a_{11} & & & & a_{12} & & & & a_{1m} \\
1 & \ddots & & \vdots & & & & \vdots & & & & \vdots \\
& \ddots & 0 & \vdots & & & & \vdots & & & & \vdots\\
& & 1 & a_{p_1 1} & & & & \vdots & & & & \vdots \\
& & & \vdots & 0 & & & a_{p_1+1\,2} & & & & \vdots \\
& & & \vdots & 1 & \ddots & & \vdots & & & & \vdots\\
& & & \vdots & & \ddots & 0 & \vdots & & & & \vdots\\
& & & \vdots & & & 1 & a_{p_2 2} & \ddots & & & \vdots\\
& & & \vdots & & & & \vdots & & \ddots & 0 & \vdots \\
& & & \vdots & & & & \vdots & & 1 & \ddots & \vdots\\
& & & a_{n1} & & & & a_{n2} & & & 1 & a_{p_m m}
\end{pmatrix}
\begin{pmatrix} x_1 \\ \vdots \\ x_{p_1} \\ x_{p_1+1} \\ \vdots \\ x_{p_2} \\ \vdots \\ x_{p_m}\end{pmatrix} + Bu
\tag{4.7.11a}
$$

$$
y = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ C_{21} & 1 & & \vdots \\ \vdots & C_{32} & & \vdots \\ \vdots & \vdots & & 0 \\ C_{m1} & C_{m2} & & 1\end{pmatrix} x + D(q)u
\tag{4.7.11b}
$$

where

$$
p_i = \sum_{j=1}^{i} \mu_j'
$$

The B and D(q) matrices have no special form.

Theorem 4.7.2. The observability indices are equal to the numbers $\{\mu_1' \ldots \mu_m'\}$ (not in order).

Proof. By Theorem 4.7.1 it is sufficient to show the theorem for the system (4.7.11). It was shown in Section 4.6 (4.6.8) - (4.6.11) that $x(t)$ is the state of the system (4.7.11) at time t. Define $\bar{y}(t) = Hy(t)$ where H is nonsingular and such that HC becomes as C in (4.7.11) but will all C_{ij} equal to zero. Suppose $\mu_1' \leq \ldots \leq \mu_m'$. This can always be achieved by a coordinate change in the state space and the output space.

Let $u_t = 0$ and $x(1)$ have the component $x_i \neq 0$ and all other components equal to zero. We have $p_{j-1} + 1 \leq i \leq p_j$ for some j $(p_0 \triangleq 0)$. This $x(1)$ will give $\bar{y}(t) = 0$ for $t = 1, \ldots, p_j - i$ and $\bar{y}_j(p_j - i+1) = x_i$, where $\bar{y}_j$ is the j:th component of $\bar{y}$. The other components of $\bar{y}(p_j - i+1)$ are zero. To determine μ_m we have to maximize $p_j - i+1$. This is done by choosing $j = m$ and $i = p_{m-1} + 1$. Then $p_j - i + 1 = \mu_m'$. Therefore $\mu_m = \mu_m'$. Suppose that $\mu_r' = \mu_m'$, $r = m, \ldots, m - \delta_1 + 1$ for some δ_1. Then we could have chosen j equal to any r, where $r = m, \ldots, m - \delta_1 + 1$. If $i = p_{j-1} + 1$ then they will all give $\bar{y}(t) = 0$, $t = 1, \ldots, \mu_r' - 1$ and $\bar{y}(\mu_r') \neq 0$. All $\bar{y}(\mu_r')$ will lie in the space spanned by $\{e_{m-\delta_1+1}, \ldots, e_m\}$, where e_i is the i:th column in I_m. This space is by definition M_1 and has dimension δ_1. Therefore $\mu_r = \mu_r'$ for $r = m, \ldots, m - \delta_1 + 1$.

To determine $\mu_{m-\delta_1}$ we have to choose an $x(1)$ which gives an η outside M_1. Therefore x_i must be nonzero for some $i = 1, \ldots, p_{m-\delta_1}$. Choose $x_i = 0$ for $i = p_{m-\delta_1} + 1, \ldots, n$ and repeat the procedure above. If $\mu_r' = \mu_{m-\delta_1}'$ for $r = m - \delta_1, \ldots, m - \delta_2 + 1$ then this will give $\mu_r = \mu_r'$ for $r = m - \delta_1, \ldots, m - \delta_2 + 1$. Furthermore M_2 will be spanned by $\{e_{m-\delta_2+1}, \ldots, e_m\}$.

If we continue like this it follows that $\mu_r = \mu'_r$ for $r = 1, \ldots, m$.

□

Corollary

$$\sum_{i=1}^{m} \mu_i = n$$

where n is the order of the system.

Now the observable versions of Theorems 4.6.3, 4.6.4 and 4.6.5, and Lemma 4.6.1 follow directly. We state the theorems. The proofs are analogous to the ones in Section 4.6.

Theorem 4.7.3. Any observable system is equivalent to a system

$$P_1(q) = \left(\begin{array}{c|c} T_1(q) & U_1(q) \\ \hline -I & 0 \end{array}\right)$$

in standard polynomial form in q. Furthermore $T_1(q)$ is row proper with row degrees equal to the observability indices (not in order).

Remark. "Row proper" and "row degree" are defined analogously to "column proper" and "column degree" (see Def. 4.6.2 and 4.6.3).

Lemma 4.7.1. Let $A(x)$ be an $m \times \ell$ polynomial matrix in x with $m \leq \ell$. Suppose $A(x)$ has full rank i.e. there is a nonzero $m \times m$ minor of $A(x)$. Then there is a unimodular polynomial matrix $R(x)$ such that $R(x)A(x)$ is row proper.

<u>Theorem 4.7.4.</u> Let the system matrix

$$P(q) = \left[\begin{array}{c|c} T(q) & U(q) \\ \hline -I & 0 \end{array}\right] \tag{4.7.12}$$

be in standard polynomial form in q with $T(q)$ row proper. Then the row degrees of $T(q)$ are the observability indices of the system (not in order).

<u>Theorem 4.7.5.</u> Let the system matrix

$$P^*(q^{-1}) = \left[\begin{array}{c|c} T^*(q^{-1}) & U^*(q^{-1}) \\ \hline -I & 0 \end{array}\right] \tag{4.7.13}$$

be in standard polynomial form in q^{-1}. Suppose that the system is <u>causal</u> and that the matrix $\left(T^*(x) \quad U^*(x)\right)$ is row proper. Then the row degrees of $\left(T^*(x) \quad U^*(x)\right)$ are the observability indices of the system (not in order).

5. REFERENCES

Åström K J (1970): Introduction to Stochastic Control Theory. Academic Press, New York.

Furhmann P A (1976): Algebraic System Theory; An Analyst's Point of View. Journal of the Franklin Institute 301, 521-540.

Fuhrmann P A (1977): On Strict System Equivalence and Similarity. International Journal of Control 25, 5-10.

Kalman R E, Falb P L, Arbib M A (1969): Topics in Mathematical System Theory. McGraw-Hill, New York.

Mac Duffee C C (1946): The Theory of Matrices. Chelsea Publishing Company, New York.

Pernebo L (1977): Notes on Strict System Equivalence. International Journal of Control 25, 21-38.

Rosenbrock H H (1970): State-space and Multivariable Theory. Nelson, London.

Rosenbrock H H (1974): Order, Degree and Complexity. International Journal of Control 19, 323-331.

Rosenbrock H H (1977): The Transformation of Strict System Equivalence. International Journal of Control 25, 11-19.

Rosenbrock H H, Hayton G E (1974): Dynamical Indices of a Transfer Function Matrix. International Journal of Control 20, 177-189.

Rowe A (1971): Some Algebraic Problems in System Theory. PhD thesis, Control System Centre, University of Manchester Institute of Science and Technology, Manchester, England.

Sinervo J, Blomberg H (1971): Algebraic Theory for Ordinary Linear Time-Invariant Difference Systems. Acta Polytechnica Scandinavica Ma 21, Helsinki, Finland.

Wolovich W A (1974): Linear Multivariable Systems. Springer--Verlag, New York.

Part II - Algebraic Design Theory

1. INTRODUCTION

The problem of controlling a linear, time invariant system has been examined extensively. The problem may be divided into three sub-problems, the servo problem, the regulator problem, and the sensitivity problem. The servo problem and the regulator problem will be treated in this thesis. The sensitivity problem means that the closed loop system should be insensitive to modelling errors. This problem will not be treated here. It will be examined what can be done if the system is known accurately.

Consider a linear, time invariant, finite dimensional, causal, dynamical system, which can be represented by the following block diagram.

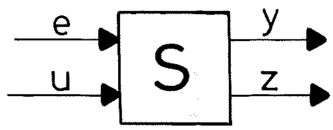


Figure 1.1. The system to be controlled.

The system is supposed to have two input vectors u and e and two output vectors y and z. The components of u are the control variables and the components of e represent the disturbances that act on the system. It will in general be assumed that e cannot be measured. The output vector y contains all variables that are to be controlled and the vector z contains all variables that can be measured. The vectors y and z may have components in common. Both continuous and discrete time systems are considered.

The control problem is assumed to be formulated as follows. Find a linear, time invariant, finite dimensional, causal controller R, such that the closed loop system fulfils requirements of the following types.

Servo requirements:

The input - output relation between a command input u_r and the output y should satisfy given specifications.

Regulator requirements:

The input - output relation between the disturbance e and the output y should satisfy given specifications.

Stability requirements:

The closed loop system should be stable.

An example of a servo specification is that y should be equal to u_r. An example of a regulator specification is that step changes in e should give no steady state error in y.

The information available to the controller R is the measured output z and the command input u_r. A block diagram showing the closed loop system is found in figure 1.2.

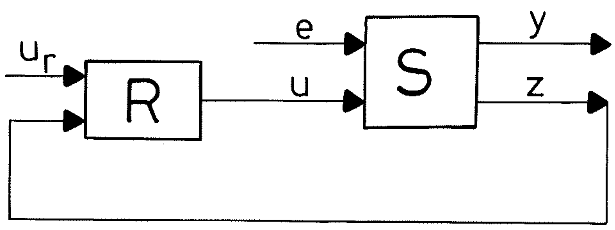


Figure 1.2. The closed loop system.

The following viewpoint will be adopted. The closed loop system is assumed to work over an infinite time interval and the behaviour right after start-up is assumed to be of minor importance. The important thing is that the servo and regulator requirements are satisfied when the effects of the initial conditions have vanished. Note that these effects will be negligible after a short time interval if the stability requirements are satisfied. Observe also that this viewpoint does not imply that the design is made for a steady state situation. The inputs e and u_r may vary arbitrarily.

A consequence of this viewpoint is that the system S and the controller R are appropriately described by their multivariable transfer functions. Furthermore, the expression "input - output relation" can be substituted by "transfer function" in the servo and regulator requirements. The transfer function is however not sufficient for the closed loop system when stability is dealt with. The reason is that some variables may become uncontrollable or unobservable when S or R are connected. These variables will occur as "common factors", that will cancel out, if the transfer function is formed. This problem can be circumvented if a polynomial matrix representation, due to Rosenbrock (1970), is used. The stability requirement is treated in chapter 4.

The solution to the control problem will be obtained from solutions to different polynomial matrix equations and rational matrix equations. Not all solutions to these equations are however valid. It turns out that the valid solutions are precisely those that are generalized polynomial matrices. The set of generalized polynomials is, in chapter 2, defined to be the set of rational functions with poles outside some subset Λ of the complex plane. It is shown to be a ring and in fact a principal ideal domain. Many useful results can be shown for matrices with entries in a principal ideal domain. Some results are given in e.g. Mac Duffee (1946). If the equations, mentioned previously, are interpreted as generalized polynomial matrix equations then necessary and sufficient conditions for the existence of valid solutions are easily obtained. These conditions will thus be necessary and sufficient for the existence of a solution to the control problem. Different servo and regulator specifications correspond to different subsets Λ.

There is a system theoretical interpretation of the use of generalized polynomials. Let Λ be regarded as the unstable region of the system. Poles and zeros of the transfer function outside Λ then correspond to stable and minimum phase properties of the system, i.e. to the part of the system that behaves well and needs no special attention. This corresponds to the fact that the units, i.e. invertible elements, in the ring of generalized polynomials are precisely the rational functions with poles and zeros outside Λ. Multiplication or division of an element in a ring by a unit does not change the properties of the element from the point of view of the ring.

The introduction of generalized polynomial matrices gives a tool for isolation of the part of a system that is difficult to handle from the part that is easily handled. The part that is difficult to handle is completely described by a polynomial matrix, called the structure matrix. Two versions are defined in chapter 3. The left structure matrix gives information about how well the system can be controlled in the servo sense. It can therefore be used to formulate the servo specifications or to check if given specifications can be fulfilled. The right structure matrix gives information about how well the disturbances can be estimated from measurements of the output z. These measurements are used by the controller in order to fulfil the regulator specifications. The right structure matrix will therefore appear in the conditions that have to be checked in order to determine if the regulator specifications can be fulfilled. The left and right structure matrices are dual concepts.

It will be shown that the control problem can be separated into two parts. First choose a feedback controller R_{fb} which satisfies the regulator and stability requirements. Then choose a feedforward controller R_{ff} to satisfy the servo requirements. The configuration is shown in figure 1.3.

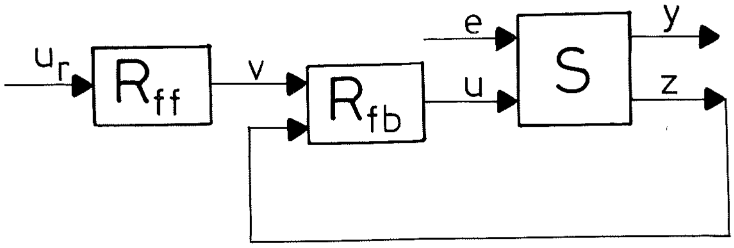


Figure 1.3. Separation of the control problem.

The problem to determine R_{fb} is solved by feedforward design and feedback realization in the following way. Consider the control configuration in figure 1.4.

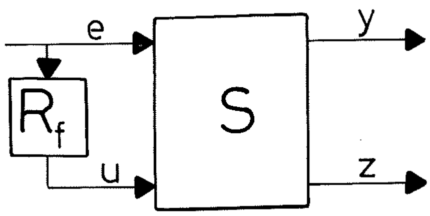


Figure 1.4. A possible control configuration if e were measurable.

The controller R_f is said to be feedback realizable if there is a causal and stabilizing feedback controller R_{fb} from z to u, such that the system in figure 1.5 has the same transfer function from e to u as the system in figure 1.4.

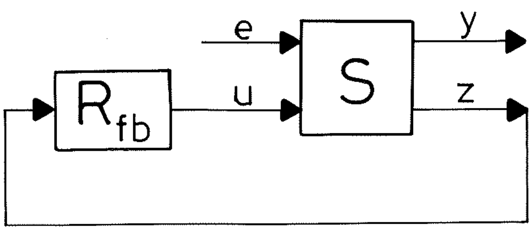


Figure 1.5. A feedback realization of the system in figure 1.4.

Necessary and sufficient conditions for R_f to be feedback realizable are given in chapter 7. The right structure matrix is used to characterize the class of feedback realizable transfer functions. It is also shown how a causal and stabilizing R_{fb} can be computed from the system S and a feedback realizable R_f. The special case of z being the state vector has been solved in Bengtsson (1977b). The results in chapter 7 extend Bengtsson's results to arbitrary z.

The control configuration in figure 1.4 with feedback realizable R_f is in chapter 8 used to find a method for checking if the regulator specifications can be fulfilled. It is also shown how a feedback realizable R_f, that satisfies the regulator specifications, can be computed, if one exists.

The results in chapters 7 and 8 are combined to compute a feedback controller R_{fb} that satisfies regulator and stability requirements if this is possible. When this is done it has to be determined how the input v in figure 1.3 shall be connected to R_{fb}. In chapter 6 it is shown that there is a "best way" to do this.

It remains to determine the feedforward controller R_{ff} in figure 1.3. Let G_{vy} be the transfer function from v to y and let K be the transfer function for the controller R_{ff}. It follows that $G_{vy}K = H$, where H is the transfer function from the command input u_r to the controlled output y.

The transfer function G_{vy} is given and a stable and causal K should be determined, such that H fulfils the servo specifications. This problem has been treated by many authors. In Wang, Davison (1973a) H is supposed to be given and a least order and causal K is found if one exists. They can however not guarantee that K becomes stable. Bengtsson, Wonham (1976) do also suppose that H is given and a stable and causal K is found if one exists. This K might not be of least order.

The results of Bengtsson, Wonham (1976) will be extended in chapter 5. It is shown how the class of possible transfer functions H can be characterized by the left structure matrix of the system S. The characterization makes it possible to check if the servo specifications can be fulfilled. It also gives a method for choosing H in this class, such that the servo specifications are fulfilled, if such an H exists. An algorithm for finding a stable and causal K, when H is given properly, is presented. It will not necessarily give a K of least order.

The design procedure in chapters 5 - 8 is applied to a simple multivariable system in chapter 9. Simulations show that the closed loop system behaves as desired. The simulations also show that the closed loop system is, in this example, insensitive to parameter changes in the system S.

2. GENERALIZED POLYNOMIALS

Generalized polynomials are defined in this chapter. They are useful when dealing with design problems for linear, time invariant systems. They provide for instance a convenient way to formalize allowable cancellations.

A generalized polynomial is defined as a rational function with all its poles outside a given subset Λ of the complex plane C. The reason for giving them a name including the word "polynomial" and not "rational function" is that the generalized polynomials algebraically have more properties in common with the polynomials than with the rational functions. They are a ring and in fact a euclidean domain just as the polynomials are, but not a field like the rational functions.

2.1. Λ-generalized polynomials

Let R and C be the fields of real and complex numbers and let Λ be an arbitrary subset of C. Furthermore let $R[\lambda]$ be the polynomials and $R(\lambda)$ the rational functions with coefficients in R. Poles and zeros of a rational function are as usual defined as the zeros in C of the denominator and numerator polynomials.

Definition 2.1. The set of Λ-generalized polynomials, denoted $R_{\Lambda}[\lambda]$, is defined as the set of rational functions with no poles in Λ.

Remark. If Λ is specialized to C then $R_{\Lambda}[\lambda]$ becomes the ordinary polynomials $R[\lambda]$. With $\Lambda=\Phi$, the empty set, $R_{\Lambda}[\lambda]$ is equal to the rational functions $R(\lambda)$. Finally if Λ is $C \setminus \{0\}$ then $R_{\Lambda}[\lambda]$ is the Laurent polynomials $R(\lambda]$ (part 1 of this thesis).

Define addition and multiplication in $R_{\Lambda}[\lambda]$ as it is defined in $R(\lambda)$. It then follows directly from the definition of a ring that $R_{\Lambda}[\lambda]$ is a ring. The invertible elements, units, in the ring $R_{\Lambda}[\lambda]$ are characterized by the following lemma.

<u>Lemma 2.1</u> The units in $R_\Lambda[\lambda]$ are all nonzero rational functions with no poles and zeros in Λ.

An integral domain is a ring that fulfils the following three conditions.

(i) There is a multiplicative identity.

(ii) Multiplication is commutative.

(iii) $a \cdot b = 0 \Rightarrow a = 0$ or $b = 0$ for all a and b in the ring.

<u>Lemma 2.2</u> The ring of Λ-generalized polynomials $R_\Lambda[\lambda]$ is an integral domain.

<u>Lemma 2.3</u> Every nonzero $a \in R_\Lambda[\lambda]$ can be uniquely factorized as $a = p \cdot r$, where p is a monic polynomial with all its zeros in Λ and r is a unit in $R_\Lambda[\lambda]$.

The simple proofs of the lemmas above are omitted.

<u>Definition 2.2</u> Let a be a nonzero element in $R_\Lambda[\lambda]$. The Λ-degree of a, denoted $\deg_\Lambda a$, is the degree of the polynomial p, defined through lemma 2.3.

<u>Remark 1.</u> If $\Lambda = C$ so that $R_\Lambda[\lambda]$ is the ring of polynomials then the degree will be denoted "deg" as usual.

<u>Remark 2.</u> Note that if a is a polynomial then in general $\deg_\Lambda a \neq \deg a$. Let for instance Λ be the left half plane then $\deg_\Lambda(\lambda^2-1) = 1$ while $\deg(\lambda^2-1) = 2$.

<u>Theorem 2.1</u> $R_\Lambda[\lambda]$ is a euclidean domain.

<u>Proof</u> An integral domain D is a euclidean domain if there is a function v from the nonzero elements of D into the nonnegative integers such that

(i) For all pairs a, b in D for which $b \neq 0$ there exist q and r in D such that

$$a = bq + r \quad \text{with } v(r) < v(b) \text{ or } r = 0$$

(ii) For all pairs a, b in D for which $a \neq 0$ and $b \neq 0$

$$v(a) \leq v(a\,b).$$

For all $a \in R_\Lambda[\lambda]$ define $v(a) = \deg_\Lambda a$. Choose a and b in $R_\Lambda[\lambda]$ with $b \neq 0$. Decompose them according to lemma 2.3 as $a = a_p a_r$ and $b = b_p b_r$, where a_p and b_p belong to $R[\lambda]$. By the division algorithm for polynomials there are x and y in $R[\lambda]$ such that

$$a_p = b_p x + y \quad \text{with } \deg y < \deg b_p \text{ or } y = 0. \tag{2.1}$$

Multiply (2.1) by a_r

$$a = b_p\, x\, a_r + y\, a_r$$

$$\Rightarrow \quad a = b\, x\, \frac{a_r}{b_r} + y\, a_r$$

Choose $q = x \frac{a_r}{b_r}$ and $r = y\, a_r$. Clearly q and r belong to $R_\Lambda[\lambda]$. If $y = 0$ then $r = 0$ and (i) is fulfilled. If $y \neq 0$ then $\deg_\Lambda r \leq \deg y < \deg b_p = \deg_\Lambda b$ and (i) is fulfilled.

Choose nonzero a and b in $R_\Lambda[\lambda]$ and decompose them as above, $a = a_p a_r$ and $b = b_p b_r$. Then

$$\deg_\Lambda a = \deg a_p \leq \deg a_p b_p = \deg_\Lambda\, a_p b_p a_r b_p = \deg_\Lambda\, a\, b$$

and (ii) is fulfilled. □

Corollary $R_\Lambda[\lambda]$ is a principal ideal domain.

Proof Every euclidean domain is a principal ideal domain. □

2.2 Λ-generalized polynomial matrices

Matrices with entries in $R_\Lambda[\lambda]$ will now be considered. These matrices have properties analogous to the properties of the polynomial matrices. The reason is that most of the results shown for polynomial matrices are in fact special cases of results that can be shown for matrices with entries in any principal ideal domian or euclidean domain. This is for instance done in Mac Duffee(1946).

Different concepts, that can be defined for matrices with entries in any principal ideal domain, will be given below for Λ-generalized polynomial matrices. We will use the conventional names, but the names will be preceeded by "Λ"-(e.g. Λ-unimodular, Λ-equivalent) to indicate that the underlying principal ideal domain is the ring of Λ-generalized polynomials. If $\Lambda = C$ so that we have ordinary polynomials then the "Λ" will be omitted.

Definition 2.3 A Λ-generalized polynomial matrix is a matrix with entries in $R_\Lambda[\lambda]$. The set of $n\times m$ matrices with entries in $R_\Lambda[\lambda]$ is denoted $R_\Lambda^{n\times m}[\lambda]$.

Definition 2.4 A matrix $A \in R_\Lambda^{n\times n}[\lambda]$ is Λ-unimodular if there is a $B \in R_\Lambda^{n\times n}[\lambda]$ such that $AB = I$.

The following theorem is shown in Mac Duffee (1946) for a principal ideal domain.

Theorem 2.2 A matrix $A \in R_\Lambda^{n\times n}[\lambda]$ is Λ-unimodular if and only if det A is a unit in $R_\Lambda[\lambda]$, i.e. is a nonzero rational function with no poles or zeros in Λ.

Definition 2.5 The matrices A and B in $R_\Lambda^{n\times m}[\lambda]$ are Λ-equivalent if there are Λ-unimodular $U \in R_\Lambda^{n\times n}[\lambda]$ and $V \in R_\Lambda^{m\times m}[\lambda]$ such that $A = U\,B\,V$.

Theorem 2.3 Every $A = R^{n\times m}[\lambda]$ is Λ-equivalent to a matrix $S \in R_\Lambda^{n\times m}[\lambda]$ in Λ-Smith form.

$$S = (D \quad 0) \quad \text{if } m > n$$
$$S = D \quad \text{if } m = n$$
$$S = \begin{bmatrix} D \\ 0 \end{bmatrix} \quad \text{if } m < n$$

Here $D = \text{diag}(i_1,\ldots,i_k,0,\ldots,0)$, where $\{i_j\}_1^k$ are monic polynomials with all their zeros in Λ. Furthermore $i_j | i_{j+1}$ for $j = 1,2,\ldots, k-1$ and $\{i_j\}_1^k$ are called the Λ-invariant factors of A. The Λ-Smith form is unique.

Proof The existence of an equivalent Smith form with invariant factors, that are unique up to units, is shown for a general principal ideal domain in e.g. Mac Duffee (1946). In our case we can choose the units so that the Λ-invariant factors become monic polynomials with all their zeros in Λ. They will then clearly be unique.

□

Since $R_\Lambda^{n\times m}[\lambda] \subset R^{n\times m}(\lambda)$ the definitions of rank and linearly independent column or row vectors as stated for rational matrices also apply to Λ-generalized polynomial matrices. Let $R_\Lambda^n[\lambda]$ denote the vectors of n-tuples of Λ-generalized polynomials.

Definition 2.6 Let $\{a_i\}_1^k$ belong to $R_\Lambda^n[\lambda]$. $\{a_i\}_1^k$ are linearly independent if the only $b_i \in R(\lambda)$, $i = 1,\ldots,k$ that satisfy

$$b_1 a_1 + \ldots + b_k a_k = 0$$

are $b_i = 0$, $i = 1, \ldots, k$.

Definition 2.7 The rank of $A \in R_\Lambda^{n\times m}[\lambda]$ can be defined in either of the following three ways

(i) The highest order of any nonzero minor of A.

(ii) The maximal number of linearly independent columns that can be found among the columns of A.

(iii) The maximal number of linearly independent rows that can be found among the rows of A.

The definitions are equivalent.

Remark Linear independence for complex vectors is defined in the usual way i.e. through definition 2.6 with C^n substituted for $R_\Lambda^n[\lambda]$ and C for $R(\lambda)$. Definition 2.7 can then be applied to complex matrices. This gives the usual definition for complex matrices.

A rational matrix is left (right) invertible if and only if its columns (rows) are linearly independent. If a Λ-generalized polynomial matrix has linearly independent columns (rows) then consequently there is a rational left (right) inverse. It might, however, not be possible to find a left (right) inverse, which is a Λ-generalized polynomial matrix.

Definition 2.8. The matrix $A \in R_{\Lambda}^{n\times m}[\lambda]$ is right Λ-invertible if there is a $B \in R_{\Lambda}^{m\times n}[\lambda]$ such that $AB = I$.

Theorem 2.4. Let $A \in R_{\Lambda}^{n\times m}[\lambda]$, $n \le m$. The following statements are equivalent.

(i) A is right Λ-invertible.

(ii) The Λ-Smithform of A is [I 0].

(iii) $A(\lambda_o)$ has linearly independent rows for every $\lambda_o \in \Lambda$.

(iv) There is a $B \in R_{\Lambda}^{(m-n)\times m}[\lambda]$ such that $\begin{bmatrix} A \\ B \end{bmatrix}$ is Λ-unimodular.

Remark. If $\Lambda = \Phi$, i.e. A is a rational matrix, then (iii) is meaningless and should be substituted by:

(iii)' A has linearly independent rows in the sense of definition 2.6.

We will prove the theorem in the case $\Lambda \neq \Phi$, but only minor modifications have to be made if $\Lambda = \Phi$.

Proof. (i) $\Leftrightarrow$ (ii). Let S be the Λ-Smith form of A. Then there are Λ-unimodular N and M such that $S = N\,A\,M$. Suppose that A is right Λ-invertible and let $B \in R_{\Lambda}^{m\times n}[\lambda]$ be a right Λ-inverse of A. Then $M^{-1}B\,N^{-1}$ belongs to $R_{\Lambda}^{m\times n}[\lambda]$ and is a right Λ-inverse to S. Conversely suppose that S is right Λ-invertible and $Z \in R_{\Lambda}^{m\times n}[\lambda]$ be a right Λ-inverse to $S(\lambda)$. Then $M\,Z\,N$ belongs to $R_{\Lambda}^{m\times n}[\lambda]$ and is a right Λ-inverse to A. We have shown that A is right Λ-invertible if and only if S is right Λ-invertible.

If S is equal to $[I \quad 0]$ then $Z = \begin{bmatrix} I \\ 0 \end{bmatrix}$ is a right Λ-inverse to S. Conversely suppose that S is right Λ-invertible and that $Z \in R_{\Lambda}^{m\times n}[\lambda]$ is a right Λ-inverse. According to theorem 2.3 S has the form $[D \quad 0]$, where $D = \text{diag}\,(d_1,\ldots,d_n)$ and the nonzero d_i are monic polynomials with no zeros outside Λ. Decompose Z as $Z = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$. Then

$$DZ_1 = I \tag{2.2}$$

It follows that all d_i must be nonzero. (2.2) is then equivalent to

$$Z_1 = D^{-1} \tag{2.3}$$

The left member of (2.3) has no poles in Λ since $Z \in R_\Lambda^{m\times n}[\lambda]$ and D^{-1} has no poles outside Λ. Therefore D^{-1} is a polynomial matrix, i.e. d_i^{-1} are polynomials. Since d_i by definition are monic polynomial they must all be equal to 1. I.e. $S = [I \quad 0]$.

(ii) ⇒ (iv). We have $A = N[I_n \quad 0]M$, where N and M are Λ-unimodular. Choose $B = [0 \quad I_{m-n}]M$. Then

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} N & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} M \tag{2.4}$$

Clearly $B \in R_\Lambda^{(m-n)\times n}[\lambda]$ and $\begin{bmatrix} A \\ B \end{bmatrix}$ is Λ-unimodular.

(iv) ⇒ (iii). Choose $\lambda_o \in \Lambda$. We know that $N \triangleq \begin{bmatrix} A \\ B \end{bmatrix}$ belongs to $R_\Lambda^{m\times m}[\lambda]$ Therefore $N(\lambda_o)$ is a well defined matrix with entries in C. Since N is Λ-unimodular it follows from theorem 2.2 that $\det N(\lambda_o)$ is a nonzero complex number. Consequently $N(\lambda_o)$, and therefore also $A(\lambda_o)$, has linearly independent rows.

(iii) ⇒ (ii). Let $S = [D \quad 0]$, where $D = \mathrm{diag}(d_1 \ldots d_m)$, be the Λ-Smith form of A. Then there are Λ-unimodular matrices N and M such that

$$[D \quad O] = N\,A\,M \tag{2.5}$$

It was shown above that N and M are nonsingular complex matrices for every λ_o in Λ. Since A has linearly independent rows for every $\lambda_o \in \Lambda$ it follows that also the right member of (2.5) has linearly independent rows for every $\lambda_o \in \Lambda$. The polynomials d_i

must therefore be nonzero for all λ_o Λ. It follows from theorem 2.3 that the polynomials d_i have no zeros outside Λ. Consequently they are nonzero for all Λ in C. They must therefore be independent of λ. They are then by definition equal to 1.

□

2.3 The structure matrix

It was mentioned previously that a Λ-generalized polynomial matrix is not necessarily (left or right) Λ-invertible. In this section it will be shown how an arbitrary Λ-generalized polynomial matrix can be factorized into a product of a matrix, called the structure matrix, and a Λ-invertible matrix. This factorization will be of great importance when discussing dynamical systems. The structure matrix represents in some sense the part of the matrix that is difficult to handle, while the Λ-invertible part is easily handled. The factorization is essentially unique.

One way to do the factorization is given in theorem 2.5 below. This factorization will be of interest in chapter 3, but it is from the modified version in theorem 2.6 that the structure matrix is defined.

Theorem 2.5 Let $A \in R_\Lambda^{n\times m}[\lambda]$ have rank r. Then A can be factorized as A = B C, where $B \in R_\Lambda^{n\times r}[\lambda]$ has rank r and C is a right Λ-invertible matrix in $R_\Lambda^{r\times m}[\lambda]$. The matrix B is unique up to multiplication from the right by Λ-unimodular matrices.

Proof Let S be the Λ-Smith form of A. Then S can be written

$$S = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}, \tag{2.6}$$

where rank D = rank A = r. There are Λ-unimodular matrices N and M such that A = N S M. Partition N and M compatibly with the partitioning of S:

$$A = [N_1 \quad N_2] \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}. \tag{2.7}$$

This is equivalent to

$$A = N_1 D\, M_1 \tag{2.8}$$

Choose $B = N_1 D$ and $C = M_1$. Then B belongs to $R_\Lambda^{n\times r}[\lambda]$ and C belongs to $R_\Lambda^{r\times m}[\lambda]$. Furthermore C is right Λ-invertible by theorem 2.4. ((iv) $\Rightarrow$ (i)) and B has rank r since A has. We have thus shown that there is a factorization of the desired type.

Let $A = B_1 C_1$ and $A = B_2 C_2$ be two different factorizations. Since rank $A = r$ and B_1 has r columns it follows that B_1 has linearly independent columns. Therefore there is a rational matrix $\bar{B}_1 \in R^{r\times n}(\lambda)$ such that $\bar{B}_1 B_1 = I$. Multiply the equality $B_1 C_1 = B_2 C_2$ from the left by $\bar{B}_1$. This gives

$$C_1 = R\, C_2\ , \tag{2.9}$$

where $R = \bar{B}_1 B_2$ is a rational $r\times r$ matrix. Rank $R = r$ since rank $C_1 = r$. Therefore R^{-1} exists as a rational matrix. By theorem 2.4 ((i) $\Rightarrow$ (ii)) there exist Λ-unimodular N and M such that

$$C_2 = N[I \quad O]M \tag{2.10}$$

Insert this into (2.9)

$$C_1 = R\, N[I \quad O]M$$

$$\Leftrightarrow \quad C_1 M^{-1} = R[I \quad O]\begin{bmatrix} N & O \\ O & I \end{bmatrix}$$

$$\Rightarrow \quad C_1 M^{-1}\begin{bmatrix} N^{-1} & O \\ O & I \end{bmatrix} = [R \quad O] \tag{2.11}$$

The left member of (2.11) is a Λ-generalized matrix. Therefore R is also a Λ-generalized polynomial matrix. From (2.9) we get

$$C_2 = R^{-1} C_1 \tag{2.12}$$

We can now proceed as above to conclude that R^{-1} is also a Λ-generalized polynomial matrix. Consequently R is Λ-unimodular.

Introduce (2.9) into

$$B_1C_1 = B_2C_2. \tag{2.13}$$

This gives

$$B_1R\,C_2 = B_2C_2 \tag{2.14}$$

or since C_2 is right Λ-invertible

$$B_1\,R = B_2 \tag{2.15}$$

where R is Λ-unimodular.

□

The factorization of theorem 2.5 can be modified so that the matrix B becomes a polynomial matrix, which is essentially unique. To show this we need the following two lemmas.

<u>Lemma 2.4</u> Let the polynomial matrix $A \in R^{m\times n}[\lambda]$, $n \leq m$, have linearly independent columns. Then A can be written $A = B\,C$, where $B \in R^{m\times n}[\lambda]$ has linearly independent colums and $C \in R^{n\times n}[\lambda]$ has nonzero determinant. Furthermore the invariant factors of B have no zeros outside Λ and det C has no zeros inside Λ.

<u>Proof</u> Let $\begin{bmatrix} D \\ 0 \end{bmatrix}$, where $D = \text{diag}(d_i,\ldots d_n)$, be the Smithform of the polynomial matrix A. Factorize every $d_i = \hat{d}_i\tilde{d}_i$, where $\hat{d}_i$ has all its zeros in Λ and $\tilde{d}_i$ has no zeros in Λ. Define $\hat{D} = \text{diag}(\hat{d}_i,\ldots,\hat{d}_n)$ and $\tilde{D} = \text{diag}(\tilde{d}_i,\ldots,\tilde{d}_m)$. Then $D = \hat{D}\,\tilde{D}$ and there are unimodular N and M such that

$$A = N\begin{bmatrix} \hat{D} \\ 0 \end{bmatrix} \tilde{D}\,M \tag{2.16}$$

Since A has linearly independent columns it follows that all d_i are nonzero. Therefore all $\hat{d}_i$ and $\tilde{d}_i$ are also nonzero. Define

$$B = N\begin{bmatrix} \hat{D} \\ 0 \end{bmatrix} \quad \text{and} \quad C = \tilde{D}\,M. \tag{2.17}$$

The Smith form of B is $\begin{bmatrix} \hat{D} \\ 0 \end{bmatrix}$. It follows that B and C fulfil the conditions of the lemma.

□

<u>Lemma 2.5</u> Let the polynomial matrix $A \in R^{m\times n}[\lambda]$; $n \leq m$, have linearly independent columns and invariant factors with no zeros outside Λ. Let B belong to $R^{m\times k}[\lambda]$. If there is an $L \in R_{\Lambda}^{n\times k}[\lambda]$ such that

$$A\ L = B \tag{2.18}$$

then L is a polynomial matrix.

<u>Proof</u> Suppose there is a Λ-generalized polynomial matrix L satisfying (2.18). It will then be shown that L must be a polynomial matrix.

There are unimodular polynomial matrices N och M such that

$$\begin{bmatrix} D \\ 0 \end{bmatrix} = N\ A\ M \tag{2.19}$$

is the Smithform of A. Here $D = \text{diag}(d_i, \ldots, d_n)$, where d_i are nonzero polynomials with no zeros outside Λ. Introduce (2.19) into (2.18)

$$\begin{pmatrix} D \\ 0 \end{pmatrix} M^{-1}L = N\ B \tag{2.20}$$

where N B is polynomial. Partition N B as $N\ B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$. Then

$$D\ M^{-1}L = B_1$$

$$\Leftrightarrow\ M^{-1}L = D^{-1}B_1 \tag{2.21}$$

The left member has no poles in Λ and the right member has no poles outside Λ. $M^{-1}L$ will therefore be polynomial. Consequently $L = M\ M^{-1}L$ is polynomial.

□

The desired factorization of a Λ-generalized polynomial matrix is given by the following theorem.

<u>Theorem 2.6</u> Let $A \in R^{n\times m}[\lambda]$ have rank r. Then A can be factorized

as $A = \hat{A}\,\tilde{A}$, where $\tilde{A}$ is a right Λ-invertible matrix in $R_\Lambda^{r\times m}[\lambda]$ and $\hat{A}$ is a <u>polynomial</u> matrix in $R^{n\times r}[\lambda]$ with rank r. The invariant factors of A have no zeros outside Λ. The matrix $\hat{A}$ is unique up to multiplication from the right by unimodular polynomial matrices.

<u>Proof</u> The desired factorization of $A = \hat{A}\,\tilde{A}$ is of the type described in theorem 2.5, with the additional condition that $\hat{A}$ is a polynomial matrix, whose invariant factors have no zeros outside Λ.

Let A = B C be a factorization of the type described in theorem 2.5. We have to show that there is a Λ-unimodular R such that B R is a polynomial matrix with no zeros of its invariant factors outside Λ.

For i = 1,...,r, choose r_i equal to the least common denominator of the elements in column i of B. Then $R_1 = \mathrm{diag}(r_1 \ldots r_r)$ is Λ-unimodular and

$$B_1 = B\,R_1 \tag{2.22}$$

is a polynomial matrix. Furthermore B_1 has rank r and therefore linearly independent columns.

Apply lemma 2.4 to B_1

$$B_1 = B_2 R_2 \tag{2.23}$$

where B_2 is a polynomial matrix, whose invariant factors have no zeros outside Λ and R_2 is Λ-unimodular. Combination of (2.22) and (2.23) gives

$$B_2 = B\,R \tag{2.24}$$

where $R = R_1 R_2^{-1}$ is Λ-unimodular. Choose $\hat{A} = B_2$ and $\tilde{A} = R^{-1}C$. Clearly $\tilde{A}$ is right Λ-invertible since C is. We have thus found a factorization of the desired type.

Let $A = \hat{A}_1\tilde{A}_1$ and $A = \hat{A}_2\tilde{A}_2$ be two different factorizations of the type described in the theorem. By theorem 2.5 there is a Λ-unimodular matrix R such that

$$\hat{A}_2 = \hat{A}_1 R \tag{2.25}$$

By lemma 2.5 R is polynomial. Lemma 2.5 applied to

$$\hat{A}_1 = \hat{A}_2 R^{-1} \tag{2.26}$$

shows that also R^{-1} is a polynomial matrix. Therefore R is a unimodular polynomial matrix.

□

The matrix $\hat{A}$ turns out to be very important later in this thesis. It is called the structure matrix. The precise definition is given below.

Definition 2.9 The left Λ-structure matrix of a Λ-generalized polynomial matrix A is defined to be the polynomial matrix $\hat{A}$ in theorem 2.6.

2.4 Divisors

The divisors of Λ-generalized polynomial matrices will now be explored. They will be defined to be nonsquare matrices in general. Mac Duffee (1946) and Rosenbrock (1970) consider only square divisors.

Definition 2.10 If $A \in R_\Lambda^{n\times m}[\lambda]$ can be factorized as B C, where B and C are Λ-generalized polynomial matrices and B has linearly independent columns, then B is a left Λ-divisor of A.

Lemma 2.6 Let A and B belong to $R_\Lambda^{n\times m}[\lambda]$ and $R_\Lambda^{n\times k}[\lambda]$, respectively, and let r be the rank of $[A \quad B]$. Then there is a Λ-unimodular matrix $N \in R_\Lambda^{(m+k)\times(m+k)}[\lambda]$ such that

$$[A \quad B]\, N = [L \quad 0] \tag{2.27}$$

where $L \in R_\Lambda^{n\times r}[\lambda]$ and L has linearly independent columns.

Proof There are Λ-unimodular N and M such that

$$M[A \quad B]N = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \tag{2.28}$$

where the right member is the Λ-Smith form of [A B] and D is an r×r nonsingular matrix. Multiplying (2.28) from the left by M^{-1} gives (2.27) with

$$L = M^{-1} \begin{bmatrix} D \\ 0 \end{bmatrix} \tag{2.29}$$

□

Definition 2.11 Let A and B belong to $R_\Lambda^{n\times m}[\lambda]$ and $R_\Lambda^{n\times k}[\lambda]$, respectively. Let L belong to $R_\Lambda^{n\times \ell}[\lambda]$, for some ℓ. Then L is a greatest common left Λ-divisor (g.c.l.Λ.d.) to A and B if it is a left Λ-divisor to both A and B and if every other left Λ-divisor to A and B is also a left Λ-divisor to L.

Theorem 2.7 There exists a g.c.l.Λ.d. to every pair $A \in R_\Lambda^{n\times m}[\lambda]$ and $B \in R_\Lambda^{n\times k}[\lambda]$. If rank [A B] is r then any g.c.l.Λ.d. L has r columns and can be expressed as

$$L = AX + BY \tag{2.30}$$

for some $X \in R_\Lambda^{m\times r}[\lambda]$ and $Y \in R_\Lambda^{k\times r}[\lambda]$. Furthermore L is unique up to multiplication from the right by a Λ-unimodular matrix.

Proof By lemma 2.6 there is a Λ-unimodular N such that

$$[A \quad B]N = [L \quad 0] \tag{2.31}$$

where L has r linearly independent colums. It will be shown that L is a g.c.l.Λ.d to A and B. The equality (2.31) is equivalent to

$$[A \quad B] = [L \quad 0]N^{-1}. \tag{2.32}$$

Partition N^{-1} as

$$N^{-1} = \begin{bmatrix} \hat{N}_1 & \hat{N}_2 \\ \hat{N}_3 & \hat{N}_4 \end{bmatrix} \tag{2.33}$$

Then (2.32) gives

$$[A \quad B] = L[\hat{N}_1 \quad \hat{N}_2]. \tag{2.34}$$

Since L has linearly independent columns it is a left Λ-divisor of both A and B. Partition N as

$$N = \begin{bmatrix} N_1 & N_2 \\ N_3 & N_4 \end{bmatrix},$$

where N_1 is m×r. Then (2.31) gives

$$L = AN_1 + BN_3, \tag{2.36}$$

which is of the form (2.30). Let M be an arbitrary left Λ-divisor of A and B so that

$$A = M\, A_0 \tag{2.37 a}$$
$$B = M\, B_0. \tag{2.37 b}$$

Insertion of (2.37) into (2.36) gives

$$L = M(A_o N_1 + B_o N_3). \tag{2.38}$$

This shows that M is also a left Λ-divisor of L. It follows that L is a g.c.l.Λ.d. of A and B.

Let L_1 and L_2 be two g.c.l.Λ.d. of A and B. Since L_1 is a g.c.l.Λ.d. and L_2 is a left Λ-divisor we have

$$L_1 = L_2 U \tag{2.39}$$

for some Λ-generalized polynomial matrix U. Analogously

$$L_2 = L_1 V. \tag{2.40}$$

Equations (2.39) and (2.40) give

$$L_1 = L_1 V\ U \tag{2.41}$$
$$L_2 = L_2 U\ V \tag{2.42}$$

or since L_1 and L_2 have linearly independent columns

$$V\ U = I \tag{2.43}$$
$$U\ V = I. \tag{2.44}$$

Consequently U and V are both square and Λ-unimodular.

It has thus been shown that the g.c.l.Λ.d is unique up to multiplication from the right by a Λ-unimodular matrix. Since there is one g.c.l.Λ.d with r columns which can be expressed as (2.30) this is true for every g.c.l.Λ.d.

□

Definition 2.12 $A \in R_\Lambda^{n\times m}[\lambda]$ and $B \in R_\Lambda^{n\times k}[\lambda]$ are relatively left Λ-prime if the g.c.l.Λ.d. is Λ-unimodular.

Theorem 2.8 Let A and B belong to $R_\Lambda^{n\times m}[\lambda]$ and $R_\Lambda^{n\times k}[\lambda]$, respectively. The following statements are equivalent.

(i) A and B are relatively left Λ-prime.

(ii) The Λ-Smithform of [A B] is [I 0].

(iii) [A B] has linearly independent rows for every $\lambda_o \in \Lambda$.

Remark If $\Lambda = \phi$, i.e. A and B are rational matrices, then (iii) should be substituted by:

(iii)' [A B] has linearly independent rows in the sense of definition 2.6.

Proof (i) ⇔ (ii). By lemma 2.6 there are N and L such that

$$[A\quad B]N = [L\quad 0]$$

where N is Λ-unimodular and L has linearly independent columns. By the proof of theorem 2.7 L is the g.c.l.Λ.d of A and B. Let S be the Λ-Smithform of L. Then [S 0] is the Λ-Smithform of

[A B] and this is equal to [I 0] if and only if the Λ-Smithform of L is I, i.e. if and only if L is Λ-unimodular.
The equivalence between (ii) and (iii) follows from theorem 2.4.
□

2.5 Matrix equations

The treatment of the feedback problem in chapter 6 and 7 leads to two different kinds of Λ-generalized polynomial matrix equations. Variants of the corresponding equations for polynomial matrices have been studied by many authors, e.g. Mac Duffee (1946), Rosenbrock (1970), Roth (1952). In this section the results will be stated in a form suitable for chapter 7 and 8.

Theorem 2.9 Let A, B and C belong to $R_\Lambda^{n\times m}[\lambda]$, $R_\Lambda^{n\times k}[\lambda]$ and $R_\Lambda^{n\times \ell}[\lambda]$, respectively. There is a solution $X \in R_\Lambda^{m\times \ell}[\lambda]$ and $Y \in R_\Lambda^{k\times \ell}[\lambda]$ to the equation

$$C = AX + BY \tag{2.45}$$

if and only if the g.c.l.Λ.d. of A and B is also a left Λ-divisor of C.

Proof Let L be the g.c.l.Λ.d. of A and B. Suppose that L is a left Λ-divisor of C so that

$$C = L\,C_o \tag{2.46}$$

By theorem 2.7 there are X_o and Y_o such that

$$L = AX_o + BY_o. \tag{2.47}$$

Multiplying (2.47) from the right by C_o gives (2.45) with $X = X_oC_o$ and $Y = Y_oC_o$.

Conversely suppose that (2.45) has a solution. Since L is the g.c.l.Λ.d. of A and B there are A_o and B_o such that

$$A = LA_o \tag{2.48 a}$$

$$B = LB_o \tag{2.48 b}$$

Insert (2.48) into (2.45)

$$C = L(A_oX + B_oY) \tag{2.49}$$

and L is a left Λ-divisor of C.

□

Corollary Let $A \in R_\Lambda^{n\times m}[\lambda]$ and $B \in R_\Lambda^{n\times k}[\lambda]$ be given. The equation

$$I = AX + BY \tag{2.50}$$

has a solution $X \in R_\Lambda^{m\times n}[\lambda]$ and $Y \in R_\Lambda^{k\times n}[\lambda]$ if and only if A and B are relatively left Λ-prime.

If the equation (2.45) has a solution then it is interesting to find all solutions. Since the equation is linear all solutions $X \in R_\Lambda^{m\times \ell}[\lambda]$ and $Y \in R_\Lambda^{k\times \ell}[\lambda]$ to the homogenous equation

$$AX + BY = 0 \tag{2.51}$$

are needed. This problem was solved for polynomial matrices by Forney (1975) via the introduction of a minimal polynomial basis. Our problem can be solved analogously. A polynomial basis is first defined. A comparison with Forney´s definition shows that every minimal polynomial basis is a polynomial basis.

Definition 2.13 Let $Z \in R^{n\times m}(\lambda)$ have rank r. $Q \in R^{m\times(m-r)}[\lambda]$ is a polynomial basis for the nullspace of Z if $ZQ = 0$ and the Smith-form of Q is $[I_{m-r} \quad 0]^T$.

Remark Forney has shown that there exists a minimal polynomial basis for the nullspace of every $Z \in R^{n\times m}(\lambda)$. Consequently there exists a polynomial basis for the nullspace of every $Z \in R^{n\times m}(\lambda)$.

Theorem 2.10 Let $Z \in R^{n\times m}(\lambda)$ have rank r and let Q be a polynomial basis for the nullspace of Z. Any solution $P \in R_{\Lambda}^{m \times k}[\lambda]$ to

$$Z\ P = 0 \tag{2.52}$$

can be written

$$P = Q\ R \tag{2.53}$$

for some R in $R_{\Lambda}^{(m-r) \times k}[\lambda]$.

Proof Since the rational matrix Z has m columns and rank r then the nullspace has dimension m-r. Regard the polynomial matrix Q as a rational matrix. Q has m-r columns, which are linearly independent since the Smithform is $[I_{m-r} \quad 0]^T$, and Q is contained in the nullspace of Z. Therefore Q is a basis for the nullspace of Z. Consequently P can be written as (2.53) for some rational matrix R. It remains to be shown that R is in fact a Λ-generalized polynomial matrix.

There are unimodular polynomial matrices M and N such that

$$M\ Q\ N = \begin{bmatrix} I \\ 0 \end{bmatrix}. \tag{2.54}$$

Introduction of (2.54) into (2.53) gives

$$P = M^{-1}\begin{bmatrix} I \\ 0 \end{bmatrix}N^{-1}R$$

$$\Leftrightarrow \quad M\ P = \begin{bmatrix} N^{-1} & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} I \\ 0 \end{bmatrix} R$$

$$\Leftrightarrow \quad \begin{bmatrix} N & 0 \\ 0 & I \end{bmatrix}M\ P = \begin{bmatrix} R \\ 0 \end{bmatrix} \tag{2.55}$$

Since the left member of (2.55) is a Λ-generalized polynomial matrix this is also true for R. □

This theorem can now be used to compute all solutions to the equation (2.45).

<u>Theorem 2.11</u> Let A, B and C belong to $R_{\Lambda}^{n\times m}[\lambda]$, $R_{\Lambda}^{n\times k}[\lambda]$ and $R_{\Lambda}^{n\times \ell}[\lambda]$, respectively and suppose that the equation

$$C = AX + BY \tag{2.56}$$

has a solution $X = X_o \in R_{\Lambda}^{m\times \ell}[\lambda]$ and $Y = Y_o \in R_{\Lambda}^{k\times \ell}[\lambda]$. Let $\begin{bmatrix} P \\ Q \end{bmatrix}$ be a polynomial basis for the nullspace of $[A \quad B]$. Then any solution X in $R_{\Lambda}^{m\times \ell}[\lambda]$ and Y in $R_{\Lambda}^{k\times \ell}[\lambda]$ can be written

$$X = X_o + P\ R \tag{2.57 a}$$

$$Y = Y_o + Q\ R \tag{2.57 b}$$

for some R in $R_{\Lambda}^{q\times \ell}[\lambda]$, where $q = m + k - r$ and r is the rank of $[A \quad B]$.

<u>Proof</u> Any solution to

$$AX + BY = 0$$

$$\Leftrightarrow [A \quad B]\begin{bmatrix} X \\ Y \end{bmatrix} = 0 \tag{2.58}$$

can by theorem 2.10 be written

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} P \\ Q \end{bmatrix} R \tag{2.59}$$

for some Λ-generalized polynomial matrix R. □

The following linear equation, which resembles the equation (2.45) is also of interest.

$$C = AX + YB \tag{2.60}$$

This equation has been studied by Roth (1952) in the case of polynomial matrices. Roth gives necessary and sufficient conditions for the existence of a solution. The properties of the polynomials that Roth uses in his proof are properties of any principal ideal domain. Therefore his proof will be valid for matrices with entries in any principal ideal domain. We state the theorem for matrices with entries in the ring of Λ-generalized polynomials.

Theorem 2.12 Let A, B and C belong to $R_\Lambda^{n\times m}[\lambda]$, $R_\Lambda^{k\times \ell}[\lambda]$ and $R_\Lambda^{n\times \ell}[\lambda]$, respectively. The equation

$$C = AX + YB \tag{2.61}$$

has a solution $X \in R_\Lambda^{m\times \ell}[\lambda]$ and $Y \in R_\Lambda^{n\times k}[\lambda]$ if and only if the matrices

$$\begin{pmatrix} A & C \\ O & B \end{pmatrix} \text{ and } \begin{pmatrix} A & O \\ O & B \end{pmatrix} \tag{2.62}$$

are Λ-equivalent.

2.6 Duality

The dual forms of the definitions, lemmas och theorems in this chapter are easily obtained by applying them to the transposed matrices. L is for instance a greatest common right Λ-divisor (g.c.r.Λ.d) to A and B if and only if L^T is a g.c.l.Λ.d to A^T and B^T. The equation C = XA + YB has a solution if and only if the equation $C^T = A^T X^T + B^T Y^T$ has a solution. $\hat{A}$ is a right Λ-structure matrix of A if and only if $\hat{A}^T$ is a left Λ-structure matrix of A^T. Other dual forms are obtained analogously and will not be listed here.

3. THE SYSTEM DESCRIPTION

Consider the system S, represented by the block diagram in figure 3.1.

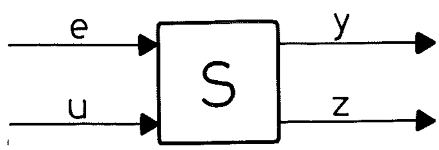


Figure 3.1 The system to be controlled.

The system is multivariable with two input vectors u and e and two output vectors y and z. The components of u are the control variables and the components of e are the disturbances that act on the system. It is in general supposed that the disturbance vector e can not be measured. The output vector y contains all variables to be controlled and z all variables that are measured. The vectors y and z may have some components in common.

The system S is assumed to be a linear, time invariant, finite dimensional, causal system. Furthermore, the system is supposed to be stabilizable from u, detectable from z and have no direct feedthrough from u to z. The controller R for the system is required to be a linear, time invariant, finite dimensional, causal controller that stabilizes the system. It will be connected to the system according to figure 3.2, where u_r is the command input.

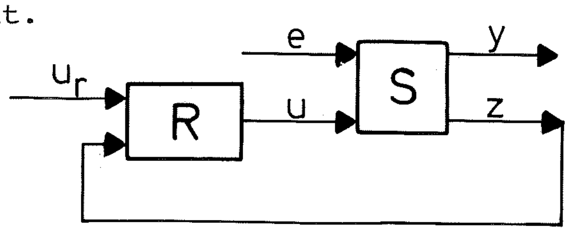


Figure 3.2 The closed loop system.

The assumptions on the system will be discussed in this chapter. A mathematical description of the system will be presented and the assumptions will be stated in algebraic terms. Furthermore, two kinds of structure matrices will be defined for the system S. The left and right structure matrices for a generalized polynomial matrix were defined in chapter 2. These definitions are used to define the two structure matrices for a dynamical system. The left structure matrix for a system is associated with the signal transmission from u to y, while the right structure matrix is associated with the transmission from e to z. The structure matrices turn out to be important tools for design. They contain all the relevant information about the nonminimum phase properties of the system. In the discrete time case they also contain information about the time delay structure of the system. In the continuous time case the structure matrix contains the corresponding information, which may be called the structure of inherent integrations c.f. Sain, Massey (1969).

3.1 An input-output description of the system

The closed loop system in figure 3.2 is supposed to work over an infinite time interval and the transient behaviour right after startup is not considered important. Since the closed loop system is required to be stable the effect of initial conditions will vanish quickly. There is therefore no need to include the initial conditions in the system description. Observe that this viewpoint does not imply that the design is made for a steady state situation. The disturbance e and the command input u_r may vary arbitrarily.

As pointed out previously the system S is assumed to be stabilizable from u and detectable from z. These assumptions must be satisfied, because otherwise it is not possible to achieve a stable closed loop system. Introduce the total input vector ν and the total output vector η as

$$\nu = \begin{pmatrix} u \\ e \end{pmatrix} \qquad \eta = \begin{pmatrix} y \\ z \end{pmatrix} \tag{3.1}$$

The part of the system S that is not controllable from ν is by the stabilizability assumption stable. The corresponding variables will quickly go to zero and the uncontrollable part of the system does not have to be included in the system description. The part of the system S that is not observable from η is by the detectability assumption stable. The corresponding variables will be bounded and do not affect y. The unobservable part of the system therefore does not have to be included in the system description.

To summerize, we have concluded that the linear, time invariant, finite dimensional system S, with input ν and output η , is appropriately described by its controllable and observable part. Furthermore the initial conditions can be assumed to be zero. It is well known that such a system can be described by a rational transfer operator.

Let the transfer operator from ν to η be denoted $G^*(\mu)$. Then $G^*(\mu)$ is a matrix with entries that are rational functions. In the continuous time case μ is the differential operator, defined through

$$px = \frac{dx}{dt}. \tag{3.2}$$

In the discrete time case μ is the forward shift operator q, defined through

$$qx(t) = x(t + 1). \tag{3.3}$$

The problem of defining a suitable class of time functions for $G^*(\mu)$ to operate on will not be treated here. A detailed discussion of appropriate function classes for discrete time systems is given in part 1 of this thesis. Here it is merely supposed that such a class is already defined and that the usual isomorfism between the transfer operators and transfer functions, i.e. the corresponding matrices of a complex variable, is valid. Addition and multiplication of transfer functions correspond to addition and composition of transfer operators. Because of this isomorfism $G^*(\mu)$ will sometimes be regarded as a transfer operator and sometimes as a transfer function. Consequently μ is sometimes regarded as an operator and sometimes as a complex number. No notational difference will be made between these two interpretations of μ and $G^*(\mu)$.

3.2 Stability and causality

Usually a continuous time transfer function $G^*(p)$ is said to be stable if it has no poles in the closed right half plane. Analogously a discrete time transfer function $G^*(q)$ is said to be stable if it has no poles outside the unit disc. In general we can define an unstable region Λ^* and say that a transfer function $G^*(\mu)$ is stable if it has no poles in Λ^*. The unstable region Λ^* will be allowed to be a quite general region in which it is undesirable for the transfer function to have poles. Two examples for continuous time systems are given in figure 3.3. The corresponding regions for discrete time are shown in figure 3.4.

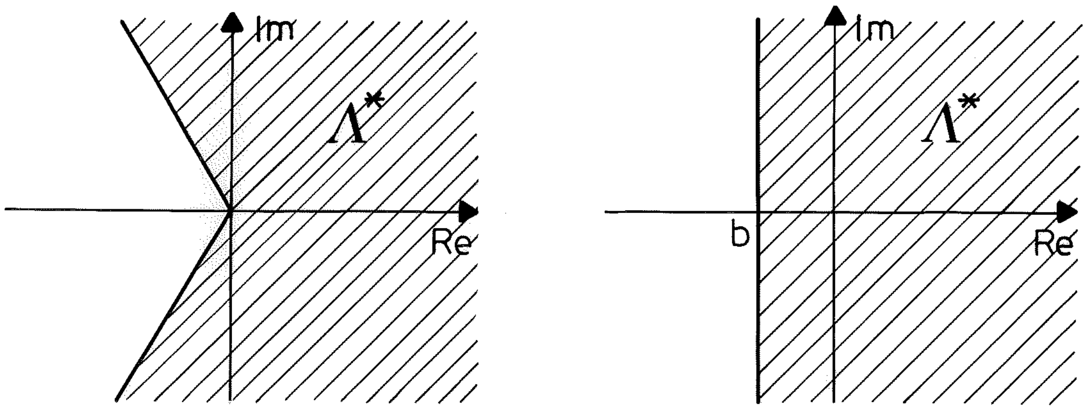


Figure 3.3. Possible choices of unstable regions Λ^* for a continuous time system.

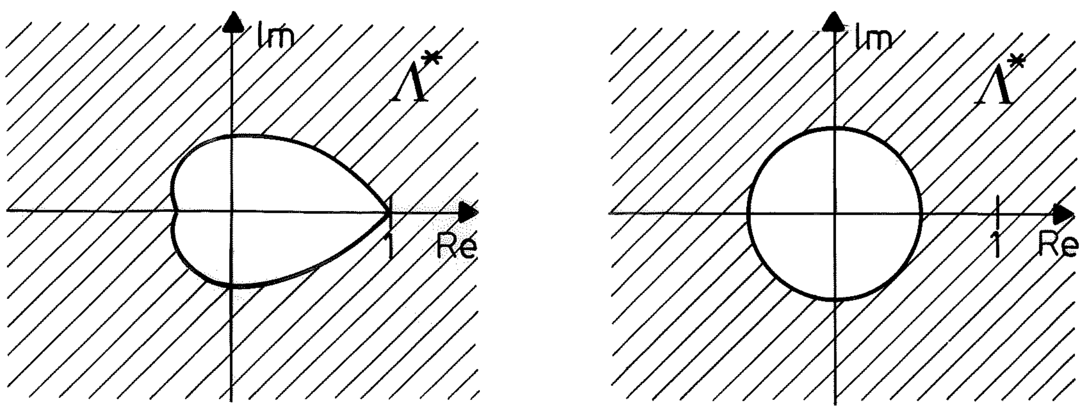


Figure 3.4. Possible choices of unstable regions Λ^* for a discrete time system.

A system can not always be appropriately described by its transfer function but has to be described in polynomial matrix form, Rosenbrock (1970), Wolovich (1974). In such a case not only the poles of its transfer function, but also its decoupling zeros, i.e. uncontrollable or unobservable modes, have to be outside Λ^* for the system to be stable. We have concluded that our system S and our controller can be described by their transfer functions, but this is not sufficient for the closed loop system. The reason is that some variables may become uncontrollable or unobservable when the system and controller are connected. The closed loop system thus has to be described in polynomial matrix form.

Causality means that the output does not depend on future values of the input. For discrete time systems causality is then obviously equivalent to $\lim_{\mu\to\infty} G^*(\mu)$ being finite. With some abuse of language we will also say that a continuous time system is causal if $\lim_{\mu\to\infty} G^*(\mu)$ is finite. This means that the output of a causal system does not depend on the derivatives of the input.

Causality is obviously a necessary requirement on any transfer function, that describes a physically realizable discrete time system. It is also a necessary requirement on any transfer function, that describes a physically realizable continuous time system because pure differentiators do not exist in practice.

A system is thus said to be causal of $\lim_{\mu\to\infty} G^*(\mu)$ if finite. In other words the system is causal if $G^*(\mu)$ has no pole at infinity. On the other hand a system, which can be described by its transfer function, is said to be stable if $G^*(\mu)$ has no poles in Λ^*. Let the "infinity point" belong to Λ^*. It is then possible to treat stability and causality simultaneously by saying that a system is stable and causal if $G^*(\mu)$ has no poles in Λ^*.

The unstable region Λ^* is required to fulfil the following conditions:

- Λ^* is symmetric with respect to the real axis. (3.4)

- There is at least one point on the real axis, that does not belong to Λ^*. (3.5)
- Λ^* contains the "infinity point". (3.6)

Condition (3.4) is required because the pole and zero configuration of a rational function with real coefficients, like the transfer function, is symmetric with respect to the real axis. Condition (3.5) ensures that also systems with an odd number of poles can be stable, because such a system must have at least one real pole. Finally, (3.6) implies that a system with no poles in Λ^* is both stable and causal.

It would be desirable to use the theory for Λ-generalized polynomial matrices to treat stable and causal transfer functions. There is, however, one problem. The region Λ is, in chapter 2 required to be a subset of C. It is thus not allowed to include the "infinity point". The theory in chapter 2 can be changed so that Λ includes the "infinity point". This would, however, make the theory more complicated. Furthermore, there would be no ordinary polynomials included in the set of Λ-generalized polynomials for such a Λ. This would be a great disadvantage, because the ordinary polynomials are more practical to work with. In section 2.3 the structure matrix is for instance defined as a polynomial matrix. The region Λ will therefore not be allowed to include the "infinity point". The problem will instead be solved in the following way. The region Λ^* is mapped onto a region, which does not include the "infinity point".

Let a be a point on the real axis in the complement of Λ^* and define the mapping

$$\lambda = f(\mu) = \frac{1}{\mu - a} \tag{3.7}$$

Let $\bar{C}$ be the complex plane C extended with the "infinity point". Then f is a bijective mapping from $\bar{C}$ to $\bar{C}$. It maps the "infinity point" to the origin and a to the "infinity point". Define Λ through

$$\Lambda = f(\Lambda^*) \tag{3.8}$$

Then Λ has the following properties.

- It is symmetric with respect to the real axis. (3.9)
- It does not contain the "infinity point". (3.10)
- It contains the origin. (3.11)

These properties follow directly from (3.4)-(3.6) and the definition of f in (3.7).

Define $G(\lambda)$ through

$$G(\lambda) = G^*(\mu), \tag{3.12}$$

where λ and μ are connected through (3.7). Then $G(\lambda)$ has no poles in Λ if and only if $G^*(\mu)$ has no poles in Λ^*. In other words the system is causal and stable if and only if $G(\lambda)$ is a Λ-generalized polynomial matrix. We will in the rest of this thesis work with $G(\lambda)$ rather than $G^*(\mu)$. The great advantage is that stability and causality can be treated with an unstable region Λ which does not contain the "infinity point". The theory in chapter 2 is thus directly applicable.

The region Λ corresponding to Λ^* in figure 3.3 and 3.4 are shown in figure 3.5 and 3.6.

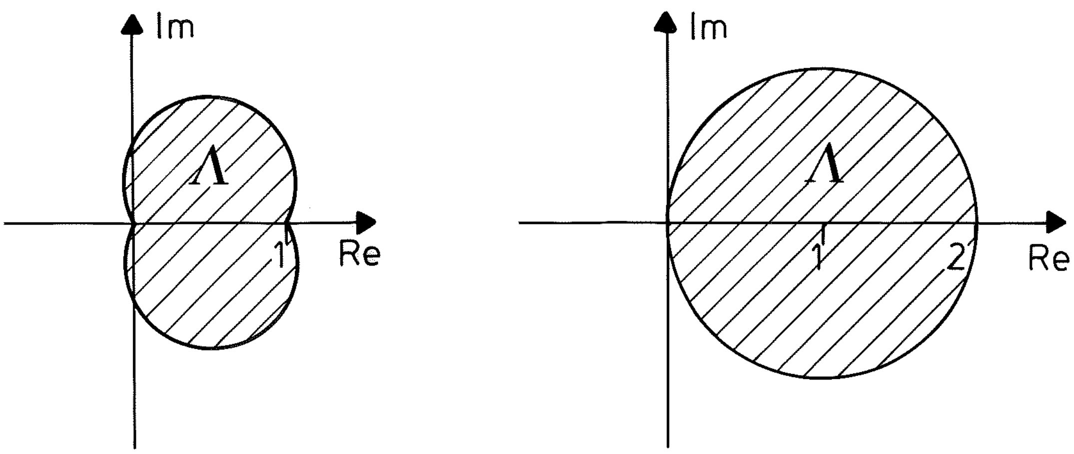


Figure 3.5. The regions Λ corresponding to Λ^* in figure 3.3 with $a = -1$ and $b = -\frac{1}{2}$.

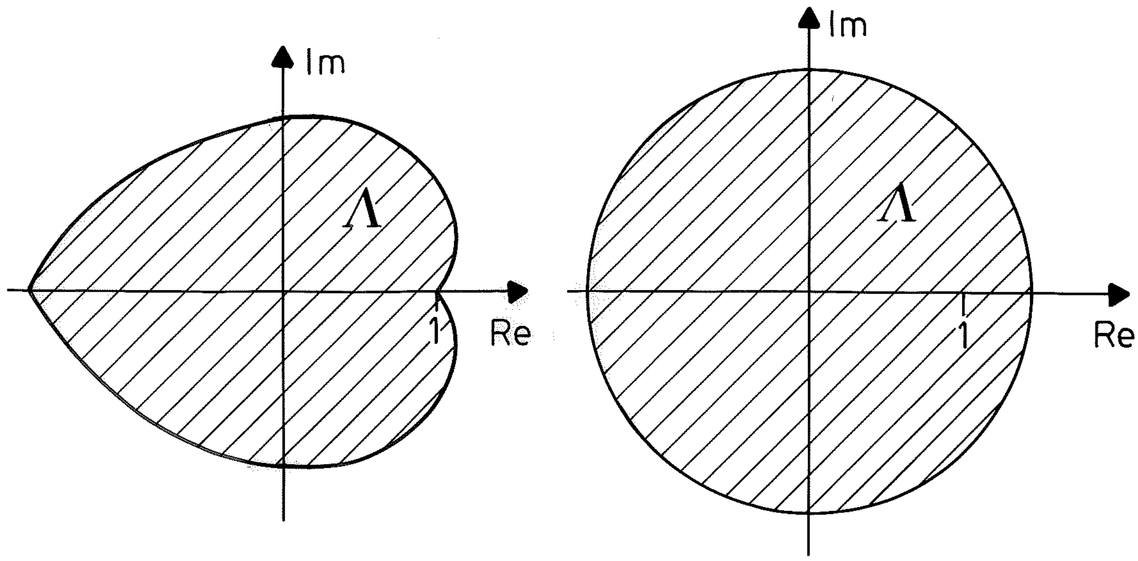


Figure 3.6. The regions Λ corresponding to Λ* in figure 3.4 with a = 0.

In most practical cases a will be chosen as an inner point of the complement of Λ*. Then Λ will be bounded as it is in figure 3.5 and 3.6.

G(λ) will be called the transfer function for the system. Of course G(λ) depends on a, but a is a stable point and we will show that our results are in fact independent of the choice of a.

The mapping (3.7) maps the "infinity point" to the origin independently of a. A system is therefore causal if and only if G(0) is finite and this condition is independent of a. A system, described by its transfer function G(λ), will be said to be Λ-stable if G(λ) has no poles in Λ.

3.3 Fractional representations.

Let G(λ) be the transfer function, defined through (3.12) and (3.7) with same real a, for the system S. Then we have

$$\eta = G(\lambda)\nu. \tag{3.13}$$

It is shown e.g. in Rosenbrock (1970) that $G(\lambda)$ can be factorized as $G(\lambda) = T^{-1}(\lambda)\, U(\lambda)$, where $T(\lambda)$ and $U(\lambda)$ are relatively left prime polynomial matrices and $T(\lambda)$ is square with nonzero determinant. Furthermore the zeros of det $T(\lambda)$ are the poles of $G(\lambda)$. Because $G(\lambda)$ is causal it follows that

$$\det T(0) \neq 0 \tag{3.14}$$

If the function spaces for ν and η are defined properly (3.13) is equivalent to

$$T(\lambda)\eta = U(\lambda)\nu. \tag{3.15}$$

This is discussed in detail for discrete time systems in part 1 of this thesis. The representation (3.15) is unique up to multiplication from the left by unimodular matrices. Now, instead of a in (3.7), choose another real number $\bar{a}$ outside Λ^*.
Define

$$\bar{\lambda} = \frac{1}{\mu - \bar{a}} \tag{3.16}$$

and let $\bar{G}(\bar{\lambda})$ be the corresponding transfer function. Then

$$\bar{\lambda} = \frac{\lambda}{1 - \lambda(\bar{a} - a)} \tag{3.17}$$

and

$$\bar{G}(\bar{\lambda}) = G(\lambda) \tag{3.18}$$

In analogy with (3.15) we have

$$\bar{T}(\bar{\lambda})\eta = \bar{U}(\bar{\lambda})\nu, \tag{3.19}$$

where $\bar{T}(\bar{\lambda})$ and $\bar{U}(\bar{\lambda})$ fulfil the same conditions as $T(\lambda)$ and $U(\lambda)$. We will show how the representations (3.15) and (3.19) are related. For this we need the following lemma.

Lemma 3.1 Let Λ be a subset of C and let A belong to $R_\Lambda^{r\times r}[\lambda]$ and have nonzero determinant. Let B belong to $R_\Lambda^{r\times k}[\lambda]$ and suppose that A and B are relatively left Λ-prime. Furthermore, let A_1 and B_1 fulfil the same conditions and assume that

$$A^{-1}B = A_1^{-1}B_1 \tag{3.20}$$

Then there is a Λ-unimodular matrix $Q \in R_\Lambda^{r\times r}[\lambda]$, such that

$$A_1 = Q\,A \tag{3.21 a}$$

$$B_1 = Q\,B\,. \tag{3.21 b}$$

Proof By the corollary of theorem 2.9 there are $X \in R_\Lambda^{r\times r}[\lambda]$ and $Y \in R_\Lambda^{k\times r}[\lambda]$, such that

$$\begin{aligned} & I = AX + BY \\ \Leftrightarrow\ & A^{-1} = X + A^{-1}B\ Y \\ \Leftrightarrow\ & A^{-1} = X + A_1^{-1}B_1Y \\ \Leftrightarrow\ & A_1A^{-1} = A_1X + B_1Y \end{aligned} \tag{3.22}$$

Define

$$Q = A_1A^{-1} \tag{3.23}$$

It follows from (3.22) that $Q \in R_\Lambda^{r\times r}[\lambda]$. Analogously we have that $Q^{-1} = A\,A_1^{-1} \in R_\Lambda^{r\times r}[\lambda]$. Therefore Q is Λ-unimodular and from (3.23)

$$A_1 = Q\ A, \tag{3.24}$$

which is (3.21 a). Introduce (3.24) into (3.20)

$$\begin{aligned} & A^{-1}B = A^{-1}Q^{-1}B_1 \\ \Leftrightarrow\ & B = Q^{-1}B_1 \\ \Leftrightarrow\ & B_1 = Q\,B\,, \end{aligned} \tag{3.25}$$

which is (3.21 b). □

Theorem 3.1 Consider the two representations (3.15) and (3.19). There is a Λ-unimodular matrix Q, such that

$$\bar{T}(\bar{\lambda}) = Q(\lambda)T(\lambda) \tag{3.26 a}$$

$$\bar{U}(\bar{\lambda}) = Q(\lambda)U(\lambda), \tag{3.26 b}$$

where $\bar{\lambda}$ is given by

$$\bar{\lambda} = \frac{\lambda}{1-\lambda(\bar{a}-a)} \tag{3.27}$$

Proof Since $T(\lambda)$ and $U(\lambda)$ are relatively left prime the matrix $[T(\lambda_o)\ U(\lambda_o)]$ has linearly independent rows for all $\lambda_o \in C$. It then follows by theorem 2.8 ((iii) ⇒ (i)) that they are relatively left Λ-prime.

From (3.27) it follows that $\bar{\lambda}$, as a function of λ, has a pole at $\lambda_1 = \frac{1}{\bar{a}-a}$. This λ_1 does not belong to Λ, by (3.7) and (3.8), since $\bar{a}$ does not belong to Λ^*. Define

$$T_o(\lambda) = \bar{T}\left(\frac{\lambda}{1-\lambda(\bar{a}-a)}\right) \tag{3.28 a}$$

$$U_o(\lambda) = \bar{U}\left(\frac{\lambda}{1-\lambda(\bar{a}-a)}\right) \tag{3.28 b}$$

Then $T_o(\lambda)$ and $U_o(\lambda)$ are Λ-generalized polynomial matrices since they are rational matrices and their only pole is outside Λ.

Let λ_2 be an arbitrary element in Λ and define $\bar{\lambda}_2$ through (3.27) with $\lambda = \lambda_2$. Then $\bar{\lambda}_2$ is finite and $[T_o(\lambda_2)\ U_o(\lambda_2)] = [\bar{T}(\bar{\lambda}_2)\ \bar{U}(\bar{\lambda}_2)]$ has linearly independent rows since $\bar{T}(\bar{\lambda})$ and $\bar{U}(\bar{\lambda})$ are relatively left prime. It follows that $T_o(\lambda)$ and $U_o(\lambda)$ are relatively left Λ-prime.

By (3.18) we have that

$$T_o^{-1}(\lambda)U_o(\lambda) = T^{-1}(\lambda)U(\lambda) \tag{3.29}$$

It now follows by lemma 3.1 that there is a Λ-unimodular $Q(\lambda)$ such that

$$T_o(\lambda) = Q(\lambda)T(\lambda) \tag{3.30 a}$$

$$U_o(\lambda) = Q(\lambda)U(\lambda) \tag{3.30 b}$$

But (3.30) is equivalent to (3.26). □

Theorem 3.1 gives the desired connection between two representations of the type (3.15), corresponding to the same Λ^*, but different a.

The representation (3.15) is called a fractional representation of $G(\lambda)$ and there is a dual form of it.

$$P(\lambda)\xi = \nu \tag{3.31 a}$$

$$\eta = V(\lambda)\xi \tag{3.31 b}$$

Here $P(\lambda)$ and $V(\lambda)$ are relatively right prime polynomial matrices and $P(\lambda)$ is square and has a nonzero determinant. The zeros of $P(\lambda)$ are the poles of $G(\lambda)$. The dual form of theorem 3.1 is of course valid. In other words let

$$\bar{P}(\bar{\lambda})\bar{\xi} = \nu \tag{3.32 a}$$

$$\eta = \bar{V}(\bar{\lambda})\bar{\xi}$$

be another polynomial fractional representation of the same system with the same Λ^* but with another a-value, say $\bar{a}$. Then there is a Λ-unimodular $Z(\lambda)$, such that

$$\bar{P}(\bar{\lambda}) = P(\lambda)Z(\lambda) \tag{3.33 a}$$

$$\bar{V}(\bar{\lambda}) = V(\lambda)Z(\lambda) \tag{3.33 b}$$

with $\bar{\lambda}$ given by (3.27).

Finally we will show how the two different types of fractional representations (3.15) and (3.31), corresponding to the same system, Λ^* and a, are related. This problem is solved in Rosenbrock (1970). The system matrices are strictly system equivalent (s.s.e.) i.e. there are polynomial matrices $M(\lambda)$, $N(\lambda)$, $X(\lambda)$ and $Y(\lambda)$, where $M(\lambda)$ and $N(\lambda)$ are unimodular, such that

$$\left(\begin{array}{cc|c} I_{k-r} & 0 & 0 \\ 0 & P & I \\ \hline 0 & -V & 0 \end{array}\right) = \begin{pmatrix} M & 0 \\ X & I \end{pmatrix} \left(\begin{array}{cc|c} I_{k-s} & 0 & 0 \\ 0 & T & U \\ \hline 0 & -I & 0 \end{array}\right) \begin{pmatrix} N & Y \\ O & I \end{pmatrix} \quad (3.34)$$

for some $k \geq n$, where $n = \deg \det T(\lambda) = \det \det P(\lambda)$. Here r and s are the dimensions of P and T, respectively.

3.4 Assumptions on the system to be controlled

It has already been stated that the system S as well as its controllers are supposed to be linear, time invariant, finite dimensional, causal systems. It has also been concluded that they are adequately described by their transfer function matrices. Let $G(\lambda)$, where λ is given by (3.7), be the transfer function for S or any controller. Then consequently the elements of $G(\lambda)$ are supposed to be rationel functions with constant coefficients. Furthermore, by the causality requirement, $G(\lambda)$ has no poles at the origin.

In addition to these general assumptions the system S, but not its controllers, is supposed to fulfil the following three assumptions.

- There is no direct feedthrough from u to z. (3.35)
- The system is Λ-stabilizable from u. (3.36)
- The system is Λ-detectable from z. (3.37)

The condition (3.35) is not necessary, but on the other hand it is considered as no serious restriction. The assumption (3.35) prevents algebraic loops when feedback is applied from z to u. The conditions (3.36) and (3.37) are necessary because the controlled system is required to be Λ-stable.

The conditions (3.35)-(3.37) will now be made more precise. Let the transfer function G(λ), where $\lambda = (\mu-a)^{-1}$ and μ is p of q, be partitioned as

$$\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} G_{uy}(\lambda) & G_{ey}(\lambda) \\ G_{uz}(\lambda) & G_{ez}(\lambda) \end{pmatrix} \begin{pmatrix} u \\ e \end{pmatrix} \tag{3.38}$$

<u>Definition 3.1</u> The system S has no direct feedthrough from u to z if

$$G_{uz}(0) = 0 \tag{3.39}$$

<u>Remark</u> Let the transfer function G*(μ), where μ is p or q, be partitioned in the same way. Then (3.39) is equivalent to

$$\lim_{\mu \to \infty} G^*_{uz}(\mu) = 0, \tag{3.40}$$

which is independent of a. Therefore (3.39) is independent of a.

As pointed out in section 3.3 the following two fractional representations will be used to represent the transfer function G(λ).

$$A(\lambda)\eta = B(\lambda)u + C(\lambda)e \tag{3.41}$$

and

$$D(\lambda)\xi = \nu \tag{3.42 a}$$

$$y = E(\lambda)\xi \tag{3.42 b}$$

$$z = F(\lambda)\xi \tag{3.42 c}$$

Furthermore the corresponding fractional representations with other a-values will be used.

In (3.41) $A(\lambda)$ and $[B(\lambda)\ C(\lambda)]$ are relatively left prime polynomial matrices with $A(\lambda)$ square and nonsingular. It was shown in section 3.3 that all representations of the type (3.41), with the same a-value, can be obtained by multiplication of (3.41) from the left by unimodular matrices. In particular there is a representation with $A(\lambda)$ upper block triangular.

$$\begin{pmatrix} A_1(\lambda) & A_3(\lambda) \\ 0 & A_2(\lambda) \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} B_1(\lambda) \\ B_2(\lambda) \end{pmatrix} u + \begin{pmatrix} C_1(\lambda) \\ C_2(\lambda) \end{pmatrix} e \tag{3.43}$$

Theorem 3.2 Let the system S be represented by (3.43). The system has no direct feedthrough from u to z if and only if

$$B_2(0) = 0 \tag{3.44}$$

This condition is independent of the a-value used in (3.43).

Proof From (3.43) it follows that

$$G_{uz}(\lambda) = A_2^{-1}(\lambda) B_2(\lambda) \tag{3.45}$$

Since the system is causal it follows from (3.14) that

$$\det A_2(0) \neq 0. \tag{3.46}$$

Therefore $A_2^{-1}(0)$ exists and is nonsingular and (3.45) shows that (3.44) is equivalent to (3.39). Furthermore (3.39) has been shown to be independent of a.

□

The dual form of theorem 3.2 is obtained from the representation (3.42).

The two fractional representations (3.41) and (3.42) are both special cases of the polynomial matrix representation

$$T(\lambda)\xi = U_1(\lambda)u + U_2(\lambda)e \tag{3.47 a}$$

$$y = V_1(\lambda)\xi \tag{3.47 b}$$

$$z = V_2(\lambda)\xi \tag{3.47 c}$$

Here $T(\lambda)$ is square and nonsingular. Furthermore $T(\lambda)$ and $[U_1(\lambda)\ U_2(\lambda)]$ are relatively left prime and $T(\lambda)$ and $[V_1^T(\lambda)\ V_2^T(\lambda)]^T$ are relatively right prime.

<u>Definition 3.2</u> Let $L(\lambda)$ be the g.c.l.d of $T(\lambda)$ and $U_1(\lambda)$. The system S is Λ-stabilizable from u if det $L(\lambda)$ has no zeros in Λ.

<u>Remark</u> The definition means that all modes, that are uncontrollable from u, are Λ-stable if the system is Λ-stabilizable from u. This is discussed in detail for discrete time systems in part 1 of this thesis.

It has to be shown that the property of Λ-stabilizability is well defined, i.e. that it does not depend on which of the two representations (3.41) and (3.42) and which a-value that is used.

As pointed out previously all representations of the type (3.41) with the same a-value are obtained by multiplication of (3.41) from the left by unimodular matrices. This clearly does not affect the zeros of the determinant of the g.c.l.d. and therefore it does not affect the property of Λ-stabilizability.

The following lemma gives an alternative way to characterize the property of Λ-stabilizability.

<u>Lemma 3.2</u> Consider the representation (3.47). The system S is Λ-stabilizable from u if and only if $[T(\lambda_o)\ U_1(\lambda_o)]$ has linearly independent rows for all $\lambda_o \in \Lambda$.

<u>Proof</u> Let $T_o(\lambda)$ and $U_o(\lambda)$ be relatively left prime polynomial matrices such that $T(\lambda) = L(\lambda)T_o(\lambda)$ and $U_1(\lambda) = L(\lambda)U_o(\lambda)$, where

$L(\lambda)$ is the g.c.l.d of $T(\lambda)$ and $U_1(\lambda)$. The result follows from

$$[T(\lambda) \quad U_1(\lambda)] = L(\lambda)[T_o(\lambda) \quad U_o(\lambda)] \tag{3.48}$$

since $[T_o(\lambda_o) \quad U_o(\lambda_o)]$ has linearly independent rows for all $\lambda_o \in C$ by theorem 2.8 applied to polynomial matrices. □

Theorem 3.3 The representation (3.42) is Λ-stabilizable from u if and only if (3.41) is.

Proof By (3.34) it follows that

$$\left[D \begin{pmatrix} I_\ell \\ 0 \end{pmatrix} \right] = M\,[A \quad B] \begin{bmatrix} N & Y_1 \\ 0 & I_\ell \end{bmatrix}, \tag{3.49}$$

where ℓ is the dimension of u and Y_1 is the ℓ first columns of Y. The result now follows from lemma 3.2. □

The following theorem shows that the property of Λ-stabilizability is independent of a.

Theorem 3.4 The representation (3.15) is Λ-stabilizable if and only if (3.19) is $\bar{\Lambda}$-stabilizable, where $\bar{\Lambda}$ is the corresponding unstable region for (3.19).

Proof By theorem 3.1 there is a Λ-unimodular matrix $Q(\lambda)$, such that

$$[\bar{T}(\bar{\lambda}) \quad \bar{U}_1(\bar{\lambda})] = Q(\lambda)[T(\lambda) \quad U_1(\lambda)] \tag{3.50}$$

where

$$\bar{\lambda} = \frac{\lambda}{1-\lambda(\bar{a}-a)} . \tag{3.51}$$

Then (3.51) is a bijection between Λ and $\bar{\Lambda}$. The matrix $Q(\lambda_0)$ is by definition nonsingular for all $\lambda_o \in \Lambda$. Therefore it follows from (3.50) that $[T(\lambda_o) \quad U_1(\lambda_o)]$ has linearly independent rows for all $\lambda_o \in \Lambda$ if and only if $[\bar{T}(\bar{\lambda}_o) \quad \bar{U}_1(\bar{\lambda}_o)]$ has linearly independent rows for all $\bar{\lambda}_o \in \bar{\Lambda}$. The theorem now follows by lemma 3.2.

□

Theorem 3.3 and 3.4 show that the property of Λ-stabilizability is independent of which fractional representation and which a-value we use in representing the system.

Λ-detectability is a concept dual to the concept of Λ-stabilizability.

Definition 3.3 Let $L(\lambda)$ be the g.c.r.d of $T(\lambda)$ and $V_2(\lambda)$ in the representation (3.47). The system is Λ-detectable from z if det $L(\lambda)$ has no zeros in Λ.

It follows by duality that the property of Λ-detectability is independent of which fractional representation and which a-value that is used to represent the system.

3.5. The structure matrices

In this section the left and right Λ-structure matrices for a system S will be defined and it will be shown that they are essentially unique. The definitions are based on the definitions of left and right Λ-structure matrices for a Λ-generalized polynomial matrix in section 2.3. The structure matrices will turn out to be key concepts in the solutions of the design problems in the following chapters.

The left Λ-structure matrix shows how well the output y can be controlled in the servo sense, using a stable and causal precompensator. This situation will be treated in chapter 5. The right Λ-structure matrix shows how well the disturbance e can be reconstructed from measurements of the output z using a stable and causal reconstructor. This will be used in the solution of the feedback realization problem in chapter 7.

In section 3.3 it was shown that the system S can be represented by a polynomial matrix fractional representation of the type

$$A(\lambda)\eta = B(\lambda)u + C(\lambda)e, \qquad \lambda = \frac{1}{\mu - a}. \tag{3.52}$$

Here $A(\lambda)$, $B(\lambda)$ and $C(\lambda)$ are polynomial matrices. $A(\lambda)$ is square and nonsingular. $A(\lambda)$ and $[B(\lambda) \;\; C(\lambda)]$ are relatively left prime. It was pointed out in section 3.4 that a representation with $A(\lambda)$ upper block triangular always can be obtained by multiplication of (3.52) from the left by a unimodular matrix.

$$\begin{pmatrix} A_1(\lambda) & A_3(\lambda) \\ 0 & A_2(\lambda) \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} B_1(\lambda) \\ B_2(\lambda) \end{pmatrix} u + \begin{pmatrix} C_1(\lambda) \\ C_2(\lambda) \end{pmatrix} e \tag{3.53}$$

By theorem 2.6 (the dual form) $C_2(\lambda)$ can be factorized as $C_2(\lambda) = \tilde{C}(\lambda)\hat{C}(\lambda)$. If $C_2(\lambda)$ has dimension $p \times q$ and rank k, then $\hat{C}(\lambda)$ belongs to $R^{k \times q}[\lambda]$ and all the zeros of its invariant factors are in Λ. $\tilde{C}(\lambda)$ belongs to $R_{\Lambda}^{p \times k}[\lambda]$ and is left Λ-invertible. The matrix $\hat{C}(\lambda)$ is unique up to multiplication from the left by unimodular polynomial matrices. It is easy to show that $\tilde{C}(\lambda)$ is a polynomial matrix since $C_2(\lambda)$ and $\hat{C}(\lambda)$ are.

<u>Definition 3.3</u> The matrix $\hat{C}(\lambda)$ is a right Λ-structure matrix for the system S.

It will be shown that $\hat{C}(\lambda)$ is essentially unique. It was pointed out above that for a given representation (3.53) it follows from theorem 2.6 that $\hat{C}(\lambda)$ is unique up to multiplication from the left by unimodular polynomial matrices. It remains to be shown how structure matrices, defined from other representations of the form (3.53) with the same a and with different a, are related.

<u>Lemma 3.3</u> Let P, R and Q be square, nonsingular, rational matrices such that

$$P = Q\,R \tag{3.54}$$

If P and R are block traingular

$$P = \begin{pmatrix} P_1 & P_3 \\ 0 & P_2 \end{pmatrix} \qquad R = \begin{pmatrix} R_1 & R_3 \\ 0 & R_2 \end{pmatrix}, \tag{3.55}$$

where P_1 and R_1 are square matrices of the same dimension, then Q is also block triangular

$$Q = \begin{pmatrix} Q_1 & Q_3 \\ 0 & Q_2 \end{pmatrix}, \tag{3.56}$$

where Q_1 is square of the same dimension as P_1 and R_1.

The trivial proof is omitted.

Theorem 3.5 Let

$$\begin{pmatrix} T_1(\lambda) & T_3(\lambda) \\ 0 & T_2(\lambda) \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} U_1(\lambda) \\ U_2(\lambda) \end{pmatrix} u + \begin{pmatrix} Y_1(\lambda) \\ Y_2(\lambda) \end{pmatrix} e \tag{3.57}$$

be another representation of the type (3.53) for the same system S and with the same a. Let $\hat{Y}(\lambda)$ be the right Λ-structure matrix obtained from (3.57). Then

$$\hat{Y}(\lambda) = N(\lambda)\hat{C}(\lambda) \tag{3.58}$$

where $N(\lambda)$ is a unimodular polynomial matrix and $\hat{C}(\lambda)$ is the right Λ-structure matrix obtained from (3.53).

Proof The representation (3.57) can be obtained by multiplying (3.53) from the left by a unimodular polynomial matrix $Q(\lambda)$. It follows from lemma 3.3 that $Q(\lambda)$ can be partitioned as

$$Q(\lambda) = \begin{pmatrix} Q_1(\lambda) & Q_3(\lambda) \\ 0 & Q_2(\lambda) \end{pmatrix}. \tag{3.59}$$

It follows that $Q_1(\lambda)$ and $Q_2(\lambda)$ both are unimodular. We have

$$Y_2(\lambda) = Q_2(\lambda)C_2(\lambda). \tag{3.60}$$

Consequently

$$\tilde{Y}(\lambda)\hat{Y}(\lambda) = Q_2(\lambda)\tilde{C}(\lambda)\hat{C}(\lambda), \tag{3.61}$$

where $\tilde{Y}(\lambda)$ and $\tilde{C}(\lambda)$ are left Λ-invertible. It follows that also $Q_2(\lambda)\tilde{C}(\lambda)$ is left Λ-invertible. Therefore the left and right members of (3.61) are both factorizations of the type considered in theorem 2.6. The equality (3.58) then follows from theorem 2.6. □

Theorem 3.6 Let

$$\begin{pmatrix} T_1(\bar{\lambda}) & T_3(\bar{\lambda}) \\ 0 & T_2(\bar{\lambda}) \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} U_1(\bar{\lambda}) \\ U_2(\bar{\lambda}) \end{pmatrix} u + \begin{pmatrix} Y_1(\bar{\lambda}) \\ Y_2(\bar{\lambda}) \end{pmatrix} e \tag{3.62}$$

be a representation of the type (3.53) for the same system S, but with $\bar{a}$ instead of a and with

$$\bar{\lambda} = \frac{\lambda}{1-\lambda(\bar{a}-a)} \tag{3.63}$$

Let $\hat{Y}(\bar{\lambda})$ be the right $\bar{\Lambda}$-structure matrix obtained from (3.62), where $\bar{\Lambda}$ is the corresponding unstable region for (3.62). Then

$$\hat{Y}(\bar{\lambda}) = N(\lambda)\hat{C}(\lambda), \tag{3.64}$$

where $N(\lambda)$ is Λ-unimodular and $\hat{C}(\lambda)$ is the right Λ-structure matrix obtained from (3.53).

Proof It follows from theorem 3.1 that (3.62) can be obtained by multiplying (3.53) from the left by a Λ-unimodular matrix $Q(\lambda)$. By lemma 3.3 $Q(\lambda)$ can be partitioned as

$$Q(\lambda) = \begin{pmatrix} Q_1(\lambda) & Q_3(\lambda) \\ 0 & Q_2(\lambda) \end{pmatrix}. \tag{3.65}$$

It follows that $Q_1(\lambda)$ and $Q_2(\lambda)$ are Λ-unimodular and that

$$Y_2(\bar{\lambda}) = Q_2(\lambda)C_2(\lambda). \tag{3.66}$$

Consequently

$$\tilde{Y}(\bar{\lambda})\hat{Y}(\bar{\lambda}) = Q_2(\lambda)\tilde{C}(\lambda)\hat{C}(\lambda), \tag{3.67}$$

where $\tilde{Y}(\lambda)$ is left $\bar{\Lambda}$-invertible and $\tilde{C}(\lambda)$ is left Λ-invertible. Since $Q_2(\lambda)$ is Λ-unimodular then $Q_2(\lambda)\tilde{C}(\lambda)$ is left Λ-invertible. Therefore the right member of (3.67) is a factorization of the type considered in theorem 2.5 (the dual form of it).

Define

$$\tilde{Z}(\lambda) = \tilde{Y}(\bar{\lambda}) \tag{3.68 a}$$

$$\hat{Z}(\lambda) = \hat{Y}(\bar{\lambda}), \tag{3.68 b}$$

where λ and $\bar{\lambda}$ are connected through (3.63). The mapping (3.63) is a bijection between Λ and $\bar{\Lambda}$. $\tilde{Y}(\bar{\lambda})$ has by theorem 2.4 (the dual form of (i) $\Rightarrow$ (iii)) linearly independent columns for all $\bar{\lambda}$ in $\bar{\Lambda}$. Therefore $\tilde{Z}(\lambda)$ has linearly independent columns for all λ in Λ. By theorem 2.4 (the dual form of (iii) $\Rightarrow$ (i)) $\tilde{Z}(\lambda)$ is left Λ-invertible. Since (3.63) has no pole in Λ it follows that both $\tilde{Z}(\lambda)$ and $\hat{Z}(\lambda)$ are Λ-generalized polynomial matrices. Consequently $\tilde{Z}(\lambda)\hat{Z}(\lambda)$ is a factorization of the type considered in theorem 2.5. Introduce (3.68) into (3.67), then

$$\tilde{Z}(\lambda)\hat{Z}(\lambda) = \left(Q_2(\lambda)\tilde{C}(\lambda)\right)\hat{C}(\lambda), \tag{3.69}$$

where both members have been shown to be factorizations of the type considered in theorem 2.5. It then follows by theorem 2.5 that there is a Λ-unimodular $N(\lambda)$ such that

$$\hat{Z}(\lambda) = N(\lambda)\hat{C}(\lambda) \tag{3.70}$$

This together with (3.68 b) gives (3.64). □

Theorem 3.5 and 3.6 show that the structure matrix is essentially unique. For Λ-stable systems there is a close relationship between the transfer function and the right Λ-structure matrix.

Consider a Λ-stable system with transfer function $G(\lambda)$ partitioned as

$$G(\lambda) = \begin{bmatrix} G_{uy}(\lambda) & G_{ey}(\lambda) \\ G_{uz}(\lambda) & G_{ez}(\lambda) \end{bmatrix}. \tag{3.71}$$

Since $G(\lambda)$ is Λ-stable it has no poles in Λ and is therefore a Λ-generalized polynomial matrix. Therefore $G_{ez}(\lambda)$ can be factorized according to theorem 2.6 (the dual form) as

$$G_{ez}(\lambda) = \tilde{G}(\lambda)\hat{G}(\lambda), \tag{3.72}$$

where $\tilde{G}(\lambda)$ is left Λ-invertible and $\hat{G}(\lambda)$ is a polynomial matrix with all zeros of its invariant factors in Λ.

<u>Theorem 3.7</u>. Consider a <u>Λ-stable</u> system. The matrix $\hat{G}(\lambda)$, defined in (3.72) is a right Λ-structure matrix for the system.

<u>Proof</u> Consider the fractional representation (3.53) of the Λ-stable transfer function $G(\lambda)$. The right Λ-structure matrix $\hat{C}(\lambda)$ is defined from a factorization of $C_2(\lambda)$ according to theorem 2.6.

$$C_2(\lambda) = \tilde{C}(\lambda)\hat{C}(\lambda) \tag{3.73}$$

From (3.53) we obtain

$$G_{ez}(\lambda) = A_2^{-1}(\lambda)C_2(\lambda). \tag{3.74}$$

or by (3.73)

$$G_{ez}(\lambda) = A_2^{-1}(\lambda)\tilde{C}(\lambda)\hat{C}(\lambda). \tag{3.75}$$

Since $G(\lambda)$ is Λ-stable it follows that det $A_2(\lambda)$ has no zeros in Λ. Therefore $A_2^{-1}(\lambda)$ is Λ-unimodular. This implies that $A_2^{-1}(\lambda)\tilde{C}(\lambda)$ is left Λ-invertible since $\tilde{C}(\lambda)$ is. The right member of (3.75) is thus a factorization of the type considered in theorem 2.6. Since (3.72) is another such factorization it follows by theorem 2.6 that there is a unimodular polynomial matrix $N(\lambda)$ such that

$$\hat{G}(\lambda) = N(\lambda)\hat{C}(\lambda). \tag{3.76}$$

Therefore $\hat{G}(\lambda)$ is a right Λ-structure matrix for the system. □

The left Λ-structure matrix is defined from a fractional representation of the type (3.31). In analogy with (3.53) we have

$$\begin{pmatrix} D_1(\lambda) & D_3(\lambda) \\ 0 & D_2(\lambda) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} I \\ 0 \end{pmatrix} u + \begin{pmatrix} 0 \\ I \end{pmatrix} e \tag{3.77 a)}$$

$$y = \begin{pmatrix} E_1(\lambda) & E_2(\lambda) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \tag{3.77 b)}$$

$$z = \begin{pmatrix} F_1(\lambda) & F_2(\lambda) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \tag{3.77 c)}$$

Let the polynomial matrix $E_1(\lambda)$ be factorized according to theorem 2.6 as

$$E_1(\lambda) = \hat{E}(\lambda)\tilde{E}(\lambda) \tag{3.78}$$

where both $\hat{E}(\lambda)$ and $\tilde{E}(\lambda)$ are polynomial matrices. $\hat{E}(\lambda)$ has linearly independent columns and the zeros of its invariant factors all belong to Λ. $\tilde{E}(\lambda)$ is right Λ-invertible.

Definition 3.5 The matrix $\hat{E}(\lambda)$ is the left Λ-structure matrix for the system S.

In analogy with theorem 3.5 it can be shown that all left Λ-structure matrices obtained from representations of the type (3.77), corresponding to the same a, are related via multiplication from the right by unimodular polynomial matrices.

Let $\hat{E}(\lambda)$ be a left Λ-structure matrix obtained from (3.77). Consider another representation of the type (3.77) but with another a, say $\bar{a}$, and the corresponding unstable region $\bar{\Lambda}$. Let $\bar{\lambda}$ be given by

$$\bar{\lambda} = \frac{\lambda}{1-\lambda(\bar{a}-a)} , \qquad (3.79)$$

which is a bijection between Λ and $\bar{\Lambda}$. Let $\hat{K}(\bar{\lambda})$ be the left Λ-structure matrix obtained from this representation. As in theorem 3.6 it is shown that there is a Λ-unimodular $N(\lambda)$ such that

$$\hat{K}(\bar{\lambda}) = \hat{E}(\lambda)N(\lambda). \qquad (3.80)$$

If the system is Λ-stable then the left Λ-structure matrix for the Λ-generalized polynomial matrix $G_{uy}(\lambda)$ is a left Λ-structure matrix for the system S.

4. THE CLOSED LOOP SYSTEM

It is assumed that the system S is controlled by a linear, time invariant, finite dimensional, causal controller R, as is shown in the block diagram in figure 4.1.

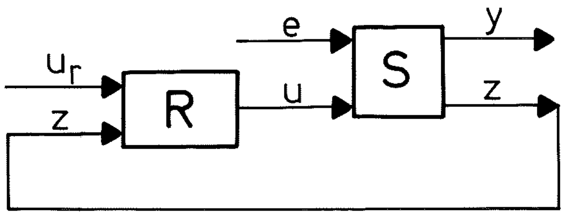


Figure 4.1 - The control configuration.

The block diagram can be written as is shown in figure 4.2.

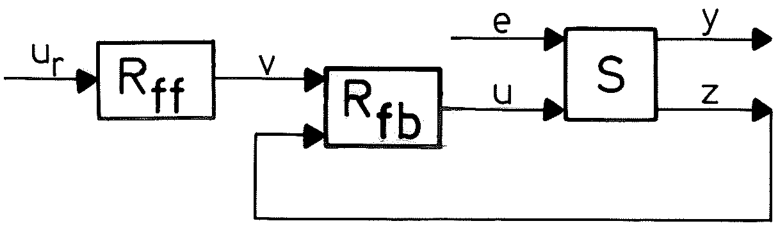


Figure 4.2 - An equivalent representation of the control configuration.

The two controller structures are equivalent in the sense that every controller, that can be represented as in figure 4.1, can also be represented as in figure 4.2 and vice versa. It turns out in the following chapters that the structure in figure 4.2 is more convenient to use.

The controllers R_{ff} and R_{fb} should be determined such that the closed loop system in figure 4.2 fulfils certain requirements. One requirement is that the closed loop system should be stable in some sense. In the previous chapter Λ-stability was defined for a system that is described by its transfer function. In section 4.1 Λ-stability will be defined for the more general case when a system is described by a set of difference or differential equations. This definition applies to the closed loop system in figure 4.2. In section 4.2 it is shown that there always exists a feedback controller that makes the closed loop system Λ-stable and in section 4.3 other requirements on the closed loop system are discussed.

4.1 Stability

In section 3.2 a system, described by its transfer function $G^*(\mu)$, was defined to be stable if $G^*(\mu)$ has no poles in a given subset $\Lambda_o^* = \Lambda^* \setminus \{\infty\}$ of the complex plane. If the transfer function is expressed in the operator λ as $G(\lambda)$, then the system is said to be Λ-stable if $G(\lambda)$ has no poles in the region Λ, given by (3.8). A Λ-stable system is always causal because of (3.11).

The closed loop system in figure 4.2 is not completely described by its transfer function. The reason is that some variables may become uncontrollable or unobservable when the system and the controllers are connected or equivalently that there will be cancellations when the transfer function for the closed loop system is formed. Therefore the definitions in section 3.2 do not apply.

The controller R_{ff} is assumed to be a minimal realization of its transfer function and the definitions in section 3.2 therefore apply to R_{ff}. Consider the rest of the system, which is described by the block diagram in figure 4.3.

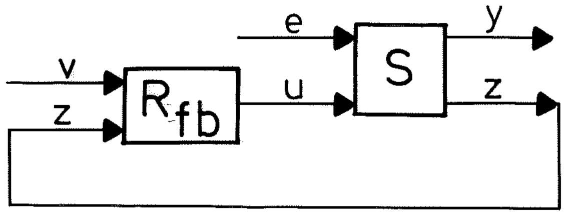


Figure 4.3 - The feedback loop.

Both S and R_{fb} are linear, finite dimensional, time invariant systems. Each of them is therefore described by a set of linear difference or differential equations. The total set of difference or differential equations then describes the closed loop system completely. It therefore also describes the stability properties. In this section stability for the closed loop system will be expressed as a condition on the fractional representations in the operator λ.

The difference or differential equations, describing the system S, are represented by the equations (3.53) or (3.77). Assume for simplicity of notation that they are represented by (3.53), i.e.

$$\begin{pmatrix} A_1(\lambda) & A_3(\lambda) \\ 0 & A_2(\lambda) \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} B_1(\lambda) \\ B_2(\lambda) \end{pmatrix} u + \begin{pmatrix} C_1(\lambda) \\ C_2(\lambda) \end{pmatrix} e. \qquad (4.1)$$

The design procedures, which will be described in the following chapters, lead to a controller R_{fb}, which can be represented as

$$T(\lambda)\xi = U(\lambda)z + Y(\lambda)v \qquad (4.2\ a)$$
$$u = V(\lambda)\xi. \qquad (4.2\ b)$$

Here $T(\lambda)$, $U(\lambda)$, $Y(\lambda)$ and $V(\lambda)$ are polynomial matrices. $T(\lambda)$ is square, nonsingular and $\det T(0) \neq 0$.

The relation between the polynomial matrices in (4.1) and (4.2) and the difference or differential equations, describing S and R_{fb}, will now be established in order to define stability for the closed loop system in figure 4.3.

Let $G^*(\mu)$ be the transfer function for S, expressed in the operator μ. Recall that μ is the difference operator q for discrete time systems and the differential operator p for continuous time systems. It is shown in Rosenbrock (1970) that a least order system of difference or differential equations with transfer function $G^*(\mu)$ is given by

$$A^*(\mu)\eta = B^*(\mu)u + C^*(\mu)e. \qquad (4.3)$$

Here $A^*(\mu)$ and $[B^*(\mu)\ C^*(\mu)]$ are relatively left prime polynomial matrices. $A^*(\mu)$ is square and nonsingular.

The relation between μ and λ is given by

$$\lambda = \frac{1}{\mu - a} \tag{4.4 a}$$

$$\Leftrightarrow \mu = \frac{1+\lambda a}{\lambda}, \tag{4.4 b}$$

where a is a real number outside the unstable region Λ^*. Let Λ be given by (3.8) and define

$$\Lambda_o = \Lambda \setminus \{0\}. \tag{4.5}$$

Write (4.1) as

$$A(\lambda)\eta = B(\lambda)u + C(\lambda)e. \tag{4.6}$$

<u>Lemma 4.1</u> There is a Λ_o-unimodular matrix $Q(\lambda)$, such that

$$A^*(\mu) = Q(\lambda)A(\lambda) \tag{4.7 a}$$

$$B^*(\mu) = Q(\lambda)B(\lambda) \tag{4.7 b}$$

$$C^*(\mu) = Q(\lambda)C(\lambda), \tag{4.7 c}$$

where μ and λ are related through (4.4).

<u>Proof</u> Since $A(\lambda)$ and $[B(\lambda)\ C(\lambda)]$ are relatively left prime the matrix $[A(\lambda_o)\ B(\lambda_o)\ C(\lambda_o)]$ has linearly independent rows for all $\lambda_o \in C$. It then follows from theorem 2.8 that they are relatively left Λ_o-prime.

Define the matrices

$$\bar{A}(\lambda) = A^*\left(\frac{1+\lambda a}{\lambda}\right) \tag{4.8 a}$$

$$\bar{B}(\lambda) = B^*\left(\frac{1+\lambda a}{\lambda}\right) \tag{4.8 b}$$

$$\bar{C}(\lambda) = C^*\left(\frac{1+\lambda a}{\lambda}\right). \tag{4.8 c}$$

The matrix $[\bar{A}(\lambda)\ \bar{B}(\lambda)\ \bar{C}(\lambda)]$ is a Λ_o-generalized polynomial matrix and has linearly independent rows for all $\lambda \in C \setminus \{0\}$ since $[A^*(\mu)\ B^*(\mu)\ C^*(\mu)]$ has linearly independent rows for all $\mu \in C$. Therefore $\bar{A}(\lambda)$ and $[\bar{B}(\lambda)\ \bar{C}(\lambda)]$ are relatively left Λ_o-prime. Since (4.3) and (4.6) have the same transfer function it follows from lemma 3.1 that there is a Λ_o-unimodular $Q(\lambda)$ such that

$$\bar{A}(\lambda) = Q(\lambda)A(\lambda) \tag{4.9 a}$$

$$[\bar{B}(\lambda)\ \bar{C}(\lambda)] = Q(\lambda)[B(\lambda)\ C(\lambda)]. \tag{4.9 b}$$

This is equivalent to (4.7).

□

The difference or differential equations for the controller (4.2) are obtained in the following way.

Insert (4.4 a) into (4.2) and determine diagonal matrices $N_1(\mu)$ and $M_1(\mu)$, where the diagonal entries are suitable powers of $\mu-a$, such that the matrices

$$T_1(\mu) = N_1(\mu)T\left(\frac{1}{\mu-a}\right) M_1(\mu) \tag{4.10 a}$$

$$U_1(\mu) = N_1(\mu)U\left(\frac{1}{\mu-a}\right) \tag{4.10 b}$$

$$Y_1(\mu) = N_1(\mu)Y\left(\frac{1}{\mu-a}\right) \tag{4.10 c}$$

$$V_1(\mu) = V\left(\frac{1}{\mu-a}\right) M_1(\mu) \tag{4.10 d}$$

become polynomial matrices. The matrices $N_2(\lambda) = N_1\left(\frac{1+\lambda a}{\lambda}\right)$ and $M_2(\lambda) = M_1\left(\frac{1+\lambda a}{\lambda}\right)$ are then Λ_o-unimodular Λ_o-generalized polynomial matrices.

Consider the system

$$T_1(\mu)\xi_1 = U_1(\mu)z + Y_1(\mu)v \tag{4.11 a}$$

$$u = V_1(\mu)\xi_1. \tag{4.11 b}$$

If this system has any decoupling zeros at the point a these can be removed as is shown in Rosenbrock (1970). The resulting system is

$$T^*(\mu)\xi^* = U^*(\mu)z + Y^*(\mu)v \tag{4.12 a}$$

$$u = V^*(\mu)\xi^*, \tag{4.12 b}$$

which is in polynomial form and has no decoupling zeros at the point a. Furthermore

$$T^*(\mu) = N_3(\mu)T_1(\mu)M_3(\mu) \tag{4.13 a}$$

$$U^*(\mu) = N_3(\mu)U_1(\mu) \tag{4.13 b}$$

$$Y^*(\mu) = N_3(\mu)Y_1(\mu) \tag{4.13 c}$$

$$V^*(\mu) = V_1(\mu)M_3(\mu),$$

where $N_3(\mu)$ and $M_3(\mu)$ are square, nonsingular, rational matrices with poles only at the point a and det $N_3(\mu)$ and det $M_3(\mu)$ are both negative powers of μ-a. The matrices $N_4(\lambda) = N_3\left(\frac{1+\lambda a}{\lambda}\right)$ and $M_4(\lambda) = M_3\left(\frac{1+\lambda a}{\lambda}\right)$ are then Λ_o-unimodular.

It follows from (4.10) and (4.13) that the matrices in (4.2) and in (4.12) are related as

$$T^*(\mu) = N(\lambda)T(\lambda)M(\lambda) \tag{4.14 a}$$

$$U^*(\mu) = N(\lambda)U(\lambda) \tag{4.14 b}$$

$$Y^*(\mu) = N(\lambda)Y(\lambda) \tag{4.14 c}$$

$$V^*(\mu) = V(\lambda)M(\lambda), \tag{4.14 d}$$

where $N(\lambda)$ and $M(\lambda)$ are Λ_o-unimodular Λ_o-generalized polynomial matrices. The equations (4.2) and (4.12) both give the same transfer function.

It was pointed out previously that the design procedures of the following chapters will result in a controller of the form (4.2). This result should be interpreted as the difference or differential equations (4.12).

Consequently the closed loop system is described by the equations

$$\begin{pmatrix} A_1^*(\mu) & A_2^*(\mu) & -B_1^*(\mu) & 0 \\ A_3^*(\mu) & A_4^*(\mu) & -B_2^*(\mu) & 0 \\ 0 & -U^*(\mu) & 0 & T^*(\mu) \\ 0 & 0 & I & -V^*(\mu) \end{pmatrix} \begin{pmatrix} y \\ z \\ u \\ \xi^* \end{pmatrix} = \begin{pmatrix} C_1^*(\mu) \\ C_1^*(\mu) \\ 0 \\ 0 \end{pmatrix} e + \begin{pmatrix} 0 \\ 0 \\ Y^*(\mu) \\ 0 \end{pmatrix} v, \tag{4.15}$$

where (4.3) has been partitioned.

The solutions to this system of equations are of the type $p_1(t)e^{\mu_1 t}$, $p_2(t)e^{\mu_2 t}$,... , in the continuous time case, and $p_1(t)\mu_1{}^t$, $p_2(t)\mu_2{}^t$, ... , in the discrete time case. In both cases $\{p_i(t)\}$ are polynomials in the time t and $\{\mu_i\}$ are the zeros of the determinant of the matrix in the left member of (4.15). The closed loop system is thus stable if this determinant has no zeros in the unstable region $\Lambda_o^* = \Lambda^* \setminus \{\infty\}$.

The representations (4.1) and (4.2) can be written

$$\begin{pmatrix} A_1(\lambda) & A_3(\lambda) & -B_1(\lambda) & 0 \\ 0 & A_2(\lambda) & -B_2(\lambda) & 0 \\ 0 & -U(\lambda) & 0 & T(\lambda) \\ 0 & 0 & I & -V(\lambda) \end{pmatrix} \begin{pmatrix} y \\ z \\ u \\ \xi \end{pmatrix} = \begin{pmatrix} C_1(\lambda) \\ C_2(\lambda) \\ 0 \\ 0 \end{pmatrix} e + \begin{pmatrix} 0 \\ 0 \\ Y(\lambda) \\ 0 \end{pmatrix} v. \tag{4.16}$$

From (4.7) and (4.14) it follows that

$$\begin{pmatrix} A_1^*(\mu) & A_2^*(\mu) & -B_1^*(\mu) & 0 \\ A_3^*(\mu) & A_4^*(\mu) & -B_2^*(\mu) & 0 \\ 0 & -U(\mu) & 0 & T^*(\mu) \\ 0 & 0 & I & -V^*(\mu) \end{pmatrix} = \begin{pmatrix} Q_1(\lambda) & Q_2(\lambda) & 0 & 0 \\ Q_3(\lambda) & Q_4(\lambda) & 0 & 0 \\ 0 & 0 & N(\lambda) & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \cdot$$

$$\cdot \begin{pmatrix} A_1(\lambda) & A_3(\lambda) & -B_1(\lambda) & 0 \\ 0 & A_2(\lambda) & -B_2(\lambda) & 0 \\ 0 & -U(\lambda) & 0 & T(\lambda) \\ 0 & 0 & I & -V(\lambda) \end{pmatrix} \cdot \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & M(\lambda) \end{pmatrix}, \quad (4.17\text{ a})$$

where $Q(\lambda)$ has been partitioned. Let the equality (4.17 a) be written

$$K^*(\mu) = \bar{N}(\lambda)K(\lambda)\bar{M}(\lambda), \quad (4.17\text{ b})$$

which defines the matrices $K^*(\mu)$, $\bar{N}(\lambda)$, $\bar{K}(\lambda)$ and $\bar{M}(\lambda)$. Notice that $\bar{N}(\lambda)$ and $\bar{M}(\lambda)$ are Λ_o-unimodular. Then following theorem now follows from (4.17) and (4.4).

Theorem 4.1 The polynomial det $K(\lambda)$ has no zeros in Λ_o if and only if det $K^*(\mu)$ has no zeros in $\Lambda_o^* = \Lambda^* \setminus \{\infty\}$.

Lemma 4.2 With $K(\lambda)$ given by (4.17) it follows that det $K(0) \neq 0$.

Proof It follows from theorem 3.2 that $B_2(0) = 0$ since there is no direct feedthrough from u to z. Furthermore it follows from (3.14) that det $A_1(0) \neq 0$ and det $A_2(0) \neq 0$. It was pointed out after (4.2) that det $T(0) \neq 0$. All this implies that det $K(0) \neq 0$.

□

It follows from lemma 4.2 that the transfer function, calculated from (4.16), from e or v to any of y, z or u has no poles at the origin, i.e. is causal. This means that the facts that S and R_{fb} are causal and that S has no direct feedthrough from u to z imply that all the internal variables y, z and u of the closed loop system depend causally on the inputs e and v.

Definition 4.1 Let $K(\lambda)$ be defined by (4.17). The closed loop system in figure 4.3 is Λ-stable if det $K(\lambda)$ has no zeros in Λ.

The closed loop system is thus Λ-stable if and only if the system of difference on differential equations (4.15) is stable with respect to $\Lambda^* \setminus \{\infty\}$.

The property of Λ-stability is invariant under the following transformation on (4.16).

Add one row of (4.16) multiplied from the left by a polynomial matrix, to another row. (4.18)

This corresponds to multiplication of (4.16) from the left by a unimodular polynomial matrix and does therefore not affect the property of Λ-stability. In fact, multiplication of (4.16) from the left by any Λ-unimodular matrix preserves the property of Λ-stability.

In this section the representation (4.1) has been used to describe the system S. If the representation (3.77) is used then

$$\begin{pmatrix} D_1(\lambda) & D_3(\lambda) & 0 & -I & 0 \\ 0 & D_2(\lambda) & 0 & 0 & 0 \\ -F_1(\lambda) & -F_2(\lambda) & I & 0 & 0 \\ 0 & 0 & -U(\lambda) & 0 & T(\lambda) \\ 0 & 0 & 0 & I & -V(\lambda) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ z \\ u \\ \xi \end{pmatrix} = \begin{pmatrix} 0 \\ I \\ 0 \\ 0 \\ 0 \end{pmatrix} e + \begin{pmatrix} 0 \\ 0 \\ 0 \\ Y(\lambda) \\ 0 \end{pmatrix} v \quad (4.19\text{ a})$$

$$y = E_1(\lambda)\xi_1 + E_2(\lambda)\xi_2 \quad (4.19\text{ b})$$

is obtained instead of (4.16). It follows from (3.34) that there are polynomial matrices $M(\lambda)$, $N(\lambda)$, $X(\lambda)$ and $Y(\lambda)$, where $M(\lambda)$ and $N(\lambda)$ are unimodular, such that

$$\left(\begin{array}{ccc|ccc} I & 0 & 0 & 0 & 0 & 0 \\ 0 & A_1 & A_3 & 0 & -B_1 & 0 \\ 0 & 0 & A_2 & 0 & -B_2 & 0 \\ \hline 0 & 0 & -I & I & 0 & 0 \\ 0 & 0 & 0 & -U & 0 & T \\ 0 & 0 & 0 & 0 & I & -V \end{array}\right) =$$

$$= \begin{pmatrix} M & 0 \\ X & I \end{pmatrix} \left(\begin{array}{ccc|ccc} I & 0 & 0 & 0 & 0 & 0 \\ 0 & D_1 & D_3 & 0 & -I & 0 \\ 0 & 0 & D_2 & 0 & 0 & 0 \\ \hline 0 & -F_1 & -F_2 & I & 0 & 0 \\ 0 & 0 & 0 & -U & 0 & T \\ 0 & 0 & 0 & 0 & I & -V \end{array}\right) \begin{pmatrix} N & Y \\ 0 & I \end{pmatrix} \tag{4.20}$$

<u>Lemma 4.3</u> Let $R(\lambda)$ be the matrix in the left member of (4.19 a). Then

$$\det R(\lambda) = k \det K(\lambda) \tag{4.21}$$

for some nonzero real number k.

<u>Proof</u> $\det R(\lambda)$ is equal to the determinant of the second matrix in the right member of (4.20) and $\det K(\lambda)$ is equal to the determinant of the left member of (4.20). Therefore the lemma follows from the fact that $M(\lambda)$ and $N(\lambda)$ are unimodular.

□

The lemma shows that the closed loop system in figure 4.3 is Λ-stable if and only if $\det R(\lambda)$ has no zeros in Λ. The following theorem has thus been shown.

Theorem 4.2 The property of Λ-stability does not depend on the choice of representation for the system S.

Analogously it can be shown that the property of Λ-stability is invariant under strict system equivalence transformations (see Rosenbrock (1970)) of the controller (4.2).

4.2 Stabilizing feedback controllers

It will now be shown that there always exists a controller R_{fb}, that makes the closed loop system in figure 4.3 Λ-stable.

Let the closed loop system be described by (4.16) and choose $U(\lambda) = I$. Add the third row, multiplied by $A_2(\lambda)$ and the fourth row, multiplied by $B_2(\lambda)$, to the second row. This gives

$$\begin{pmatrix} A_1 & A_3 & -B_1 & 0 \\ 0 & 0 & 0 & A_2T-B_2V \\ 0 & -I & 0 & T \\ 0 & 0 & I & -V \end{pmatrix} \begin{pmatrix} y \\ z \\ u \\ \xi \end{pmatrix} = \begin{pmatrix} C_1 \\ C_2 \\ 0 \\ 0 \end{pmatrix} e + \begin{pmatrix} 0 \\ A_2Y \\ Y \\ 0 \end{pmatrix} v \qquad (4.22)$$

Let $K(\lambda)$ be the matrix in the left member of (4.22). Then

$$\det K(\lambda) = \det A_1(\lambda) \det\Big(A_2(\lambda)T(\lambda) - B_2(\lambda)V(\lambda)\Big). \qquad (4.23)$$

It follows from the fact that the system S is Λ-detectable from z that $\det A_1(\lambda)$ has no zeros in Λ.

Let $L(\lambda)$ be the g.c.l.d. of $A_2(\lambda)$ and $B_2(\lambda)$. Then $\det L(\lambda)$ has no zeros in Λ since S is Λ-stabilizable from u. Therefore it is possible to choose $T(\lambda)$ and $V(\lambda)$, such that

$$L(\lambda) = A_2(\lambda)T(\lambda) - B_2(\lambda)V(\lambda). \qquad (4.24)$$

With this choice det $K(\lambda)$ has no zeros in Λ and the closed loop system is Λ-stable. It remains to be shown that det $T(0) \neq 0$ to ensure that the controller is causal.

It follows from theorem 3.2 that $B_2(0) = 0$, since there is no direct feedthrough from u to z. Therefore (4.24) gives

$$L(0) = A_2(0)T(0). \tag{4.25}$$

It follows from (3.14) that det $A_2(0) \neq 0$. This implies that det $L(0) \neq 0$ since $L(\lambda)$ is a left divisor to $A_2(\lambda)$. Consequently it follows from (4.25) that det $T(0) \neq 0$.

The result is summerized in the following theorem.

Theorem 4.3 There always exists a causal controller R_{fb}, such that the closed loop system in figure 4.3 is Λ-stable.

4.3 The control problem

Consider the closed loop system in figure 4.4.

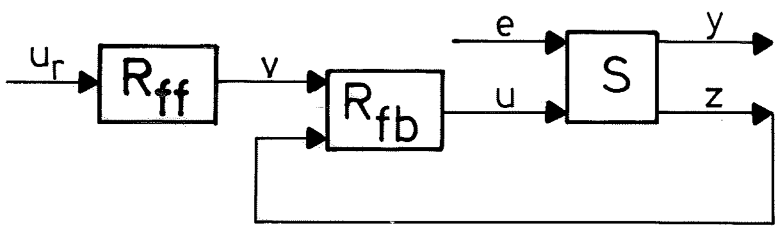


Figure 4.4 - The closed loop system.

It is assumed that the problem is to determine R_{ff} and R_{fb} such that both the servo and the regulator problems are solved in a satisfactory way.

The servo requirements are supposed to be stated in the following general way:

The transfer function from u_r to y should satisfy given specifications. (4.26)

An example of a specification is that the transfer function from u_r to y should be diagonal.

The regulator requirements are assumed to be stated in the following general way:

The transfer function from e to y should satisfy given specifications. (4.27)

An example of a specification is that the transfer function from e to y should have all its poles within a certain subset of the complex plane. Another example is that it is such that it does not transmit certain kinds of disturbances, for instance steps or ramps.

The class of available controllers is defined as follows:

The transfer function for R_{ff} must be Λ-stable and R_{fb} must be causal and such that the closed loop system is Λ-stable. (4.28)

It follows from theorem 4.3 that this class is not empty.

The servo and regulator problems are solved in the following sense. The class $\mathcal{H}$ of transfer functions from u_r to y and the class $\mathcal{F}$ of transfer functions from e to y, that can be obtained with controllers satisfying (4.28), are characterized. It is shown that an $H(\lambda) \in \mathcal{H}$ and $F(\lambda) \in \mathcal{F}$ can be obtained independently of each other. This leads to a separation of the servo problem from the regulator problem. Furthermore necessary and sufficient conditions for the existence of certain types of transfer functions in $\mathcal{H}$ and $\mathcal{F}$ are given. A method to calculate the controllers R_{ff} and R_{fb} for given $H(\lambda) \in \mathcal{H}$ and $F(\lambda) \in \mathcal{F}$ is also presented.

The control problem will be solved in the following steps. In chapter 5 the class $\mathcal{H}_o$ of transfer functions from u_r to y, that can be obtained with a fixed controller R_{fb}, is determined. In chapter 6 it is shown how the input v to R_{fb} should be chosen in order to maximize $\mathcal{H}_o$. This maximal class is called $\mathcal{H}$ and is in fact independent of the feedback controller R_{fb}. This fact implies that the servo problem can be separated from the regulator problem. In chapter 7 the class $\mathcal{D}$ of transfer functions, that can be obtained from e to u, is characterized. This class is called the class of feedback realizable transfer functions. In chapter 8 the characterization of the class $\mathcal{D}$ is used to determined the class $\mathcal{F}$ of transfer functions from e to y.

In this and the previous chapter the different matrix arguments μ, λ and $\bar{\lambda}$ were used. In the following chapters only the argument λ will be used. It can therefore be omitted and the matrix $G(\lambda)$ will for instance be written G.

5. THE SERVO PROBLEM

In this chapter it is assumed that a feedback controller R_{fb}, which makes the closed loop system Λ-stable, has been determined. The servo problem will be solved for this fixed feedback controller.

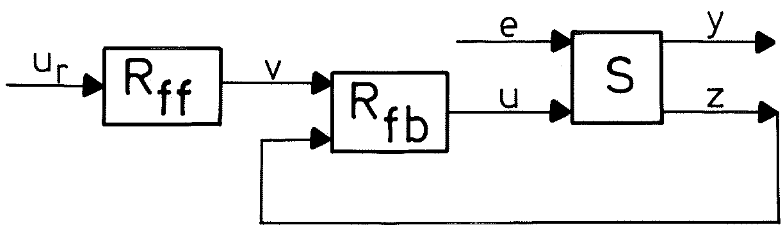


Figure 5.1 The control configuration.

Let G_o be the transfer function from v to y. Then G_o is a fixed, Λ-stable transfer function. Let K be the transfer function for R_{ff} and let H be the transfer from u_r to y, then

$$G_oK = H. \tag{5.1}$$

The transfer function K is by (4.28) required to be Λ-stable.

The servo problem will be solved in the following sense. The class $\mathcal{H}_o$ of transfer functions H, that can be obtained with Λ-stable K for a given G_o, will be characterized. Furthermore, a method to compute a Λ-stable K, if $H \in \mathcal{H}_o$ is given, will be presented.

Another method to compute a stable and causal K for a given H is shown in Bengtsson, Wonham (1976). They use a geometric state space approach rather than the polynomial matrix approach that will be used here. Another version of the same problem is solved in Wang, Davison (1973 a). They give a method to find a causal controller K of least possible order. The controller will not necessarily be stable.

The equation (5.1) occurs in many control problems. In Wang, Davison (1973 b) it is shown that the dynamic observer problem can be formulated as (5.1). The disturbance localization problem (see Wonham, Morse (1970)), with e measured, can also be formulated as (5.1). A generalization of this problem will be solved in chapter 8. The dynamic decoupling problem is solved in Bengtsson, Wonham (1976). A slightly different version of this problem is solved in this chapter, where also the connection to the concept of right invertibility (see Morse, Wonham (1971)) is given.

5.1 The class of transfer functions from the command input to the controlled output.

Let G_o be the transfer function from v to y in figure 5.1. Since the feedback controller R_{fb} is supposed to stabilize the system S it follows that G_o is Λ-stable. A transfer function is Λ-stable if and only if it is a Λ-generalized polynomial matrix. The problem is thus to characterize all transfer functions H in (5.1) that can be obtained with a Λ-generalized polynomial matrix K when G_o is fixed.

Since both G_o and K are Λ-generalized polynomial matrices then so is H. Therefore both G_o and H can be factorized, according to theorem 2.6, as

$$G_o = \hat{G}\,\tilde{G} \tag{5.2}$$

$$H = \hat{H}\,\tilde{H} \tag{5.3}$$

where $\hat{G}$ and $\hat{H}$ are left Λ-structure matrices and $\tilde{G}$ and $\tilde{H}$ are right Λ-invertible.

Theorem 5.1 Equation (5.1) has a stable and causal solution K if and only if $\hat{G}$ is a left divisor of $\hat{H}$.

Proof Suppose that $\hat{G}$ is a left divisor of $\hat{H}$. Then there is a polynomial matrix F, such that

$$\hat{H} = \hat{G} F \tag{5.4}$$

and (5.1) can be written

$$\hat{G} \tilde{G} K = \hat{G} F \tilde{H}. \tag{5.5}$$

Since $\tilde{G}$ is right Λ-invertible there is a Λ-generalized polynomial matrix G^+, such that

$$\tilde{G} G^+ = I. \tag{5.6}$$

The matrix

$$K = G^+ F \tilde{H} \tag{5.7}$$

is then a Λ-generalized polynomial matrix that is a solution to (5.5).

Conversely suppose that there is a Λ-generalized polynomial matrix K such that (5.1) is valid. Then

$$\hat{G} \tilde{G} K = H. \tag{5.8}$$

The Λ-generalized polynomial matrix $\tilde{G}$ K can be factorized as

$$\tilde{G} K = \hat{K} \tilde{K}, \tag{5.9}$$

where $\hat{K}$ is a left Λ-structure matrix and $\tilde{K}$ is right Λ-invertible. Insert (5.9) into (5.8)

$$\hat{G} \hat{K} \tilde{K} = H. \tag{5.10}$$

Both $\hat{G}$ and $\hat{K}$ have linearly independent columns for all λ outside Λ. Therefore $\hat{G} \hat{K}$ too has linearly independent columns for all λ outside Λ. Consequently $\hat{G} \hat{K}$ is a polynomial matrix with all zeros of its invariant factors in Λ. This, together with the fact that $\tilde{K}$ is right Λ-invertible, implies that $\hat{G} \hat{K}$

is a left Λ-structure matrix of H. By theorem 2.6 there is a unimodular polynomial matrix N, such that

$$\hat{H} = \hat{G}\,\hat{K}\,N. \tag{5.11}$$

Since $\hat{G}$ by definition has linearly independent columns, it follows from (5.11) that $\hat{G}$ is a left divisor of $\hat{H}$. □

Corollory 1. There is a stable and causal solution K to (5.1) if and only if there exists a Λ-generalized polynomial matrix M, such that

$$H = \hat{G}\,M. \tag{5.12}$$

Corollary 2. There is a stable and causal solution K to (5.1) with H = I if and only if $\hat{G}$ is unimodular.

Definition 5.1 Define $\mathcal{H}_o$ as

$$\mathcal{H}_o = \{\ H = \hat{G}\,M \ |\ M \text{ is a } \Lambda\text{-generalized polynomial matrix}\}.$$

It follows from corollary 1 that $\mathcal{H}_o$ is the class of all transfer functions H, such that there is a stable and causal solution K to (5.1) for a given G_o. The left Λ-structure matrix $\hat{G}$ contains the zeros of G_o in the unstable region of the complex plane. Corollary 1 shows that every H must contain those zeros. They can not be cancelled by a stable K. In the discrete time case $\hat{G}$ also contains the time delays between v and y. These time delays can not be annihilated by a causal K. The analogue interpretation is valid for the continuous time case. Finally the rank of $\hat{G}$ is equal to the rank of G_o. Therefore (5.12) shows that the rank of H can not be larger than the rank of G_o.

It has thus been shown that the class $\mathcal{H}_o$ of transfer functions from u_r to y, in figure 5.1, is characterized by the left Λ-structure matrix $\hat{G}$ of G_o. The problem is that G_o depends on the feedback R_{fb}. However, in chapter 6 it will be shown how v should be inserted into R_{fb} in order to maximize H_o. This maximal class is called H and is characterized by the left Λ-structure matrix of the system S. Consequently H does not depend on the feedback controller R_{fb}.

Example 5.1 Consider the system

$$y = G_o^* v, \tag{5.13}$$

where

$$G_o^* = \begin{pmatrix} \frac{1}{p+1} & \frac{2}{p+3} \\ \frac{1}{p+1} & \frac{1}{p+1} \end{pmatrix} . \tag{5.14}$$

This system is in Rosenbrock (1966) shown to be difficult to control because of the non minimum phase effects.

Let unstable region Λ^* be the closed right half plane and define λ as

$$\lambda = \frac{1}{p+2} . \tag{5.15}$$

The region Λ, given by (3.8), is then the closed disc in figure 5.2.

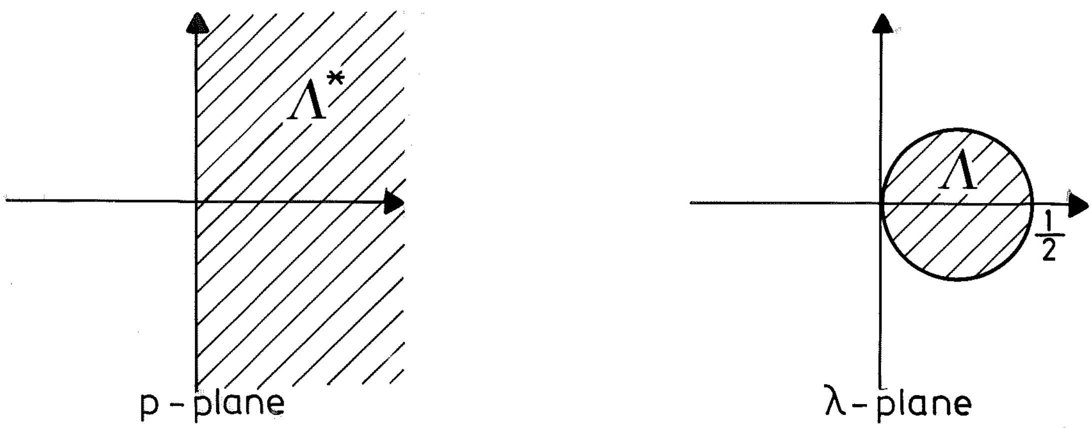


Figure 5.2 - The choice of unstable regions.

The transfer function G_o, expressed in λ, is obtained from (5.14) and (5.15)

$$G_o = \begin{pmatrix} \frac{\lambda}{1-\lambda} & \frac{2\lambda}{1+\lambda} \\ \frac{\lambda}{1-\lambda} & \frac{\lambda}{1-\lambda} \end{pmatrix} \tag{5.16}$$

The left Λ-structure matrix of G_o can now be computed by factorizing G_o according to theorem 2.6.

$$G_o = \begin{pmatrix} \lambda & 2\lambda(1-\lambda) \\ \lambda & \lambda(1+\lambda) \end{pmatrix} \begin{pmatrix} \frac{1}{1-\lambda} & 0 \\ 0 & \frac{1}{(1-\lambda)(1+\lambda)} \end{pmatrix} \triangleq \hat{G}_o \tilde{G}_o, \tag{5.17}$$

where $\tilde{G}_o$ is a right Λ-invertible Λ-generalized polynomial matrix. That $\hat{G}_o$ is the left Λ-structure matrix is seen in the following way. Add column one, multiplied by $2(\lambda-1)$, to column two. This is obtained by multiplication of $\hat{G}_o$ from the right by a unimodular matrix.

$$\hat{G} \triangleq \begin{pmatrix} \lambda & 2\lambda(1-\lambda) \\ \lambda & \lambda(1+\lambda) \end{pmatrix} \begin{pmatrix} 1 & 2(\lambda-1) \\ 0 & 1 \end{pmatrix} = \lambda \begin{pmatrix} 1 & 0 \\ 1 & 3\lambda-1 \end{pmatrix} . \tag{5.18}$$

The matrix $\hat{G}$ has linearly independent columns and the zeros of the invariant factors are $\lambda_1 = 0$ and $\lambda_2 = \frac{1}{3}$, which both belong to Λ. Therefore $\hat{G}$ (and $\hat{G}_o$) is a left Λ-structure matrix of G_o.

Assume that the transfer function H^* is given by

$$H^* = \begin{pmatrix} \frac{1}{p+10} & 0 \\ 0 & \frac{1}{p+10} \end{pmatrix} . \tag{5.19}$$

Introduction of (5.15) into (5.19) gives

$$H = \begin{pmatrix} \frac{\lambda}{1+8\lambda} & 0 \\ 0 & \frac{\lambda}{1+8\lambda} \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} \frac{1}{1+8\lambda} & 0 \\ 0 & \frac{1}{1+8\lambda} \end{pmatrix} . \qquad (5.20)$$

The left Λ-structure matrix $\hat{H}$ of H is then

$$\hat{H} = \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \qquad (5.21)$$

The matrix $\hat{G}$ is not a left divisor of $\hat{H}$. Therefore there is no $K(\lambda)$, satisfying (5.1), for G_o and H given by (5.17) and (5.20). □

<u>Example 5.2</u>. The following system is examined in Wang, Davison (1973 a).

$$G_o^* = \begin{pmatrix} \frac{1}{p+2} & \frac{p+3}{(p+1)(p+2)} & \frac{p(p+3)}{(p+1)(p+2)} \\ \frac{1}{p+1} & \frac{p}{p+1} & 0 \end{pmatrix} \qquad (5.22)$$

Introduce λ, given by (5.15), then

$$G_o = \begin{pmatrix} \lambda & \frac{\lambda(1+\lambda)}{1-\lambda} & \frac{(1-2\lambda)(1+\lambda)}{1-\lambda} \\ \frac{\lambda}{1-\lambda} & \frac{1-2\lambda}{1-\lambda} & 0 \end{pmatrix} . \qquad (5.23)$$

Theorem 2.4 ((iii) ⇒ (i)) can be used to show that G_o is right Λ-invertible. The entries (1,3) and (2,2) in G_o are nonzero for all λ in Λ, except $\lambda = \frac{1}{2}$. Therefore the rows of G are linearly independent for all $\lambda \in \Lambda$, except possibly $\lambda = \frac{1}{2}$. Insertion of $\lambda = \frac{1}{2}$ into (5.23) gives

$$G_o\left(\frac{1}{2}\right) = \begin{pmatrix} \frac{1}{2} & \frac{2}{3} & 0 \\ 1 & 0 & 0 \end{pmatrix} , \qquad (5.34)$$

which has linear independent rows. Therefore the rows of G_o

are linearly independent for all $\lambda \in \Lambda$ and G_o is right Λ-invertible. Consequently the left Λ-structure matrix $\hat{G}$ is given by

$$\hat{G} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{5.25}$$

Therefore any Λ-stable H belongs to H_o for this system. In particular, there is a Λ-stable K, satisfying (5.1) for

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{5.26}$$

and G_o given by (5.23). □

<u>Example 5.3</u> The following system is investigated in Mac Farlane, Postlethwaite (1977).

$$G_o^* = \frac{1}{(p+1)(p+2)} \begin{pmatrix} p-1 & p \\ -6 & p-2 \end{pmatrix} \tag{5.27}$$

Let λ be given by (5.15). Then

$$G_o = \frac{\lambda}{1-\lambda} \begin{pmatrix} 1-3\lambda & 1-2\lambda \\ -6 & 1-4\lambda \end{pmatrix} . \tag{5.28}$$

Factorize G_o as

$$G_o = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} \frac{1-3\lambda}{1-\lambda} & \frac{1-2\lambda}{1-\lambda} \\ \frac{-6}{1-\lambda} & \frac{1-4\lambda}{1-\lambda} \end{pmatrix} \triangleq \hat{G}\,\tilde{G} \tag{5.29}$$

The matrix $\tilde{G}$ is a Λ-generalized polynomial matrix. Furthermore

$$\det \tilde{G} = \frac{7-19\lambda + 12\lambda^2}{(1-\lambda)^2} \tag{5.30}$$

has no zeros in Λ. Therefore $\tilde{G}$ is Λ-unimodular. In particular $\tilde{G}$ is right Λ-invertible. Consequently

$$\hat{G} = \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{5.31}$$

is the left Λ-structure matrix of G_o.

This $\hat{G}$ is a left divisor of $\hat{H}$, given by (5.21). Therefore there is a Λ-stable K, satisfying (5.1), with G_o and H given by (5.28) and (5.20).

□

Example 5.4 Consider the system

$$G_o^* = \begin{pmatrix} \frac{1}{p+1} & \frac{1}{p+3} \\ \frac{1}{p+2} & \frac{p+1}{(p+2)(p+3)} \end{pmatrix}. \tag{5.32}$$

Let λ be given by (5.15). Then

$$G_o = \lambda \begin{pmatrix} \frac{1}{1-\lambda} & \frac{1}{1+\lambda} \\ 1 & \frac{1-\lambda}{1+\lambda} \end{pmatrix}. \tag{5.33}$$

Let G_o be factorized as

$$G_o = \lambda \begin{pmatrix} 1 & 1 \\ 1-\lambda & 1-\lambda \end{pmatrix} \begin{pmatrix} \frac{1}{1-\lambda} & 0 \\ 0 & \frac{1}{1+\lambda} \end{pmatrix} =$$

$$= \lambda \begin{pmatrix} 1 \\ 1-\lambda \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{1-\lambda} & 0 \\ 0 & \frac{1}{1+\lambda} \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda \\ \lambda(1-\lambda) \end{pmatrix} \begin{pmatrix} \frac{1}{1-\lambda} & \frac{1}{1+\lambda} \end{pmatrix} \triangleq \hat{G}\,\tilde{G}, \tag{5.34}$$

where $\tilde{G}$ is right Λ-invertible and $\hat{G}$ is the left Λ-structure matrix of G_o.

The matrix $\hat{G}$ is not a left divisor of $\hat{H}$, given by (5.21). Therefore there is no Λ-stable K, satisfying (5.1), with G_o and H, given by (5.33) and (5.20). □

5.2 Invertibility

The property of left invertibility for a dynamical system is defined in Morse, Wonham (1971). It is equivalent to the property left invertibility of the transfer matrix, regarded as a rational matrix. Right invertibility is defined as the dual concept. In Rosenbrock (1970) the concept of functional controllability is defined. It is equivalent to right invertibility. The following equivalent definition will be used here.

Definition 5.2 A system is right invertible if its transfer function has linearly independent rows.

In figure 5.1, consider the subsystem that has v as input and y as output. Let this system be called S_o and let its transfer be G_o. The system S_o is thus right invertible if and only if there is a rational matrix K such that

$$G_o K = I. \tag{5.35}$$

Even if S_o is right invertible there may not be any Λ-stable (i.e. stable and causal) solution K to (5.35). Necessary and sufficient conditions for the existence of a Λ-stable solution to (5.35) is given by corollary 2 of theorem 5.1.

This does not mean that the concept of right invertibility is uninteresting. It will be shown that the dynamic decoupling problem, defined below, can be solved if and only if the system S_o is right invertible.

The following theorem shows how the concept of right invertibility for a Λ-stable system can be expressed in terms of the left Λ-structure matrix of its transfer function.

Theorem 5.1 Let G_o be the transfer function for a Λ-stable system and let $\hat{G}$ be the left Λ-structure matrix of G_o. The system is right invertible if and only if $\hat{G}$ is square.

Proof Let G_o have n rows and rank r. Then the system is right invertible if and only if n = r. By theorem 2.6, $\hat{G}$ has n rows and r columns.

□

5.3 Dynamic decoupling

Consider the system in figure 5.1. The class $\mathcal{H}_o$ of possible transfer functions from the command input u_r to the controlled output y is characterized in definition 5.1. Let H be a transfer function in this class. It is often desirable to choose H as simple as possible. In many cases it is desirable that every component of u_r should influence only one of the components of y. This is the dynamic decoupling problem, which more precisely is formulated as follows. Let G_o be the transfer function from v to y in figure 5.1. Find necessary and sufficient conditions on G_o for the existence of a diagonal H in $\mathcal{H}$.

Different versions of the dynamic decoupling problem have been treated by many authors. Bengtsson, Wonham (1976) require for instance H to be block diagonal. The dyanmic decoupling problem is solved by the following theorem.

Theorem 5.3 There exists a diagonal $H \in \mathcal{H}_o$ with nonzero diagonal elements if and only if the left Λ-structure matrix $\hat{G}$ of G_o is square.

Proof By definition 5.1 every H in $\mathcal{H}_o$ can be written

$$H = \hat{G}\,M \tag{5.36}$$

where M is some Λ-generalized polynomial matrix. $\hat{G}$ has by definition linearly independent columns. Therefore $\hat{G}$ must be square if H is square and has full rank.

Conversely suppose that $\hat{G}$ is square and let S be the Smith-form of $\hat{G}$. Then there are unimodular polynomial matrices N and L such that

$$S = N \hat{G} L \tag{5.37}$$

Furthermore $S = \text{diag}(d_1, \ldots, d_m)$, where all d_i are nonzero since $\hat{G}$ has linearly independent columns. Define

$$\tilde{d}_i = \frac{d_m}{d_i} \qquad i = 1, \ldots, m. \tag{5.38}$$

All $\tilde{d}_i$ are polynomials because of the division property of the invariant factors. Let D be given by

$$D = \text{diag}(\tilde{d}_1, \ldots, \tilde{d}_m), \tag{5.39}$$

then

$$S D = d_m I. \tag{5.40}$$

Furthermore define

$$M = L D N \tag{5.41}$$

and introduce this into (5.36)

Then

$$H = \hat{G} L D N = N^{-1} S L^{-1} L D N =$$
$$= N^{-1} S D N = N^{-1} d_m I N = d_m I, \tag{5.42}$$

which means that H is diagonal.

□

The theorem shows that the dynamic decoupling problem can be solved if and only if the system, with transfer function G_o, is right invertible. One way to choose a diagonal H in $\mathcal{H}_o$ is given by (5.42). In some cases it is possible to choose some of the polynomials in the diagonal of lower degree than in (5.42).

Introduce λ, given by (3.7), into $d_m(\lambda)$. Then $d_m^*(\mu)$ is obtained. Since d_m is a polynomial, then d_m^* is a causal rational function with all its poles at the point a. These poles can be shifted to any points in the stable region by modification of M in (5.36).

It follows from theorem 5.3 that there exists no Λ-stable K and H, with H diagonal, such that

$$G_o K = H \tag{5.43}$$

for the system in example 5.4. For the systems in examples 5.2 and 5.3 it has been shown that H, given by (5.20) belongs to the class $\mathcal{H}_o$. This is not true for the system in example 5.1, but it follows from theorem 5.3 that there exist diagonal H in $\mathcal{H}_o$.

<u>Example 5.5</u> Consider the system in example 5.1. It was shown that

$$\hat{G} = \lambda \begin{pmatrix} 1 & 0 \\ 1 & 3\lambda-1 \end{pmatrix} . \tag{5.44}$$

A transfer function H belongs by definition 5.1 to $\mathcal{H}_o$ if and only if there is a Λ-stable M, such that

$$H = \hat{G}\, M. \tag{5.45}$$

The choice

$$M = \begin{pmatrix} 3\lambda-1 & 0 \\ -1 & 1 \end{pmatrix}$$

gives

$$H = \lambda \begin{pmatrix} 3\lambda-1 & 0 \\ 0 & 3\lambda-1 \end{pmatrix} . \tag{5.46}$$

Insert (5.15) into (5.46), then

$$H^* = \begin{pmatrix} \frac{-p+1}{(p+2)^2} & 0 \\ 0 & \frac{-p+1}{(p+2)^2} \end{pmatrix} . \tag{5.47}$$

In fact, every Q^* given by

$$Q^* = \begin{pmatrix} a\,\frac{p-1}{(p+b)\,(p+c)} & 0 \\ 0 & d\,\frac{p-1}{(p+e)\,(p+f)} \end{pmatrix} , \tag{5.48}$$

where b, c, e and f belong to the open left half of the complex plane, has (5.46) as its left Λ-structure matrix. Therefore there always exists a causal and stable K* such that

$$G^* K^* = Q^* , \tag{5.49}$$

where G^* is given by (5.14).

The transfer function Q^* has a zero in the right half plane for both loops. This corresponds to the factor 3λ-1 in (5.46). In $\hat{G}$ this factor is present in only one loop. Instead there is a coupling between the loops.

The only way to remove the coupling from (5.44) is to introduce the factor 3λ -1 into the other loop as well. Observe that it is not possible to remove the factor 3λ -1 from (5.44) because that would imply an M that is not Λ-stable.

Suppose that it is more important to control the first output than the second. Then maybe the coupling is not a disadvantage. It is more important not to introduce the factor $3\lambda-1$ into the first loop.

Choose M = I and insert (5.15) into (5.45). Then

$$H^* = \begin{pmatrix} \frac{1}{p+2} & 0 \\ \frac{1}{p+2} & \frac{-p+1}{(p+2)^2} \end{pmatrix} . \tag{5.50}$$

In fact every Q^*, given by

$$Q^* = \begin{pmatrix} \frac{a}{p+b} & 0 \\ \frac{a}{p+b} & c\,\frac{p-1}{(p+d)(p+e)} \end{pmatrix} , \tag{5.51}$$

where b, d and e belong to the open left half plane, has (5.44) as its left Λ-structure matrix.

This Q^* has a nicer response in the first loop than (5.48), but the second loop is not decoupled from the first.

If it is more important to control the second output, then choose M in (5.45) as

$$M = \begin{pmatrix} 1-3\lambda & 1 \\ 1 & 0 \end{pmatrix} . \tag{5.52}$$

This gives

$$H = \lambda \begin{pmatrix} 1-3\lambda & 1 \\ 0 & 1 \end{pmatrix} , \tag{5.53}$$

or

$$H^* = \begin{pmatrix} \frac{p-1}{(p+2)^2} & \frac{1}{p+2} \\ 0 & \frac{1}{p+2} \end{pmatrix} . \qquad (5.54)$$

Now the second loop gives a nice response, but not the first. □

5.4 A method to compute the feedforward controller.

The class $\mathcal{H}_o$ of possible transfer functions between u_r and y in figure 5.1 was characterized in section 5.1. Suppose that an $H \in \mathcal{H}_o$ is given and let G_o be the Λ-stable transfer function from v to y. In this section a method to calculate a Λ-stable K, such that

$$G_oK = H, \qquad (5.55)$$

will be given.

One way to find such a K is given by (5.7) in the proof of theorem 5.1. The construction includes calculations with Λ-generalized polynomial matrices. The method given in this section includes mainly operations on polynomial matrices and is therefore more suitable for computer calculations. Furthermore this method makes it possible to calculate the complete set of solutions to (5.55). It will also be shown that there is a fixed set of poles common to all solutions K. The method is inspired by Forney (1975).

Let G_o and $H \in \mathcal{H}_o$ have dimensions $m\times \ell$ and $m\times n$ and let the matrix $[G_o \;\; -H]$ have rank r. Furthermore let $[Y^T \; X^T]^T$, where $Y \in R^{\ell\times k}[\lambda]$, $X \in R^{n\times k}[\lambda]$ and $k = \ell + n - r$, be a polynomial basis, according to definition 2.13, for the nullspace of $[G_o \;\; -H]$. Then

$$[G_o \;\; -H]\begin{pmatrix} Y \\ X \end{pmatrix} = 0. \qquad (5.56)$$

Apply lemma 2.6, with $\Lambda = C$, to X instead of [A B]. It follows that there is a unimodular polynomial matrix M, such that

$$X\, M = [D \;\; 0], \tag{5.57}$$

where D has linearly independent columns. Define the polynomial matrices N and L through

$$\begin{pmatrix} Y \\ X \end{pmatrix} M = \begin{pmatrix} N & L \\ D & 0 \end{pmatrix}. \tag{5.58}$$

The right member of (5.58) is then also a polynomial basis for the nullspace of $[G_o \;\; -H]$.

<u>Theorem 5.4</u> The polynomial matrix D is square and nonsingular and any Λ-stable solution K to (5.55) can be written

$$K = Q\, P^{-1}, \tag{5.59}$$

where

$$P = D\, R \tag{5.60 a}$$

$$Q = N\, R + L\, Z. \tag{5.60 b}$$

Here R and Z are polynomial matrices. R is square and nonsingular and det R has no zeros in Λ. Conversely any K, given by (5.59)-(5.60), is a Λ-stable solution to (5.55).

<u>Proof</u> Let K be an arbitrary Λ-stable solution to (5.55). Such a solution exists by theorem 5.1. Then K can be written

$$K = Q\, P^{-1}, \tag{5.61}$$

where $P \in R^{n\times n}[\lambda]$ and $Q \in R^{\ell\times n}[\lambda]$. P is nonsingular and det P has no zeros in Λ. Introduction of (5.61) into (5.55) gives

$$[G_o \quad -H] \begin{pmatrix} Q\,P^{-1} \\ I \end{pmatrix} = 0. \tag{5.62}$$

This is equivalent to

$$[G_o \quad -H] \begin{pmatrix} Q \\ P \end{pmatrix} = 0. \tag{5.63}$$

Therefore $[Q^T \quad P^T]^T$ is a polynomial matrix in the nullspace of $[G_o \quad -H]$. It follows from theorem 2.10, with Λ = C, that there are polynomial matrices R and Z such that

$$\begin{pmatrix} Q \\ P \end{pmatrix} = \begin{pmatrix} N & L \\ D & O \end{pmatrix} \begin{pmatrix} R \\ Z \end{pmatrix} \tag{5.64}$$

or

$$Q = N\,R + L\,Z \tag{5.65 a}$$

$$P = D\,R. \tag{5.65 b}$$

Since P is square and nonsingular and D has linearly independent columns it follows that both D and R must be square and nonsingular. Therefore

$$\det P = \det D \det R. \tag{5.66}$$

Neither det D nor det R can have zeros in Λ since det P does not.

Conversely suppose that K is given by (5.59) and (5.60). Then

$$[G_o \quad -H] \begin{pmatrix} K \\ I \end{pmatrix} = [G_o \quad -H] \begin{pmatrix} Q \\ P \end{pmatrix} P^{-1} =$$

$$= [G_o \quad -H] \begin{pmatrix} N & L \\ D & 0 \end{pmatrix} \begin{pmatrix} R \\ Z \end{pmatrix} P^{-1} = 0. \tag{5.67}$$

Furthermore K is Λ-stable since it has been shown that det D has no zeros in Λ.

□

<u>Remark 1</u> The theorem gives a characterization of all Λ-stable solutions to (5.55).

<u>Remark 2</u> This method to determine all Λ-stable solutions to (5.55) is suitable for computer calculations. The only thing that is needed is an algorithm to calculate M and D in (5.57). Such an algorithm is described in Pernebo (1978), where it is shown that this algorithm also can be used to calculate the polynomial basis in (5.56).

<u>Corollary</u> Every Λ-stable solution K to (5.55) contains the zeros of det D among its poles.

Loosely speaking the zeros of det D are those poles of H that are not poles of G_o and those zeros of G_o that are not zeros of H. This is a generalization of a result in Bengtsson (1974), where it is shown that any right inverse of G_o has the zeros of G_o among its poles.

Choose R = I in (5.60) then the zeros of det D are the only poles of K. Let the transfer function K be expressed in the operator μ through

$$K^*(\mu) = K(\lambda), \tag{5.68}$$

where

$$\lambda = \frac{1}{\lambda - a} \tag{5.69}$$

and a is an arbitrary point on the real axis outside Λ^*. The relation between Λ and Λ^* is given by (3.8). The transfer function K^* has a set of poles outside Λ^*, corresponding to the poles of K, but K^* may also have some poles at the point a because of (5.69). Therefore the parameter a is a design parameter.

The poles at a may disappear for a suitable choice of R and Z. Other poles at the zeros of det R will then appear (c.f. example 5.6). A general method to choose R and Z so that the order of K^* is minimized is not yet known to the author.

Example 5.6 Consider the system in example 5.2.

$$G_o = \begin{pmatrix} \lambda & \frac{\lambda(1+\lambda)}{1-\lambda} & \frac{(1-2\lambda)(1+\lambda)}{1-\lambda} \\ \frac{\lambda}{1-\lambda} & \frac{1-2\lambda}{1-\lambda} & 0 \end{pmatrix} \tag{5.70}$$

Choose

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{5.71}$$

It was shown in example 5.2 that there exist a Λ-stable K, satisfying

$$G_oK = H. \tag{5.72}$$

Theorem 5.4 shows how all such K can be computed. A polynomial basis for the nullspace of $(G_o \quad -H)$ is given by

$$\begin{pmatrix} Y \\ X \end{pmatrix} = \left(\begin{array}{ccc} 1 & 0 & -\lambda+1 \\ 1 & 0 & 0 \\ 1 & \lambda-1 & 0 \\ \hline 2\lambda+1 & 2\lambda^2+\lambda-1 & -\lambda^2+\lambda \\ 1 & 0 & \lambda \end{array}\right) . \tag{5.73}$$

It can be computed with the method in Pernebo (1978). Elementary column operations can now be performed on (5.73) in order to bring it to the form (5.58). The polynomial matrices D, N and L in (5.58) are found to be

$$D = \begin{pmatrix} 3 & 9 \\ 0 & 3 \end{pmatrix} \tag{5.74}$$

$$N = \begin{pmatrix} 4\lambda^2 + 4\lambda - 3 & 8\lambda - 1 \\ 2\lambda^2 + 3\lambda & 4\lambda + 3 \\ -\lambda^2 + 3\lambda + 3 & -2\lambda + 9 \end{pmatrix} \qquad (5.75)$$

$$L = \begin{pmatrix} 4\lambda^3 - 3\lambda + 1 \\ 2\lambda^3 + \lambda^2 - \lambda \\ -\lambda^3 + 4\lambda^2 - \lambda \end{pmatrix} \qquad (5.76)$$

All Λ-stable K, which satisfy (5.72), can now be obtained from (5.59) and (5.60).

Since det D has no zeros, there are no fixed poles in K. If R = I and Z = 0 are chosen in (5.60) then K is given by

$$K = N\, D^{-1} \qquad (5.77)$$

which has no poles. Insert (5.15) into K to obtain K*. Then K* has all its poles at -2. These poles can be shifted by a proper choice of R and Z.

Choose for instance

$$R = \begin{pmatrix} -\lambda + 1 & \lambda - 3 \\ 0 & \lambda + 1 \end{pmatrix} \qquad (5.78\ a)$$

$$Z = (1 \quad -1). \qquad (5.78\ b)$$

Then it follows from (5.60) that

$$P = \begin{pmatrix} -3\lambda + 3 & 12\lambda \\ 0 & 3\lambda + 3 \end{pmatrix} \qquad (5.79)$$

$$Q = \begin{pmatrix} 4\lambda - 2 & -5\lambda + 7 \\ 2\lambda & -\lambda + 3 \\ -\lambda + 3 & 2\lambda \end{pmatrix} \qquad (5.80)$$

and K is obtained from (5.59). Insert (5.15) into this K to obtain K^*. Then K^* has one pole at -1 and one at -3, but no pole at -2.

□

6. COMBINATION OF SERVO AND REGULATOR DESIGN

In chapter 5 it was shown that the class H_o of possible transfer functions from u_r to y, in figure 6.1, is characterized by the left Λ-structure matrix $\hat{G}$ to the transfer function G_o from v to y.

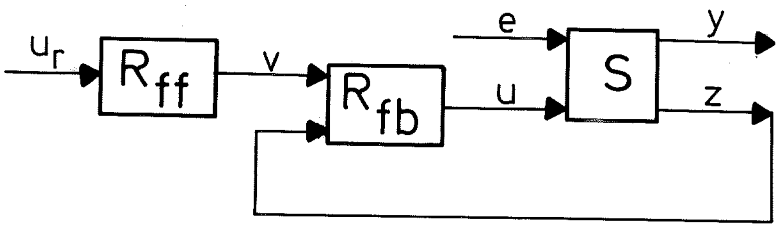


Figure 6.1 - The control configuration.

The transfer function G_o, and therefore also the class H_o, depends on the feedback controller R_{fb}. There is, however, a freedom to insert v into R_{fb} in many different ways. In this chapter it is shown that if this freedom is utilized then the class H of all possible transfer functions from u_r to y does not depend on the feedback controller R_{fb} as long as this is stabilizing. The class H is in fact characterized by the left Λ-structure matrix of the system S itself. This result implies that the servo and regulator requirements on the closed loop system can be specified independently.

6.1 The choice of input to the feedback controller

Let the equation

$$Tu = Uz + Yv \tag{6.1}$$

be a polynomial fractional representation of the feedback controller R_{fb} in figure 6.1. It is assumed that R_{fb} stabilizes the system S.

In chapter 5 it was shown that the class H_o of transfer functions from u_r to y is given by $\hat{G}$. Here $\hat{G}$ is the left Λ-structure matrix of the transfer function G_o, from v to y. In this section it will be shown how $\hat{G}$ depends on the choice of the polynomial matrix Y in (6.1). Observe that Y does not affect the stability of the closed loop system.

It was shown in section 3.5 that the system S can be described by the fractional representation (see (3.77))

$$\begin{pmatrix} D_1 & D_3 \\ 0 & D_2 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} I \\ 0 \end{pmatrix} u + \begin{pmatrix} 0 \\ I \end{pmatrix} e \tag{6.2 a}$$

$$y = \begin{pmatrix} E_1 & E_2 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \tag{6.2 b}$$

$$z = \begin{pmatrix} F_1 & F_2 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} . \tag{6.2 c}$$

Factorize E_1 as

$$E_1 = \hat{E}\,\tilde{E} \tag{6.3}$$

where $\hat{E}$ is the left Λ-structure matrix of the polynomial matrix E_1. By definition 3.5 $\hat{E}$ is the left Λ-structure matrix for the system S.

The transfer function G_o can be computed from (6.1) and (6.2). Inserting u, given by (6.2 a), and z, given by (6.2 c), into (6.1) gives

$$\left(T\,D_1 - U\,F_1\right)\xi_1 + \left(T\,D_3 - U\,F_2\right)\xi_2 = Y\,v. \tag{6.4}$$

The system of equations (6.1)-(6.2) is therefore equivalent to the following system of equations.

$$\begin{pmatrix} T\ D_1 - U\ F_1 & T\ D_3 - U\ F_2 \\ 0 & D_2 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} Y \\ 0 \end{pmatrix} v + \begin{pmatrix} 0 \\ I \end{pmatrix} e \tag{6.5 a}$$

$$\begin{pmatrix} y \\ z \\ u \end{pmatrix} = \begin{pmatrix} E_1 & E_2 \\ F_1 & F_2 \\ D_1 & D_3 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \tag{6.5 b}$$

Here the polynomial matrices T and U are supposed to be chosen so that (6.5) becomes Λ-stable. Furthermore T must be such that det $T(0) \neq 0$ for (6.1) to be causal. It follows from theorem 4.2 that such T and U always exist.

Let (6.1) be any feedback controller, such that the closed loop system (6.5) is Λ-stable. The transfer function G_o from v to y can be calculated from (6.5) as

$$G_o = E_1 (T\ D_1 - U\ F_1)^{-1} Y. \tag{6.6}$$

It is then a Λ-generalized polynomial matrix.

<u>Theorem 6.1</u> The left Λ-structure matrix $\hat{E}$ of the system S is a left divisor of the left Λ-structure matrix $\hat{G}$ of G_o for all causal and Λ-stabilizing controllers (6.1). If $Y = I$ then $\hat{E}$ and $\hat{G}$ are equal up to multiplication from the right by a unimodular matrix.

<u>Proof</u> From (6.3) and (6.6) it follows

$$G_o = \hat{E}\ \tilde{E} (T\ D_1 - U\ F_1)^{-1} Y. \tag{6.7}$$

Factorize $\tilde{E}(T\ D_1 - U\ F_1)^{-1} Y$ according to theorem 2.6 as

$$\tilde{E} (T\ D_1 - U\ F_1)^{-1} Y = \hat{X}\ \tilde{X}. \tag{6.8}$$

Insert (6.8) into (6.7). Then

$$G_o = \hat{E}\ \hat{X}\ \tilde{X}. \tag{6.9}$$

Both $\hat{E}$ and $\hat{X}$ have linearly independent columns for all λ outside Λ. Therefore $\hat{E}\ \hat{X}$ too has linearly independent columns for all λ outside Λ. Consequently $\hat{E}\ \hat{X}$ is a polynomial matrix with all zeros of its invariant factors in Λ. Since $\tilde{X}$ is right Λ-invertible it follows that $\hat{E}\ \hat{X}$ is a left Λ-structure matrix of G_o. By theorem 2.6 there is a unimodular polynomial matrix N, such that

$$\hat{G} = \hat{E}\ \hat{X}\ N. \tag{6.10}$$

Since $\hat{E}$ has linearly independent columns it is a left divisor of G.

With Y = I (6.7) becomes

$$G_o = \hat{E}\ \tilde{E}\ (T\ D_1 - U\ F_1)^{-1}. \tag{6.11}$$

Since $\det(T\ D_1 - U\ F_1)$ has no zeros in Λ it follows that $(T\ D_1 - U\ F_1)^{-1}$ is Λ-unimodular. Therefore $\tilde{E}(T\ D_1 - U\ F_1)^{-1}$ is right Λ-invertible. By theorem 2.6 there is a unimodular polynomial matrix Q, such that

$$\hat{G} = \hat{E}\ Q. \tag{6.12}$$

□

<u>Definition 6.1</u> Define $\mathcal{H}$ as

$$\mathcal{H} = \{H = \hat{E}\ M \mid M \text{ is a } \Lambda\text{-generalized polynomial matrix}\},$$

where $\hat{E}$ is the left Λ-structure matrix of the system S.

The following theorem is a direct consequence of theorem 4.3, 5.1 and 6.1.

<u>Theorem 6.2</u> There exist causal controllers R_{ff} and R_{fb} in figure 6.1, such that the closed loop system is Λ-stable and has the transfer function H from u_r to y if and only if $H \in \mathcal{H}$.

The class $\mathcal{H}$ is thus equal to the class of transfer functions that can be achieved from u_r to y. The left Λ-structure matrix $\hat{E}$ of the system S containes the zeros, which belong to the unstable region of the complex plane and are associated with the transmission from u to y. Theorem 6.2 shows that these zeros will always be present in the transfer function H from u_r to y. They can not be altered by a feedback controller R_{fb} or cancelled by a stable feedforward controller R_{ff}. In the discrete time case $\hat{E}$ contains the time delays between u and y. These time delays must be present in the transmission from u_r to y if the controllers are causal. The analogue result is valid for the continuous time case. Furthermore the rank of $\hat{E}$ is equal to the rank of the transfer function from u to y. The rank of H can not be larger than this rank.

6.2 Separation

The control problem was formulated in chapter 4. The controllers R_{ff} and R_{fb} have to be chosen from the class of admissible controllers (4.28). Furthermore, it is assumed that the controllers should fulfil servo requirements of the type (4.26) and regulator requirements of the type (4.27).

The feedforward controller R_{ff} does not influence the transfer function from e to y in figure 6.1. Neither does the matrix Y in (6.1) influence this transfer function and the Λ-stability of the closed loop system must be achieved by a proper choice of the matrices T and U in the feedback controller (6.1).

On the other hand it follows from theorem 5.1 and 6.1 that for any choice of T and U there is a matrix Y in (6.1) and a controller R_{ff}, such that any $H \in \mathcal{H}$ can be obtained as a transfer function from u_r to y in figure 6.1. Therefore the choice of T and U does not influence the possibilities to achieve the servo requirements (4.26).

It has thus been shown that the servo requirements (4.26) and the regulator requirements (4.27) can be stated independently and that the control problem can be separated in the following way.

First determine T and U in the feedback controller (6.1) so that the regulator requirements (4.27) are fulfilled and the closed loop system is Λ-stable. Then choose Y = I in (6.1). Finally determine a Λ-stable R_{ff}, such that the servo requirements (4.26) are fulfilled.

The problem to determine R_{ff} was solved in chapter 5. The problem to determine T and U in R_{fb} will be solved in chapter 7 and 8.

7. FEEDBACK REALIZATIONS

In this and the next chapter the problem to determine the feedback controller R_{fb} in figure 7.1 will be considered.

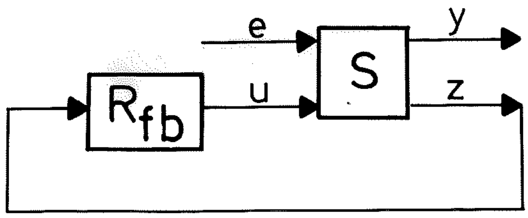


Figure 7.1 - The system with a feedback controller.

The feedback controller R_{fb} must belong to the class (4.28) of admissible controllers and should be determined so that it fulfils regulator requirements of the type (4.27). When R_{fb} has been determined it should be included in the total control system as is shown in section 6.2.

The regulator requirements in (4.27) are requirements on the transfer function F from e to y in the closed loop system in figure 7.1. Therefore, it is of interest to characterize the class $\mathcal{F}$ of all transfer functions from e to y that can be obtained with feedback controllers in the class (4.28) of admissible controllers. In this chapter it will be shown how feedback realizable transfer functions can be used to describe the class $\mathcal{F}$. It will also be shown that the right Λ-structure matrix of the system S can be used to characterize the class of feedback realizable transfer functions.

7.1 Feedback realizable transfer functions

Let the system S be given and consider the control configuration in figure 7.1.

Definition 7.1 The class $\mathcal{D}$ of feedback realizable transfer functions for the system S is equal to the class of all transfer functions from e to u that can be obtained by admissible feedback controllers in figure 7.1.

Recall that an admissible feedback controller, by (4.28), is a causal controller R_{fb} that makes the closed loop system Λ-stable. The following class of transfer functions is of interest since the regulator requirements are requirements on the transfer function from e to y in the configuration in figure 7.1.

Definition 7.2 Let $\mathcal{F}$ be the class of all transfer functions from e to y that can be obtained by admissible feedback controllers in figure 7.1.

Partition the transfer function for the system S as

$$\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} G_{uy} & G_{ey} \\ G_{uz} & G_{ez} \end{pmatrix} \begin{pmatrix} u \\ e \end{pmatrix} . \tag{7.1}$$

The following theorem is a direct consequence of definitions 7.1 and 7.2.

Theorem 7.1 Any transfer function $F \in \mathcal{F}$ can be written

$$F = G_{uy}D + G_{ey} \tag{7.2}$$

for some $D \in \mathcal{D}$. Conversely, for any $D \in \mathcal{D}$ the relation (7.2) gives an F in the class $\mathcal{F}$.

The theorem says that the class $\mathcal{F}$ can be generated by (7.2) with D in the class of feedback realizable transfer functions.

It was shown in section 3.5 that the system S can be described by the fractional representation

$$\begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} u + \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} e. \tag{7.3}$$

It follows from the fact that S is assumed to be Λ-detectable that det A_1 has no zeros in Λ. Let $\hat{C}$ be the right Λ-structure matrix for the system S. Then C_2 can be factorized as

$$C_2 = \tilde{C}\hat{C}, \tag{7.4}$$

where $\tilde{C}$ is a left Λ-invertible polynomial matrix.

Let D be a transfer function and define

$$Z = G_{uz} D + G_{ez}, \tag{7.5}$$

where G_{uz} and G_{ez} are given by (7.1). If D is the transfer function from e to u in figure 7.1, then Z is the transfer function from e to z.

The following theorem gives a characterization of the class $\mathcal{D}$ of feedback realizable transfer functions.

Theorem 7.2 The transfer function D belongs to the class $\mathcal{D}$ of feedback realizable transfer functions if and only if the following two conditions are satisfied.

- D is a Λ-generalized polynomial matrix and $\hat{C}$ is a right divisor of $\hat{D}$, where $\hat{D}$ is the right Λ-structure matrix of D. (7.6)

- $Z = G_{uz}D + G_{ez}$ is a Λ-generalized polynomial matrix and $\hat{C}$ is a right divisor of $\hat{Z}$, where $\hat{Z}$ is the right Λ-structure matrix of Z. (7.7)

Proof

Necessity of (7.6) and (7.7)

Suppose that $D \in \mathcal{D}$. By definition 7.1 there is an admissible controller R_{fb}, such that the transfer function from e to u in figure 7.1 is D. Let this R_{fb} be described by the fractional representation

$$P\xi = z \tag{7.8 a}$$
$$u = R\xi. \tag{7.8 b}$$

Insert (7.8) into (7.3). Then

$$\begin{pmatrix} A_1 & A_3P - B_1R \\ 0 & A_2P - B_2R \end{pmatrix} \begin{pmatrix} y \\ \xi \end{pmatrix} = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} e. \tag{7.9}$$

The insertion is an equivalence transformation of the type (4.18).

The closed loop system is thus described by (7.8) and (7.9). It is Λ-stable because det A_1 has no zeros in Λ and P and R are such that $\det(A_2P - B_2R)$ has no zeros in Λ. The transfer function D is given by

$$D = R(A_2P - B_2R)^{-1}\tilde{C}\hat{C} \tag{7.10}$$

The transfer function Z, defined by (7.5), is obtained from the second row of (7.3) as

$$Z = A_2^{-1}B_2D + A_2^{-1}C_2. \tag{7.11}$$

With D as in (7.10) this gives

$$Z = A_2^{-1}[B_2R(A_2P - B_2R)^{-1}\tilde{C}\hat{C} + \tilde{C}\hat{C}] = P(A_2P - B_2R)^{-1}\tilde{C}\hat{C}. \tag{7.12}$$

It follows from (7.10) and (7.12) that both D and Z are Λ-generalized polynomial matrices. That $\hat{C}$ is a right divisor of both $\hat{D}$ and $\hat{Z}$ follows from (7.10) and (7.12) as in the first half of the proof of theorem 6.1. Therefore, both (7.6) and (7.7) are satisfied.

Sufficiency of (7.6) and (7.7)

Conversely, suppose that D is given, such that (7.6) and (7.7) are satisfied. A feedback controller R_{fb} will be constructed, such that the closed loop system in figure 7.1 is Λ-stable and has a transfer function from e to u that is equal to D.

Step 1: Construction of R_{fb}

It follows from (7.6) that D can be written as

$$D = VT^{-1}\hat{C}, \qquad (7.13)$$

where T and V are relatively right prime polynomial matrices and det T has no zeros in Λ. Define the polynomial matrix Q through

$$Q = B_2V + \tilde{C}T, \qquad (7.14)$$

where $\tilde{C}$ is given by (7.4). The transfer function Z is given by (7.11). With D as in (7.13) this gives

$$Z = A_2^{-1}B_2VT^{-1}\hat{C} + A_2^{-1}C_2 = A_2^{-1}(B_2V + \tilde{C}T)T^{-1}\hat{C} = A_2^{-1}QT^{-1}\hat{C}, \qquad (7.15)$$

where (7.4) and (7.14) have been used.

Let L be the g.c.l.d. of A_2 and Q. Then there are relatively left prime polynomial matrices A_o and Q_o, such that

$$A_2 = LA_o \qquad (7.16\text{ a})$$

$$Q = LQ_o. \qquad (7.16\text{ b})$$

The equality (7.15) then gives

$$Z = A_o^{-1} Q_o T^{-1} \hat{C}. \tag{7.17}$$

It follows from (7.7) that $A_o^{-1} Q_o T^{-1}$ is a Λ-generalized polynomial matrix. Since A_o and Q_o are relatively left prime <u>det A_o has no zeros in Λ</u>.

Define the matrices $\overline{T}$, $\overline{V}$ and $\overline{Q}$ through

$$\overline{T} = \begin{pmatrix} T & 0 \\ 0 & T_1 \end{pmatrix} \tag{7.18 a}$$

$$\overline{V} = (V \quad V_1) \tag{7.18 b}$$

$$\overline{Q} = (Q_o \quad Q_1). \tag{7.18 c}$$

The polynomial matrices T_1, V_1 and Q_1 will be defined below.

Let $(S \quad 0)^T$ be the Smithform of $\tilde{C}$. Then det S is nonzero and has no zeros in Λ, since $\tilde{C}$ is left Λ-invertible. There are unimodular polynomial matrices N and M, such that

$$\tilde{C} = N \begin{pmatrix} S \\ 0 \end{pmatrix} M. \tag{7.19}$$

Define $\overline{C}$ and C_o as

$$\overline{C} = (\tilde{C} \quad C_o) = N \begin{pmatrix} S & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} M & 0 \\ 0 & I \end{pmatrix}. \tag{7.20}$$

Then <u>det $\overline{C}$</u> is nonzero and <u>has no zeros in Λ</u>.

Let L_1 be the g.c.l.d. of L and B_2. Then L_1 is also a left divisor to A_2 because of (7.16a). Since L_1 is a left divisor of both A_2 and B_2 it follows that det L_1 has no zeros in Λ, because the system is Λ-stabilizable from u.

It is shown e.g. in Rosenbrock (1970) that there are relatively right prime polynomial matrices T_1 and Y such that

$$L_1^{-1}C_o = YT_1^{-1}. \tag{7.21}$$

Here det T_1 divides det L_1. Therefore det T_1 has no zeros in Λ. It follows that det $\overline{T}$, in (7.18a), has no zeros in Λ.

Determine Q_1 and V_1, such that

$$L_1Y = LQ_1 - B_2V_1. \tag{7.22}$$

This is possible because L_1 is the g.c.l.d. of L and B_2. (See theorem 2.9 applied to polynomial matrices). The matrices T_1, V_1 and Q_1 in (7.18) are now determined.

Introduce (7.21) into (7.22), then

$$C_oT_1 = LQ_1 - B_2V_1. \tag{7.23}$$

From (7.14) and (7.16b) it follows that

$$\tilde{C}T = LQ_o - B_2V. \tag{7.24}$$

Using (7.18) and (7.20) the equalities (7.23) and (7.24) can be written

$$\overline{C}\,\overline{T} = L\overline{Q} - B_2\overline{V}. \tag{7.25}$$

It follows by theorem 3.2 that $B_2(0) = 0$ since the system S has no direct feedthrough from u to z.

Therefore

$$\bar{C}(0)\bar{T}(0) = L(0)\bar{Q}(0). \tag{7.26}$$

It has been shown that neither det $\bar{T}$ nor det $\bar{C}$ has any zero in Λ. Therefore det $\bar{T}(0) \neq 0$ and det $\bar{C}(0) \neq 0$ and it follows from (7.26) that

$$\det \bar{Q}(0) \neq 0. \tag{7.27}$$

Choose the feedback controller R_{fb} in figure 7.1 as the system

$$\bar{Q}\xi = A_o z \tag{7.28 a}$$

$$u = \bar{V}\xi. \tag{7.28 b}$$

This is a well defined system since det $\bar{Q}$ is nonzero by (7.27). The system is causal since the transfer function $\bar{V}\,\bar{Q}^{-1}A_o$ from z to u has no poles at the origin because of (7.27).

Step 2: Check of conditions

It remains to be shown that the closed loop system in figure 7.1 is Λ-stable and that the transfer function from e to u is D.

The closed loop system is described by the equations (7.3) and (7.28). Introduce (7.16a) into (7.3), which becomes

$$A_1 y + A_3 z = B_1 u + C_1 e \tag{7.29 a}$$

$$LA_o z = B_2 u + C_2 e. \tag{7.29 b}$$

The following operations on the closed loop system are all equivalence transformations of the type (4.18). Add (7.28a), multiplied from the left by L, and (7.28b), multiplied from the left by B_2, to (7.29b). This gives

$$L\overline{Q}\xi = B_2\overline{V}\xi + C_2 e. \tag{7.30}$$

Use (7.25) to rewrite (7.30) as

$$\overline{C}\,\overline{T}\xi = C_2 e. \tag{7.31}$$

Add (7.28b), multiplied from the left by B_1, to (7.29a). This gives

$$A_1 y + A_3 z - B_1\overline{V}\xi = C_1 e. \tag{7.32}$$

The closed loop system in figure 7.1 is thus equivalently described by (7.28), (7.31) and (7.32). These equations can be written as

$$\begin{pmatrix} A_1 & A_3 & -B_1\overline{V} \\ 0 & A_o & -\overline{Q} \\ 0 & 0 & \overline{C}\,\overline{T} \end{pmatrix} \begin{pmatrix} y \\ z \\ \xi \end{pmatrix} = \begin{pmatrix} C_1 \\ 0 \\ C_2 \end{pmatrix} e \tag{7.33 a}$$

$$u = \overline{V}\xi \tag{7.33 b}$$

This system is Λ-stable because det A_1, det A_o, det $\overline{C}$ and det $\overline{T}$ have all been shown to have no zero in Λ.

The transfer function X from e to u is obtained from (7.33) as

$$X = \overline{V}\,\overline{T}^{-1}\overline{C}^{-1}C_2. \tag{7.34}$$

From the left equality in (7.20) it follows that

$$I = \bar{C}^{-1}(\tilde{C} \quad C_o). \tag{7.35}$$

The first block column of (7.35) gives

$$\begin{pmatrix} I \\ 0 \end{pmatrix} = \bar{C}^{-1}\tilde{C}. \tag{7.36}$$

Use (7.4), (7.13) and (7.18) to rewrite (7.34).

$$X = \bar{V}\,\bar{T}^{-1}\bar{C}^{-1}C_2 = (V \quad V_1)\begin{pmatrix} T^{-1} & 0 \\ 0 & T_1^{-1} \end{pmatrix}\bar{C}^{-1}\tilde{C}\hat{C} =$$

$$= (VT^{-1} \quad V_1T_1^{-1})\begin{pmatrix} I \\ 0 \end{pmatrix}\hat{C} = VT^{-1}\hat{C} = D \tag{7.37}$$

□

Remark The transfer function $Z = G_{uz}D + G_{ez}$ is in (7.7) required to be a Λ-generalized polynomial matrix. Observe that this does not prevent G_{uz} and G_{ez} from having poles in Λ.

The following heuristic argument gives an explanation of the conditions (7.6) and (7.7) of theorem 7.2.

The right Λ-structure matrix $\hat{C}$ of the system S contains the zeros, which belong to the unstable region of the complex plane, associated with the transmission from e to z. These zeros cannot be cancelled by the controller R_{fb}, since such a cancellation would introduce unstable and uncontrollable modes. In the discrete time case $\hat{C}$ also contains the time delays from e to z. These time delays cannot be annihilated by a causal controller R_{fb}. The analogue result holds in the continuous time case. Finally the rank of $\hat{C}$ is equal to the rank of the transfer function from e to z. Intuitively speaking, this rank cannot be increased by any controller R_{fb}. In other words, the rank of the transfer function from e to u cannot be larger than rank $\hat{C}$.

The implication of all this is that an admissible controller R_{fb} cannot reconstruct the disturbance e itself, but only $\hat{C}e$. The input u can then be formed by letting $\hat{C}e$ pass through some Λ-stable transfer function $\bar{D}$.

This argument leads to the conclusion that the transfer function from e to u should be of the form $\bar{D}\hat{C}$. This is exactly the condition (7.6). The condition (7.7) says that this transfer function from e to u must not excite any unstable modes in the system S.

The theorem can be simplified if the system S is Λ-stable.

Corollary If the system S is Λ-stable, then D belongs to the class $\mathcal{D}$ of feedback realizable transfer functions if and only if D can be written

$$D = D_o\hat{C}, \tag{7.38}$$

where D_o is a Λ-generalized polynomial matrix and $\hat{C}$ is the right Λ-structure matrix of the system S.

Proof The transfer functions G_{ez} and G_{uz} are given by (7.1). They are both Λ-generalized polynomial matrices since S is Λ-stable. It follows from theorem 3.7 that there is a Λ-generalized polynomial matrix $\tilde{G}$ such that

$$G_{ez} = \tilde{G}\hat{C}. \tag{7.39}$$

Suppose that (7.38) is satisfied. Then (7.6) is satisfied. The transfer function Z is given by

$$Z = G_{ez} + G_{uz}D = (\tilde{G} + G_{uz}D_o)\hat{C}. \tag{7.40}$$

Here $(\tilde{G} + G_{uz}D_o)$ is a Λ-generalized polynomial matrix. Therefore (7.7) is satisfied. It follows from theorem 7.2 that $D \in \mathcal{D}$.

Conversely, if $D \in \mathcal{D}$ then (7.38) is valid by theorem 7.2.

□

Remark The special case of theorem 7.2 when z is the state vector has been shown in Bengtsson (1977b). Observe that the state vector cannot be reconstructed by a Luenberger observer since the disturbance e cannot be measured.

Example 7.1 Let the system S be described by the transfer function

$$\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \frac{p-2}{p-1} & \frac{2}{p+1} \\ \frac{1}{p-1} & \frac{p-1}{p+1} \end{pmatrix} \begin{pmatrix} u \\ e \end{pmatrix} \tag{7.41}$$

and let the unstable region Λ^* be the closed right half of the complex plane. Introduce

$$\lambda = \frac{1}{p+1}. \tag{7.42}$$

Then the corresponding unstable region Λ becomes the closed disc with radius $\frac{1}{2}$ and centre at $\frac{1}{2}$.

A fractional representation for (7.41) is given by

$$\begin{pmatrix} 1 & 1 \\ 0 & 1-2\lambda \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda \end{pmatrix} u + \begin{pmatrix} 1 \\ (1-2\lambda)^2 \end{pmatrix} e \tag{7.43}$$

The right Λ-structure matrix $\hat{C}$ is

$$\hat{C} = (1-2\lambda)^2 \tag{7.44}$$

Observe that theorem 3.7 cannot be used to calculate $\hat{C}$ because the system is unstable. The right Λ-structure matrix for the transfer function from e to z is $(1-2\lambda) \neq \hat{C}$.

A transfer function D must satisfy conditions (7.6) and (7.7) to be feedback realizable. First consider

$$D_1 = (1-2\lambda)^2. \tag{7.45}$$

The transfer function D_1 satisfies (7.6). The transfer function Z in condition (7.7) can be obtained from (7.43) as

$$Z = \frac{\lambda}{1-2\lambda} D_1 + \frac{(1-2\lambda)^2}{(1-2\lambda)} = (\lambda+1)(1-2\lambda), \tag{7.46}$$

which is Λ-stable. The right Λ-structure matrix $\hat{Z}$ of Z is

$$\hat{Z} = 1 - 2\lambda. \tag{7.47}$$

It follows that $\hat{C}$ is not a right division of $\hat{Z}$. Therefore condition (7.7) is not satisfied and D_1 is not feedback realizable. Now consider

$$D_2 = -2(1 - 2\lambda)^2. \tag{7.48}$$

The transfer function D_2 satisfies (7.6) and Z can be computed as in (7.46). Then

$$Z = \frac{\lambda}{1-2\lambda} D_2 + \frac{(1-2\lambda)^2}{(1-2\lambda)} = (1-2\lambda)^2, \tag{7.49}$$

which satisfies (7.7). Therefore D_2 is feedback realizable.

A feedback realization of D_2 can be obtained with the method in the proof of theorem 7.2.

Let D_2 be written as in (7.13), then

$$T = 1 \text{ and } V = -2 \tag{7.50}$$

The matrix Q in (7.14) is given by

$$Q = \lambda(-2) + 1 \cdot 1 = 1 - 2\lambda. \tag{7.51}$$

The matrices L, A_o and Q_o in (7.16) then become

$$L = (1-2\lambda), \quad A_o = 1 \quad \text{and} \quad Q_o = 1 \tag{7.52}$$

The matrix C_o in (7.20) vanishes since $\tilde{C} = 1$ is square. It then follows that

$$\overline{T} = T, \ \overline{V} = V \text{ and } \overline{Q} = Q_o \tag{7.53}$$

in (7.18). The feedback realization is obtained from (7.28) as

$$u = -2z. \tag{7.54}$$

The closed loop system is described by (7.43) and (7.54). This can be written

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1-2\lambda & \lambda^2 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} y \\ z \\ -u \end{pmatrix} = \begin{pmatrix} 1 \\ (1-2\lambda)^2 \\ 0 \end{pmatrix} e. \tag{7.55}$$

The closed loop system is Λ-stable since the determinant of the matrix in the left member of (7.55) is equal to 1. The transfer function from e to u can be computed from (7.55) as

$$u = -2(1 - 2\lambda)^2 e, \tag{7.56}$$

which is equal to D_2 as desired.

□

Example 7.2 Consider the discrete time system

$$Az = Bu + Ce \tag{7.57}$$

with

$$A = \begin{pmatrix} 1-0.9\lambda & 0.5\lambda \\ 0.5\lambda & 1-0.2\lambda \end{pmatrix}, \quad B = \begin{pmatrix} \lambda \\ 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1-0.2\lambda & -0.4\lambda \\ 0.2\lambda & 1-0.8\lambda \end{pmatrix}, \tag{7.58}$$

where $\lambda = q^{-1}$, the backward shift operator, and Λ is the closed unit disc with centre at the origin. It is assumed that the controlled output y is equal to the measured output z. The system is unstable because det A has a zero in Λ.

In example 8.2 of the next chapter the following feedforward controller is computed.

$$u = De, \tag{7.59}$$

where

$$D = (-0.82 - 0.042\lambda \quad 0.66 - 0.084\lambda). \tag{7.60}$$

A feedback realization of (7.59) will be computed. With the notation used in the proof of theorem 7.2 it follows that

$$\hat{C} = I \qquad \tilde{C} = C \tag{7.61 a}$$

$$T = I \qquad V = D. \tag{7.61 b}$$

Condition (7.6) of theorem 7.2 is thus satisfied. The matrix Q in (7.14) is given by

$$Q = \begin{pmatrix} \lambda \\ 0 \end{pmatrix} (-0.82 - 0.042\lambda \quad 0.66 - 0.084\lambda) + \begin{pmatrix} 1-0.2\lambda & -0.4\lambda \\ 0.2\lambda & 1-0.8\lambda \end{pmatrix} =$$

$$= \begin{pmatrix} 1 - 1.02\lambda - 0.042\lambda^2 & 0.26\lambda - 0.084\lambda^2 \\ 0.2\lambda & 1 - 0.8\lambda \end{pmatrix} =$$

$$= \begin{pmatrix} 1 - 0.9\lambda & 0.5\lambda \\ 0.5\lambda & 1 - 0.2\lambda \end{pmatrix} \begin{pmatrix} 1 - 0.12\lambda & -0.24\lambda \\ -0.3\lambda & 1 - 0.6\lambda \end{pmatrix}. \tag{7.62}$$

The matrices L, A_o and Q_o in (7.16) become

$$L = A \qquad A_o = I \qquad Q_o = \begin{pmatrix} 1 - 0.12\lambda & -0.24\lambda \\ -0.3\lambda & 1 - 0.6\lambda \end{pmatrix}. \tag{7.63}$$

It follows from (7.17) that Z becomes

$$Z = Q_o . \tag{7.64}$$

Therefore also condition (7.7) of theorem 7.2 is satisfied and the controller (7.59) - (7.60) is feedback realizable.

As in example 7.1 it follows that

$$\bar{T} = T, \ \bar{V} = V \text{ and } \bar{Q} = Q_o . \tag{7.65}$$

A feedback realization of (7.59) - (7.60) is now obtained from (7.28) as

$$\begin{pmatrix} 1 - 0.12\lambda & -0.24\lambda \\ -0.3\lambda & 1 - 0.6\lambda \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \tag{7.66 a}$$

$$u = (-0.82 - 0.042\lambda \quad 0.66 - 0.084\lambda) \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}. \tag{7.66 b}$$

The closed loop system is described by (7.57) and (7.66). This can be written

$$\begin{pmatrix} 1-0.9\lambda & 0.5\lambda & \lambda & 0 & 0 \\ 0.5\lambda & 1-0.2\lambda & 0 & 0 & 0 \\ -1 & 0 & 0 & 1-0.12\lambda & -0.24\lambda \\ 0 & -1 & 0 & -0.3\lambda & 1-0.6\lambda \\ 0 & 0 & 1 & -0.82-0.042\lambda & 0.66-0.084\lambda \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ -u \\ \xi_1 \\ \xi_2 \end{pmatrix} =$$

$$= \begin{pmatrix} 1-0.2\lambda & -0.4\lambda \\ 0.2\lambda & 1-0.8\lambda \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \tag{7.67}$$

The determinant of the matrix in the left member is $0.24\lambda^2 - \lambda + 1$, which has the zeros $\lambda_1 = \frac{5}{2}$ and $\lambda_2 = \frac{5}{3}$. The closed loop system is thus Λ-stable. A calculation confirms that the transfer function from e to u is D. The transfer function from e to z becomes

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 1-0.12\lambda & -0.24\lambda \\ -0.3\lambda & 1-0.6\lambda \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}, \tag{7.68}$$

which is the desired result in example 8.2 of the next chapter.

□

7.2 Measured disturbances

Assume that the disturbance vector can be partitioned as $e = (\varepsilon^T \ \ \delta^T)^T$, where ε can be measured, but δ not. The control configuration in figure 7.1 will then be replaced by the one in figure 7.2.

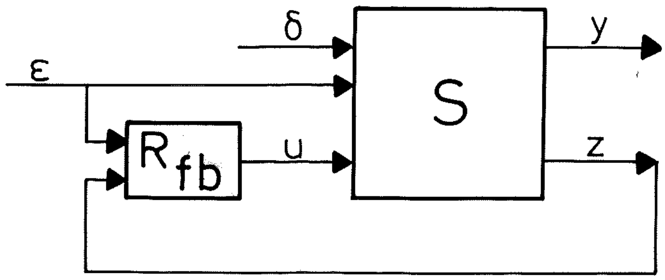


Figure 7.2 - The control configuration when a part of the disturbance vector can be measured.

Let the system S be described by the fractional representation

$$\begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} u + \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix} \begin{pmatrix} \varepsilon \\ \delta \end{pmatrix} . \tag{7.69}$$

Introduce the system $\bar{S}$ given by

$$\begin{pmatrix} A_1 & A_3 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} y \\ z \\ \varepsilon \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \\ 0 \end{pmatrix} u + \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \\ I & 0 \end{pmatrix} \begin{pmatrix} \varepsilon \\ \delta \end{pmatrix} . \tag{7.70}$$

The vector $\bar{z} = (z^T \ \ \varepsilon^T)^T$ is the measured output vector. The block diagram in figure 7.2 is then equivalent to the block diagram in figure 7.3.

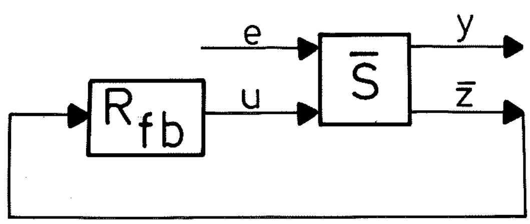


Figure 7.3. An equivalent control configuration when a part of the disturbance vector can be measured.

<u>Definition 7.3</u> The class $\mathcal{D}$ of feedback realizable transfer functions for the system S, where ε can be measured, is defined as the class of all transfer functions from e to u that can be obtained with admissible feedback controllers in figure 7.2.

The class $\mathcal{D}$ is thus equal to the class of feedback realizable transfer functions, given by definition 7.1, for the system $\bar{S}$. Let the class $\mathcal{F}$ be given by definition 7.2, but with figure 7.2 substituted for figure 7.1. Then theorem 7.1 is valid also in the case where ε is measured.

Let $\hat{C}_4$ be the right Λ-structure matrix for the matrix C_4. Then

$$\hat{C}_o = \begin{pmatrix} I & 0 \\ 0 & \hat{C}_4 \end{pmatrix} \tag{7.71}$$

is the right Λ-structure matrix for the system $\bar{S}$.

A characterization of the class $\mathcal{D}$ can now be obtained by application of theorem 7.2 to the system $\bar{S}$. The special case when the whole disturbance vector e can be measured is given below. Let G_{ez} and G_{uz} be given by (7.1).

<u>Theorem 7.3</u> Assume that the disturbance vector e can be measured. Then $D \in \mathcal{D}$ if and only if both D and $G_{uz}D + G_{ez}$ are Λ-generalized polynomial matrices.

<u>Proof</u> The right Λ-structure matrix for $\bar{S}$ is I. The matrix Z in theorem 7.2 is substituted by

$$\bar{Z} = \begin{pmatrix} G_{uz}D + G_{ez} \\ I \end{pmatrix} \tag{7.72}$$

when S is substituted by $\bar{S}$,

The matrix Z is a Λ-generalized polynomial matrix if and only if $G_{ez} + G_{uz}D$ is. The theorem now follows from theorem 7.2, applied to $\bar{S}$.

□

Example 7.3 Consider the system in example 7.1, i.e.

$$\begin{pmatrix} 1 & 1 \\ 0 & 1-2\lambda \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda \end{pmatrix} u + \begin{pmatrix} 1 \\ (1-2\lambda)^2 \end{pmatrix} e, \tag{7.73}$$

but assume that the disturbance e can be measured. The feedforward controller

$$u = D_1 e \tag{7.74 a}$$

$$D_1 = (1 - 2\lambda)^2 \tag{7.74 b}$$

was, in example 7.1, shown to be not feedback realizable if e cannot be measured.

The transfer function D_1 is Λ-stable and from (7.47) it follows that

$$G_{uz}D_1 + G_{ez} = (\lambda+1)(1-2\lambda) \tag{7.75}$$

is Λ-stable. Therefore theorem 7.3 shows that D_1 is feedback realizable if e can be measured.

Observe that the system cannot be controlled by the feedforward controller (7.74) because (7.73) is unstable. A feedback realization is therefore necessary.

Introduce the system $\bar{S}$ as

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1-2\lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ z \\ e \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda \\ 0 \end{pmatrix} u + \begin{pmatrix} 1 \\ (1-2\lambda)^2 \\ 1 \end{pmatrix} e. \tag{7.76}$$

A feedback realization of (7.74) for the system $\bar{S}$ is obtained in the following way. With the notation of the proof of theorem 7.2 it follows that

$$\hat{C} = 1 \qquad \tilde{C} = \begin{pmatrix} (1-2\lambda)^2 \\ 1 \end{pmatrix} \tag{7.77 a}$$

$$V = (1-2\lambda)^2 \qquad T = 1. \tag{7.77 b}$$

The matrix Q in (7.14) is

$$Q = \begin{pmatrix} \lambda \\ 0 \end{pmatrix} (1-2\lambda)^2 + \begin{pmatrix} (1-2\lambda)^2 \\ 1 \end{pmatrix} = \begin{pmatrix} (1+\lambda)(1-2\lambda)^2 \\ 1 \end{pmatrix}. \tag{7.78}$$

The matrices L, A_o and Q_o in (7.16) are

$$L = \begin{pmatrix} 1-2\lambda & 0 \\ 0 & 1 \end{pmatrix} \qquad A_o = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad Q_o = \begin{pmatrix} (1+\lambda)(1-2\lambda) \\ 1 \end{pmatrix}. \tag{7.79}$$

The matrix C_o in (7.20) can be taken as

$$C_o = \begin{pmatrix} 1 \\ 0 \end{pmatrix} . \tag{7.80}$$

The g.c.l.d. L_1 of L and B_2 is $L_1 = I$. Then (7.21) gives

$$Y = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad T_1 = 1. \tag{7.81}$$

The equation (7.22) becomes

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1-2\lambda & 0 \\ 0 & 1 \end{pmatrix} Q_1 - \begin{pmatrix} \lambda \\ 0 \end{pmatrix} V_1, \tag{7.82}$$

which has the solution

$$Q_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad V_1 = -2. \tag{7.83}$$

From (7.18) it follows that

$$\bar{V} = ((1-2\lambda)^2 \quad -2) \tag{7.84 a}$$

$$\bar{Q} = \begin{pmatrix} (1+\lambda)(1-2\lambda) & 1 \\ 1 & 0 \end{pmatrix} \tag{7.84 b}$$

A feedback realization is now obtained from (7.28).

$$u = \bar{V}\bar{Q}^{-1}A_o \begin{pmatrix} z \\ e \end{pmatrix} = (-2 \quad 3(1-2\lambda)) \begin{pmatrix} z \\ e \end{pmatrix} \Leftrightarrow u = -2z + 3(1-2\lambda)e. \tag{7.85}$$

The closed loop system is described by (7.73) and (7.85). It can be written

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1-2\lambda & \lambda \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} y \\ z \\ -u \end{pmatrix} = \begin{pmatrix} 1 \\ (1-2\lambda)^2 \\ -3(1-2\lambda) \end{pmatrix} e \tag{7.86}$$

The determinant of the matrix in the left member of (7.86) is equal to 1. Therefore the closed loop system is Λ-stable. The transfer function from e to u is given by

$$u = (1-2\lambda)^2 e, \tag{7.87}$$

which is the same as (7.74) as desired.

8. THE REGULATOR PROBLEM

The regulator problem was formulated in chapter 4. An admissible feedback controller R_{fb} should be determined so that the transfer function from e to y in the closed loop system in figure 8.1 satisfies given specifications.

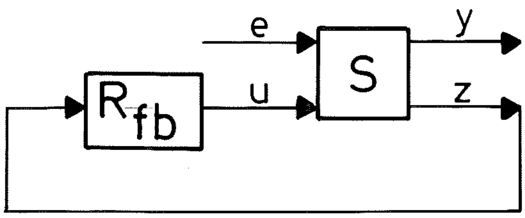


Figure 8.1 - The system with a feedback controller.

An admissible feedback controller is by (4.28) causal and such that the closed loop system is Λ-stable.

In this chapter the class $\mathcal{F}$ of all transfer functions that can be obtained from e to y, with admissible feedback controllers in figure 8.1, will be characterized. Necessary and sufficient conditions for certain types of transfer functions to belong to $\mathcal{F}$ will be given. The types, that are examined, are those with poles within a specified region of the complex plane or those, which do not transmit certain kinds of disturbances. A generalization of the minimum variance controller in Åström (1970) will also be given.

8.1 A characterization of the class of transfer functions from the disturbance to the controlled output

In this section the class $\mathcal{F}$, given by definition 7.2, will be characterized.

Let the system S be described by the fractional representation

$$A\begin{pmatrix} y \\ z \end{pmatrix} = Bu + Ce, \qquad (8.1)$$

where A and (B C) are relatively left prime polynomial matrices. Furthermore, let the transfer function for S be partitioned as

$$\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} G_{uy} & G_{ey} \\ G_{uz} & G_{ez} \end{pmatrix} \begin{pmatrix} u \\ e \end{pmatrix} . \tag{8.2}$$

It follows from theorem 7.1 that the class $\mathcal{F}$ can be obtained as all transfer functions F that are given by

$$F = G_{uy}D + G_{ey}, \tag{8.3}$$

where D belongs to the class $\mathcal{D}$ of feedback realizable transfer functions. In theorem 7.2 it was shown that $D \in \mathcal{D}$ if and only if D and

$$Z = G_{uz}D + G_{ez} \tag{8.4}$$

satisfy conditions (7.6) and (7.7).

The equalities (8.3) and (8.4) can be written

$$\begin{pmatrix} F \\ Z \end{pmatrix} = \begin{pmatrix} G_{uy} \\ G_{uz} \end{pmatrix} D + \begin{pmatrix} G_{ey} \\ G_{ez} \end{pmatrix} . \tag{8.5}$$

Use (8.1) to rewrite (8.5) as

$$\begin{pmatrix} F \\ Z \end{pmatrix} = A^{-1}BD + A^{-1}C, \tag{8.6}$$

which is equivalent to

$$A\begin{pmatrix} F \\ Z \end{pmatrix} = BD + C. \tag{8.7}$$

If A is partitioned as

$$A = (A_1 \quad A_2) \tag{8.8}$$

then (8.7) can be written

$$C = (-B \quad A_2)\begin{pmatrix} D \\ Z \end{pmatrix} + A_1F. \tag{8.9}$$

Let $\hat{C}$ be the right Λ-structure matrix of the system S. It follows from theorem 7.2 that $D \in \mathcal{D}$ if and only if there is a Λ-generalized polynomial matrix Y, such that

$$\begin{pmatrix} D \\ Z \end{pmatrix} = Y\hat{C}. \tag{8.10}$$

Insert (8.10) into (8.9), then

$$C = (-B \quad A_2)Y\hat{C} + A_1F. \tag{8.11}$$

It has then been shown that $F \in \mathcal{F}$ if and only if F satisfies (8.11) for some Λ-generalized polynomial matrix Y.

Let $\hat{K}$ be the left Λ-structure matrix of the polynomial matrix $(-B \quad A_2)$. Then there is a right Λ-invertible matrix $\tilde{K}$, such that

$$\hat{K}\tilde{K} = (-B \quad A_2). \tag{8.12}$$

The matrix $\tilde{K}$ is in fact a polynomial matrix by lemma 2.5. The equality (8.11) can be written

$$C = \hat{K}\tilde{K}Y\hat{C} + A_1F. \tag{8.13}$$

There is a Λ-generalized polynomial matrix $\bar{K}$, such that

$$\tilde{K}\bar{K} = I. \tag{8.14}$$

Define X as

$$X = \tilde{K}Y. \tag{8.15}$$

Lemma 8.1 There is a Λ-generalized polynomial matrix Y satisfying (8.13) if and only if there is a Λ-generalized polynomial matrix X satisfying

$$C = \hat{K}X\hat{C} + A_1F. \tag{8.16}$$

Proof If Y is a Λ-generalized polynomial matrix satisfying (8.13), then X, given by (8.15) is a Λ-generalized polynomial matrix satisfying (8.16). Conversely, if X is a Λ-generalized polynomial matrix satisfying (8.16), then

$$Y = \bar{K}X \tag{8.17}$$

is a Λ-generalized polynomial matrix satisfying (8.13).

□

The following theorem has now been shown.

Theorem 8.1 Let A be partitioned as (8.8) and let $\hat{K}$ be the left Λ-structure matrix of the matrix $(-B \quad A_2)$. Then $F \in \mathcal{F}$ if and only if there is a Λ-generalized polynomial matrix X, such that

$$C = \hat{K}X\hat{C} + A_1F. \tag{8.18}$$

Corollary Assume that the whole disturbance vector e can be measured, the $F \in \mathcal{F}$ if and only if there is a Λ-generalized polynomial matrix X, such that

$$C = \hat{K}X + A_1F. \tag{8.19}$$

Theorem 8.1 can be used in the following way to check if a given transfer function F_o belongs to $\mathcal{F}$. Compute the matrix $C - A_1F$. Check if $\hat{K}$ is a left Λ-divisor of $C - A_1F$. If so, compute R such that $C - A_1F = \hat{K}R$. Check if $\hat{C}$ is a right Λ-divisor of R. If so, F_o belongs to $\mathcal{F}$ otherwise it does not.

The only algorithm, that is needed to do this, is an algorithm to check if a given left Λ-structure matrix is a left Λ-divisor of a given Λ-generalized polynomial matrix.

If F = 0 then the disturbance e does not affect y at all. It follows from theorem 8.1 that F = 0 belongs to $\mathcal{F}$ if and only if C can be written

$$C = \hat{K}X\hat{C} \tag{8.20}$$

for some Λ-generalized polynomial matrix X. In fact, it follows from lemma 2.5 that if there is such a Λ-generalized polynomial matrix X, then X must be polynomial.

The problem to check if F = 0 can be obtained by state feedback was formulated and solved in Wonham, Morse (1970). Theorem 8.1 generalizes this result to an arbitrary vector z of measured outputs. The problem was in Wonham, Morse (1970) called the disturbance localization problem.

If the disturbance e can be measured it follows from (8.19) that F = 0 belongs to $\mathcal{F}$ if and only if $\hat{K}$ is a left Λ-divisor of C. This problem is discussed in Bengtsson, Wonham (1976) in the special case when the state vector can be measured.

In the following sections it will be shown how theorem 8.1 can be used to solve certain design problems.

8.2. Pole placement

The closed loop system is by (4.28) required to be Λ-stable. This implies that all $F \in \mathcal{F}$ are Λ-stable. In other words F has all its poles in the complement of Λ. In addition it might be required that the poles of F lie in a subset of the complement of Λ.

Let Ω be a subset of C, such that it fulfils (3.9) - (3.11) and $\Omega \supset \Lambda$. Assume that F is required to have all its poles in the complement of Ω, i.e. F is required to be Ω-stable. The closed loop system is still only required to be Λ-stable.

The problem to find necessary and sufficient conditions for the existence of an Ω-stable $F \in \mathcal{F}$ can be solved in the following way.

It follows from theorem 8.1 that the problem is to find a Λ-generalized polynomial matrix X and an Ω-generalized polynomial matrix F, that satisfy (8.18). Observe that the set of Ω-generalized polynomials is a subset of the set of Λ-generalized polynomials.

The following lemma is a straightforward generalization of lemma 2.5. It can be proved analogously. Therefore the proof is omitted here.

Lemma 8.2 Suppose that $\Lambda \subset \Omega$. Let $\hat{K}$ be a left Λ-structure matrix and P be an Ω-generalized polynomial matrix. If there is a Λ-generalized polynomial matrix R, such that

$$\hat{K}R = P, \tag{8.21}$$

then R is an Ω-generalized polynomial matrix.

Remark The lemma is equivalent to lemma 2.5 if $\Omega = C$.

If there exist a Λ-stable X and an Ω-stable F, which satisfy (8.18), then it follows from lemma 8.2 and its dual version that X must be Ω-stable. Consequently, there exist a Λ-stable X and an Ω-stable F, satisfying (8.18), if and only if there are Ω-stable R and F, satisfying

$$C = \hat{K}R + A_1F \tag{8.22}$$

and such that R can be written

$$R = X\hat{C} \tag{8.23}$$

for some Ω-stable X.

Lemma 8.3 Let L be the g.c.l.d of $\hat{K}$ and A_1. Then L is square and nonsingular and det L has no zeros in Λ. Furthermore, there are Ω-stable R and F satisfying (8.22) if and only if det L has no zeros in Ω.

Proof It follows from (8.12) that L is a left divisor of $(-B \quad A_2)$, since $\tilde{K}$ is a polynomial matrix. Therefore L is a left divisor of A and B. Since A is square and nonsingular and L is a left divisor of A it follows that also L is square and nonsingular. Furthermore det L has no zeros in Λ since the system is assumed to be Λ-stabilizable.

Suppose that det L has no zeros in Ω. It follows from theorem 2.7 that there are polynomial matrices R_o and F_o, such that

$$L = \hat{K}R_o + A_1F_o. \tag{8.24}$$

Multiply (8.24) by $L^{-1}C$ from the right. This gives

$$C = \hat{K}R_oL^{-1}C + A_1F_oL^{-1}C. \tag{8.25}$$

Then $R = R_oL^{-1}C$ and $F = F_oL^{-1}C$ are both Ω-stable and satisfy (8.22).

Conversely, suppose that R and F are Ω-stable and satisfy (8.22). Let K_o and A_o be polynomial matrices, such that

$$\hat{K} = LK_o \tag{8.26 a}$$

$$A_1 = LA_o. \tag{8.26 b}$$

Then it follows from (8.22) and (8.26) that

$$C = L(K_oR + A_oF) \Leftrightarrow L^{-1}C = K_oR + A_oF. \tag{8.27}$$

Consequently $L^{-1}C$ is Ω-stable. If the polynomial matrices L and C are relatively left prime, then it follows that det L has no zeros in Ω.

Suppose L_o is a common left divisor of L and C. Then L_o is also a common left divisor of A and (B C) since it has been shown that L is a common left division of A and B. The polynomial matrices A and (B C) are, however, relatively left prime since 8.1 is a fractional representation. Therefore L_o is unimodular and L and C are relatively left prime.

□

Suppose that (R,F) is an Ω-stable solution to (8.22). All solutions to (8.22) can then be obtained using theorem 2.11. Let $(P^T \quad Q^T)^T$ be a polynomial basis for the nullspace of $(\hat{K} \quad A_1)$. Then any Ω-stable solution$(\bar{R},\bar{F})$can be written

$$\bar{R} = R + PN \tag{8.28 a}$$

$$\bar{F} = F + QN, \tag{8.28 b}$$

for some Ω-stable N. It is required by (8.23) that $\bar{R}$ can be written $\bar{R} = X\hat{C}$ for some Ω-stable X. Insert this into (8.28 a)

$$X\hat{C} = R + PN \qquad R = -PN + X\hat{C}. \tag{8.29}$$

Here R, P and $\hat{C}$ are given and (8.29) should be solved for Ω-stable N and X.

It follows from theorem 2.12 that an Ω-stable solution (N,X) to (8.29) exists if and only if the two Ω-generalized polynomial matrices

$$\begin{pmatrix} P & R \\ 0 & \hat{C} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} P & 0 \\ 0 & \hat{C} \end{pmatrix} \tag{8.30}$$

are Ω-equivalent.

The results are summarized in the following theorem.

<u>Theorem 8.2</u> Let A be partitioned as in (8.8) and let $\hat{K}$ be the left Λ-structure matrix of $(-B \quad A_2)$. Furthermore let L be the

g.c.l.d. of $\hat{K}$ and A_1 and let $(P^T \ \ Q^T)^T$ be a polynomial basis for the nullspace of $(\hat{K} \ \ A_1)$. Then there is an Ω-stable $F \in \mathcal{F}$ if and only if the following two conditions are satisfied.

- The polynomial det L has no zeros in Ω. (8.31)

- Any Ω-stable solution (R,F) to $C = \hat{K}R + A_1F$ is such that

$$\begin{pmatrix} P & R \\ 0 & \hat{C} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} P & 0 \\ 0 & \hat{C} \end{pmatrix}$$

are Ω-equivalent. (8.32)

Remark 1 Observe that condition (8.31) implies, by lemma 8.3, that an Ω-stable solution (R,F) in condition (8.32) exists.

Remark 2 Observe that if the condition (8.32) is satisfied for one solution (R,F), then it is satisfied for all solutions since

$$\begin{pmatrix} P & R \\ 0 & \hat{C} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} P & R + PN \\ 0 & \hat{C} \end{pmatrix}$$

are Ω-equivalent for all Ω-stable N.

If follows from the corollary of theorem 8.1 that the case, when the disturbance e can be measured, is obtained if $\hat{C}$ is substituted by I. In this case condition (8.32) is always satisfied. Therefore condition (8.31) is necessary and sufficient for the existence of an Ω-stable $F \in \mathcal{F}$ in the case when e can be measured. If e cannot be measured, then the required information about e has to be obtained from the measured output z. Condition (8.32) is necessary and sufficient for this to be possible.

Example 8.1 Consider a discrete time system. The operator λ is then given by

$$\lambda = \frac{1}{q-a}, \tag{8.33}$$

where q is the forward shift operator. Choose a = 0, then λ becomes equal to the backward shift operator q^{-1}.

$$q^{-1}x(t) = x(t-1) \tag{8.34}$$

Let Λ be the closed unit disc and Ω the whole complex plane. The closed loop system is thus required to be asymptotically stable in the sense of Lyapunov and the transfer function from e to y is required to be a polynomial matrix in q^{-1}.

Consider the following system

$$\begin{pmatrix} \lambda-2 & (\lambda-2)(-2\lambda+3) \\ 0 & -2 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \lambda(\lambda-2) \\ \lambda \end{pmatrix} u + \begin{pmatrix} 1 \\ 1 \end{pmatrix} e. \tag{8.35}$$

For u = 0 this system gives

$$y = \frac{-2\lambda^2 + 7\lambda - 4}{2\lambda - 4}\, e = \frac{4q^2 - 7q + 2}{4q^2 - 2q}\, e. \tag{8.36}$$

A feedback controller from z to u will be determined such that the transfer function from e to y becomes Ω-stable (i.e. a polynomial in λ).

$$(-B \quad A_2) = \begin{pmatrix} -\lambda(\lambda-2) & (\lambda-2)(-2\lambda+3) \\ -\lambda & -2 \end{pmatrix} = \begin{pmatrix} -\lambda(\lambda-2) & -\lambda+3 \\ -\lambda & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & \lambda-2 \end{pmatrix} \triangleq$$

$$\triangleq \hat{K}\tilde{K} \tag{8.37}$$

The matrix $\hat{K}$ and A_1 are relatively left prime. Therefore condition (8.31) is satisfied. Since $\hat{C} = 1$, also condition (8.32) is satisfied. An Ω-stable solution to $C = \hat{K}R + A_1F$ is given by

$$R = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and } F = 1. \tag{8.38}$$

A Λ-stable (D,Z) and Ω-stable F, which solves

$$C = (-B \quad A_2)\begin{pmatrix} D \\ Z \end{pmatrix} + A_1 F \Leftrightarrow A\begin{pmatrix} F \\ Z \end{pmatrix} = BD + C, \tag{8.39}$$

is given by

$$\begin{pmatrix} D \\ Z \end{pmatrix} = \tilde{K}^{-1} R = \begin{pmatrix} -\frac{1}{\lambda-2} \\ \frac{1}{\lambda-2} \end{pmatrix} \quad \text{and } F = 1. \tag{8.40}$$

The feedforward controller

$$u = De = -\frac{1}{\lambda-2}\, e \tag{8.41}$$

thus gives

$$y = Fe = e. \tag{8.42}$$

A feedback realization of (8.41) can be computed with the method in the proof of theorem 7.2. This gives

$$u = -z. \tag{8.43}$$

Insert (8.43) into (8.35), then

$$\begin{pmatrix} \lambda-2 & (\lambda-2)(-\lambda+3) \\ 0 & \lambda-2 \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e, \tag{8.44}$$

which is equivalent to

$$\begin{pmatrix} \lambda-2 & 0 \\ 0 & \lambda-2 \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \lambda-2 \\ 1 \end{pmatrix} e. \tag{8.45}$$

This system is Λ-stable and the transfer function from e to y is equal to 1.

□

8.3. Shortest correlation regulators for discrete time systems

The results of the previous section will be specialized in the following way. The system is supposed to be a discrete time system. This means that the operator μ is equal to the forward shift operator q, defined through

$$qx(t) = x(t+1). \tag{8.46}$$

The operator λ is in chapter 3 defined through

$$\lambda = \frac{1}{\mu-a}. \tag{8.47}$$

Choose a = 0, then λ becomes the backward shift operator q^{-1}, which has the property

$$q^{-1}x(t) = x(t-1). \tag{8.48}$$

The unstable region Λ is chosen as the closed unit disc and Ω is chosen as the whole complex plane. This means that the closed loop system is required to be asymptotically stable in the sense of Lyapanov and that the transfer function F from e to y is required to be a polynomial matrix in q^{-1}.

The measured output vector z is supposed to be equal to the controlled output vector y. Therefore the system S can be described by the fractional representation

$$Ay = Bu + Ce \tag{8.49}$$

The disturbance e is supposed to be a sequence of uncorrelated stochastic variables with zero mean value. Furthermore C is supposed to be square and nonsingular and det C has no zeros inside the closed unit disc. This implies that the right Λ-structure matrix $\hat{C}$ of the system S is equal to I. The question of the validity of the assumptions on the C matrix is a question of how a stochastic disturbance on the output y can be modelled. If such a disturbance can be represented as white noise, filtered through a square, nonsingular, stable and non-minimum phase transfer function, then the assumptions on C are valid for stable systems S. This question will not be further discussed here. In the scalar case it is discussed in Åström (1970).

Since $z = y$ the development in section 8.1 can be simplified in the following way. Let D belong to $\mathcal{D}$ and F be the desired transfer function from e to y. Then

$$F = A^{-1}BD + A^{-1}C \Leftrightarrow C = -BD + AF. \tag{8.50}$$

F is the transfer function from e to y, but also from e to z, since $z = y$. It must therefore be both Λ-stable and Ω-stable, i.e. polynomial. However, if F is polynomial, then it is also Λ-stable. Since $\hat{C} = I$ it is only required that D is Λ-stable for it to be feedback realizable.

Let $\hat{B}$ be the left Λ-structure matrix of B. It follows from lemma 8.1 and 8.2 that there is a Λ-stable D and a polynomial F, satisfying (8.50) if and only if there are polynomial X and F, satisfying

$$C = \hat{B}X + AF. \tag{8.51}$$

From lemma 8.3 it follows that (8.51) has a polynomial solution (X,F) if and only if $\hat{B}$ and A are relatively left prime.

Let $(P^T \;\; Q^T)^T$ be a polynomial basis for the nullspace of $(\hat{B} \;\; A)$ and let X_o, F_o be a polynomial solution to (8.51). It follows from theorem 2.11 that any polynomial solution (X,F) to (8.51) can be written

$$X = X_o + PN \tag{8.52 a}$$

$$F = F_o + QN \tag{8.52 b}$$

for some polynomial matrix N.

The results are summarized in the following theorem.

Theorem 8.3 Let $\hat{B}$ be the left Λ-structure matrix of B and let $(P^T \; Q^T)^T$ be a polynomial basis for the nullspace of $(\hat{B} \; A)$. Then there exists a polynomial matrix $F \in \mathcal{F}$ if and only if $\hat{B}$ and A are relatively left prime. Let $F_o \in \mathcal{F}$ be polynomial. Then any polynomial matrix $F \in \mathcal{F}$ can be written as

$$F = F_o + QN \tag{8.53}$$

for some polynomial matrix N.

The expression (8.53) can be used to compute an $F \in \mathcal{F}$ with desirable properties. Three types of "shortest correlation regulators" will be given. The computations will only be briefly sketched.

Type 1. The concepts of column degree and column properness are defined in part 1 of this thesis. It is possible to find an F, which has lowest possible column degree in each column, in the following way.

Suppose that Q is column proper. This is no restriction because it can always be achieved by multiplication from the right by a unimodular matrix. Let f_i be column i in F_o. If f_i does not have lowest possible degree then a linear combination (by polynomials) of those columns in Q, which do not have higher degree than f_i, is subtracted from f_i so that its degree decreases. This is repeated until it is no longer possible to decrease the degree of f_i. Then it can be shown that f_i has lowest possible degree. The same scheme is applied to the other columns of F_o. As a result a polynomial matrix $F_1 \in \mathcal{F}$ is obtained and every column of F_1 has lowest possible degree.

If f_i is column i in F_1 and e_i is component i in e, then

$$y = F_1 e = f_1 e_1 + f_2 e_2 + \ldots + f_k e_k. \tag{8.54}$$

Let d_i be the column degree of f_i, then $f_i e_i$ is a moving average of length d_i+1. This means that $y(t)$ and $e_i(t-\tau)$ are uncorrelated if $\tau \geq d_i+1$. (The different components of e are supposed to be uncorrelated). Furthermore this is the shortest correlation between y and e_i, that can be obtained by any controller, since d_i is the lowest possible degree of column i.

□

Type 2. The columns of Q can be shown to be linearly independent because $\hat{B}$ has linearly independent columns. Suppose that Q is lower left triangular in the sense that the first nonzero entry in any column and the entries above it has only entries, that are zero, to the right. This is no restriction since it can be achieved by multiplication from the right by a unimodular matrix. Let d_i be the degree of the first nonzero entry in column i and let m_i be the number of the corresponding row.

Let F_o be a polynomial matrix in F. The first m_1-1 rows of F_o are not affected if multiples of the columns of Q are added to the columns of F_o, since the first m_1-1 rows of Q are zero. Add multiples (by polynomials) of the first column in Q to the columns in F_o so that all entries in row m_1 of F_o get a degree less than d_1. Repeat the procedure with the second column in Q to get a degree less than d_2 in row m_2 of F_o. This does not affect the first m_2-1 rows of F_o. Repeat the procedure with all the columns of Q.

The result is a matrix $F_1 \in F$ which has lowest possible degree of any row j, provided that the degree of the rows above it have been minimized sequentially.

Consequently this controller gives a shortest possible correlation between y_1 and e. It gives a shortest possible correlation between y_2 and e, provided that the correlation between y_1 and e has been minimized, and so on. Clearly this controller favours the first components of y. Therefore it is supposed that the components of y are ordered so that y_1 is most important, then y_2 and so on. □

The controllers of types 1 and 2 can be seen as generalizations of the minimum variance controller for single input single output systems in Åström (1970). If further restrictions are applied to the system, then a controller, which minimizes the variance of the output y, can be obtained. Assume that

$$\hat{B} = q^{-k}I, \tag{8.55}$$

for some positive integer k. This is a severe restriction on the system, because it means that the time delay in every path through the system S, from any input to any output, is the same and equal to k. Furthermore it implies that the system is minimum phase and has at least as many inputs as outputs.

Type 3. Suppose that $\hat{B} = q^{-k}I$. Then $\hat{B}$ and A are relatively left prime and a polynomial F_o in $\mathcal{F}$ exists. Furthermore, it is easy to see that P and Q in (8.52) are given by

$$P = -A \tag{8.56}$$

$$Q = q^{-k}I. \tag{8.57}$$

Any polynomial F in $\mathcal{F}$ can by (8.53) be written

$$F = F_o + q^{-k}N. \tag{8.58}$$

Consequently N can be chosen such that all entries of F have a degree less than k. This controller is in Borisson (1975) shown to minimize the variance of y^TMy, for any positively definite constant matrix M.

□

<u>Example 8.2</u> Consider the system in example 7.2, i.e.

$$Ay = Bu + Ce \tag{8.59}$$

with

$$A = \begin{pmatrix} 1-0.9\lambda & 0.5\lambda \\ 0.5\lambda & 1-0.2\lambda \end{pmatrix}, \quad B = \begin{pmatrix} \lambda \\ 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1-0.2\lambda & -0.4\lambda \\ 0.2\lambda & 1-0.8\lambda \end{pmatrix}. \tag{8.60}$$

The system is unstable because det A has a zero inside Λ, but it is Λ-stabilizable from u. Furthermore, det C has no zeros in Λ.

For this system $\hat{B} = B$ and the matrices $\hat{B}$ and A are relatively left prime. Therefore it follows from theorem 8.3 that there exists a polynomial $F \in \mathcal{F}$. A polynomial solution to (8.51) is given by

$$X_o = (1.1-0.266\lambda-0.0196\lambda^2 \quad -0.58-0.004\lambda-0.0056\lambda^2) \tag{8.61 a}$$

$$F_o = \begin{pmatrix} 1-0.4\lambda+0.056\lambda^2 & -0.32\lambda+0.016\lambda^2 \\ -0.3\lambda+0.14\lambda^2 & 1-0.6\lambda+0.04\lambda^2 \end{pmatrix}. \tag{8.61 b}$$

A polynomial basis $(P^T \; Q^T)^T$ for the nullspace of $(\hat{B} \;\; A)$ is

$$Q = \begin{pmatrix} -\lambda+0.2\lambda^2 \\ 0.5\lambda^2 \end{pmatrix} \qquad P = 1-1.1\lambda-0.07\lambda^2. \tag{8.62}$$

A shortest correlation regulator of type 1 can be constructed in the following way.

Add -0.28 times Q to the first column of F_o and -0.08 times Q to the second column of F_o. This corresponds to a matrix

$$N = (-0.28 \quad -0.08) \tag{8.63}$$

in (8.52) and gives

$$F_1 = F_o + QN = \begin{pmatrix} 1-0.12\lambda & -0.24\lambda \\ -0.3\lambda & 1-0.6\lambda \end{pmatrix} \tag{8.64 a}$$

$$X_1 = X_o + PN = (0.82 + 0.042\lambda \quad -0.66 + 0.084\lambda). \tag{8.64 b}$$

It is not possible to further decrease the column degrees of F_1. Therefore a controller of type 1 gives F_1 as the transfer function from e to y in the closed loop system.

Furthermore, the row degrees of F_1 cannot be decreased by addition of a multiple of Q. Therefore, also a controller of type 2 will give F_1 as a transfer function from e to y.

It follows from (8.50) and (8.51) that

$$D = -X_1, \tag{8.65}$$

since $\hat{B} = B$. Therefore a feedback realization of

$$u = -X_1 e \tag{8.66}$$

is a controller of type 1 (and of type 2) for the system (8.60). Such a feedback realization was computed in example 7.2, where it also was confirmed that the transfer function from e to y becomes F_1. □

8.4. Disturbance rejection

Suppose that the disturbance e can be represented as the impulse response of a dynamical system, i.e.

$$e = T\delta, \tag{8.67}$$

where δ is an impulse and T a rational transfer function. Disturbances like steps, ramps and sinusoids can be represented in this way. In this section a regulator, which prevents such disturbances from being transmitted to y, will be given. The configuration is shown in figure 8.2.

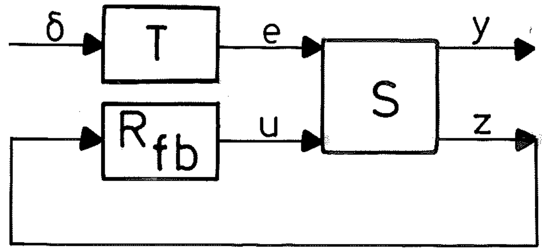


Figure 8.2 - The closed loop system with a disturbance generator.

The transfer function T can be written as

$$T = R^{-1}M, \tag{8.68}$$

where R and M are relatively left prime polynomial matrices. Let Ω be a subset of C, such that it fulfils (3.9) - (3.11) and $\Lambda \subset \Omega$. The disturbance rejection problem is formulated as follows. Find a causal controller R_{fb}, such that the transfer functions from δ to y and from e to y in figure 8.2 both are Ω-stable and such that the closed loop system in figure 8.1 is Λ-stable. Observe that it is not required that the total system in figure 8.2 is Λ-stable. The transfer function T might not be Λ-stable and the controller R_{fb} can do nothing to change that.

It follows from theorem 8.1 that $F \in \mathcal{F}$ if and only if there is a Λ-stable X such that

$$C = \hat{K}X\hat{C} + A_1F, \tag{8.69}$$

where $\hat{K}$ and A_1 are given by (8.12) and (8.8). Let E be the transfer function from δ to y. Then E is given by

$$E = FR^{-1}M \tag{8.70}$$

It is required that both E and F are Ω-stable. This implies that

$$Y = FR^{-1} \tag{8.71}$$

must be Ω-stable, since R and M are relatively left prime. It follows that E and F are Ω-stable if and only if F can be written as

$$F = YR \tag{8.72}$$

for some Ω-stable Y. Introducing (8.72) into (8.69) gives

$$C = \hat{K}X\hat{C} + A_1YR. \tag{8.73}$$

It follows from (8.73) that if Y is Ω-stable then $\hat{K}X\hat{C}$ must be Ω-stable. Lemma 8.2 and its dual version then imply that X must be Ω-stable. Consequently the disturbance rejection problem has a solution if and only if there are Ω-stable X and Y that satisfy (8.73).

Equation (8.73) can be rewritten if the Kronecker product is used. Let c, x and y be the column vectors consisting of the column vectors of C, X and Y. Then equation (8.73) can be written

$$[\hat{C}^T \otimes \hat{K} \quad R^T \otimes A_1]\begin{pmatrix} x \\ y \end{pmatrix} = c, \tag{8.74}$$

where $\otimes$ denotes the Kronecker product. Let $\hat{L}$ be the left Ω-structure matrix of $[\hat{C}^T\otimes\hat{K} \quad R^T\otimes A_1]$. Then there is a right Ω-invertible matrix $\tilde{L}$, such that

$$\hat{L}\tilde{L} = [\hat{C}^T \otimes \hat{K} \quad R^T \otimes A_1]. \tag{8.75}$$

It can be shown as in lemma 8.1 that there are Ω-stable x and y, satisfying 8.51 if and only if there is an Ω-stable m, satisfying

$$\hat{L}m = c. \tag{8.76}$$

It follows from lemma 2.5 that if there is an Ω-stable m, satisfying (8.76), then m is a polynomial vector.

The results are summarized in the following theorem.

Theorem 8.4 Let the disturbance e be the impulse response of a system with transfer function $T = R^{-1}M$ and let A be partitioned as in (8.8). Furthermore, let $\hat{K}$ be the left Λ-structure matrix of $(-B \;\; A_2)$ and let $\hat{L}$ be the left Ω-structure matrix of $(\hat{C}^T \otimes \hat{K} \;\; R^T \otimes A_1)$. Finally, let c be the column vector of columns in C. Then the disturbance rejection problem has a solution if and only if one of the following two equivalent conditions is satisfied.

(i) There exist Ω-stable X and Y satisfying

$$C = \hat{K}X\hat{C} + A_1YR. \tag{8.77}$$

(ii) There exists a polynomial vector m, satisfying

$$\hat{L}m = c, \tag{8.78}$$

i.e. $\hat{L}$ is a left divisor of c.

It was pointed out previously that the controller R_{fb} in figure 8.2 does not Ω-stabilize the system T. It only makes the unstable modes of T unobservable from y. This is done in the following way. If, for instance, e is a ramp, then the controller R_{fb} generates a ramp, which annihilates the effect of the ramp e. In other words, the controller R_{fb} contains a model of the system T, that generates the disturbance. In Bengtsson (1977a) it is shown that such an internal model must always be present in order to annihilate the effect of the disturbance.

Example 8.3 Consider the discrete time system

$$\begin{pmatrix} 1 & 1 \\ 0 & 1+2\lambda \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda \end{pmatrix} u + \begin{pmatrix} 1 \\ (1-2\lambda)^2 \end{pmatrix} e, \tag{8.79}$$

where $\lambda = q^{-1}$, the backward shift operator, and let Λ and Ω both be the closed unit disc.

The system is unstable and has a non-minimum phase behaviour in the transmission from e to z. Assume that it is required that a feedback from z to u should be determined such that step disturbances are not transmitted from e to y.

A step disturbance can be described as

$$e = \frac{1}{1-\lambda}\,\delta, \qquad \delta = \begin{cases} 1, & t = 0 \\ 0, & t \neq 0 \end{cases} \tag{8.80}$$

With the notation of this section it is found that

$$R = (1-\lambda) \qquad M = 1 \qquad \hat{C} = (1-2\lambda)^2 \tag{8.81 a}$$

$$A_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 1 \\ 1+2\lambda \end{pmatrix} \tag{8.81 b}$$

$$(-B \quad A_2) = \begin{pmatrix} -1 & 1 \\ -\lambda & 1+2\lambda \end{pmatrix} = \hat{K}. \tag{8.81 c}$$

Equation (8.77) becomes

$$\begin{pmatrix} 1 \\ (1-2\lambda)^2 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ -\lambda & 1+2\lambda \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} (1-2\lambda)^2 + \begin{pmatrix} 1 \\ 0 \end{pmatrix} Y(1-\lambda), \tag{8.82}$$

which is equivalent to

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -(1-2\lambda)^2 & (1-2\lambda)^2 & 1-\lambda \\ -\lambda & 1+2\lambda & 0 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ Y \end{pmatrix}. \tag{8.83}$$

This equation has the following solution

$$\begin{pmatrix} X_1 \\ X_2 \\ Y \end{pmatrix} = \begin{pmatrix} 1-2\lambda \\ 1-\lambda \\ 1+4\lambda^2 \end{pmatrix}. \tag{8.84}$$

It follows from (8.9) and (8.77) that the feedforward controller

$$u = De \tag{8.85}$$

is given by

$$D = X_1\hat{C} = (1-2\lambda)^3 \tag{8.86}$$

since $\hat{K} = (-B \quad A_2)$.

A feedback realization of (8.85) - (8.86) is given by

$$u = \frac{1-2\lambda}{1-\lambda}\, z \tag{8.87}$$

$$\Leftrightarrow (1-\lambda)u = (1-2\lambda)z. \tag{8.88}$$

The closed loop system is described by equations (8.79) and (8.88), which can be written

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1+2\lambda & \lambda \\ 0 & 1-2\lambda & 1-\lambda \end{pmatrix} \begin{pmatrix} y \\ z \\ -u \end{pmatrix} = \begin{pmatrix} 1 \\ (1-2\lambda)^2 \\ 0 \end{pmatrix} e. \tag{8.89}$$

The determinant of the matrix in the left member of (8.89) is equal to 1. Therefore the closed loop system is Λ-stable.

The transfer function from e to y is

$$y = (1+4\lambda^2)(1-\lambda)e. \tag{8.90}$$

This transfer function makes the step disturbance unobservable from y. Insert (8.80) into (8.90), then

$$y = (1+4\lambda^2)\delta. \tag{8.91}$$

The feedback controller (8.87) thus has the desired properties.

Observe that the controller (8.87) includes an integrator and that it is an integrator that generates the disturbance in (8.80). The controller thus includes a model of the system, that generates the disturbance. This "internal model" creates an input, which annihilates the effect of the disturbance on the output y.

□

8.5. Decoupling

In this section necessary and sufficient conditions for the existence of a diagonal $F \in \mathcal{F}$ will be given. The problem is solved only in the case when either the disturbance or the state can be measured. If e can be measured then $\hat{C} = I$ and if the measured output z is the state of the system then $\hat{C} = \lambda I$. In both cases equation (8.18) can be written

$$C = \hat{R}X + A_1F, \qquad (8.92)$$

where

$$\hat{R} = \hat{K} \qquad (8.93)$$

if $\hat{C} = I$ and

$$\hat{R} = \lambda\hat{K} \qquad (8.94)$$

if $\hat{C} = \lambda I$. In both cases $\hat{R}$ is a left Λ-structure matrix.

Assume that F is diagonal

$$F = \text{diag}\ (f_1, \ldots, f_k). \qquad (8.95)$$

Let c_i, x_i and a_i be column i in C, X and A_1. Then (8.92) can be written

$$c_i = \hat{R}x_i + a_if_i, \quad i = 1, \ldots, k. \qquad (8.96)$$

Consequently the decoupling problem has a solution if and only if there are Λ-stable x_i and f_i satisfying (8.96) for i = 1,..., k.

Let $\hat{R}_i$ be the left Λ-structure matrix of $(\hat{R}\ a_i)$. Then it can be shown as in section 8.4 that there are Λ-stable solutions x_i, f_i to (8.96) if and only if there are polynomial vectors m_i, satisfying

$$c_i = \hat{R}_i m_i, \quad i = 1, \ldots, k. \qquad (8.97)$$

The following theorem has been shown.

<u>Theorem 8.5</u> Let a_i and c_i be the i:th column vectors of A_1 and C and let $\hat{K}$ be the left Λ-structure matrix of $(-B\ \ A_2)$. Furthermore, let $\hat{R} = \hat{K}$, if the disturbance can be measured, or $\hat{R} = \lambda\hat{K}$, if the state can be measured. Finally, let $\hat{R}_i$ be the left Λ-structure matrix of $(\hat{R}\ \ a_i)$. Then there exists a diagonal $F \in \mathcal{F}$ if and only if there are polynomial vectors $\{m_i\}$, satisfying

$$c_i = \hat{R}_i m_i, \quad i = 1, \ldots, k, \qquad (8.98)$$

where k is the number of components in e (and in y).

9. A DESIGN EXAMPLE

The design method described in chapters 5 - 8 will be applied to a continuous time system S characterized by

$$y = G^*(p)w, \tag{9.1}$$

where

$$G^*(p) = \begin{pmatrix} \frac{1}{p+1} & \frac{2}{p+3} \\ \frac{1}{p+1} & \frac{1}{p+1} \end{pmatrix} . \tag{9.2}$$

This system was, in Rosenbrock (1966), taken as an example of a system that presents control difficulties due to non-minimum phase behaviour. The system has a zero in the right half of the complex plane. In examples 5.1 and 5.5 the left Λ-structure matrix for this system was calculated and it was shown what can be done with feedforward control.

9.1. The control requirements

The controlled output y is equal to the measured output z and it is assumed that a nonmeasurable disturbance e is added to the input. The control configuration is shown in figure 9.1.

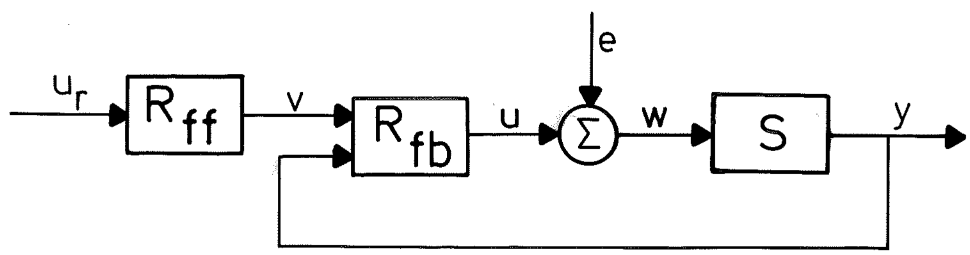


Figure 9.1. The control configuration.

The control problem is to find linear controllers R_{fb} and R_{ff}, such that the following requirements are fulfilled.

- The transfer function from u_r to y should be diagonal. (9.3)

- A step disturbance e should give no steady state error in the output y. (9.4)

- The closed loop system should be asymptotically stable in the sense of Lyapunov. (9.5)

Requirement (9.5) implies that the unstable region Λ^* should be chosen as the closed right half of the complex plane. Choose λ as

$$\lambda = \frac{1}{p+2} \ . \qquad (9.6)$$

This gives an unstable region Λ, which is the closed disc with radius $\frac{1}{4}$ and centre at $\frac{1}{4}$ (see example 5.1). Furthermore $\Omega = \Lambda$.

A fractional representation for the system S is given by

$$Ay = Bu + Ce, \qquad (9.7)$$

where

$$A = \begin{pmatrix} 1-\lambda^2 & 0 \\ 0 & 1-\lambda \end{pmatrix} \quad \text{and } B = C = \lambda \begin{pmatrix} 1+\lambda & 2-2\lambda \\ 1 & 1 \end{pmatrix} . \qquad (9.8)$$

The right Λ-structure matrix $\hat{C}$ of the system is equal to C.

The determinant of $\hat{C}$ is $\lambda^2(3\lambda-1)$, which has a zero at $\lambda = \frac{1}{3}$. This means that the disturbance e itself cannot be obtained from measurements of y. Only e, filtered through a non-minimum phase system, can be obtained.

9.2. The feedback controller

The step disturbance e can be described as an integrated impulse, i.e.

$$e = R^{-1}M\delta, \tag{9.10}$$

where

$$R = \begin{pmatrix} 1-2\lambda & 0 \\ 0 & 1-2\lambda \end{pmatrix}, \quad M = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \tag{9.11}$$

and δ is a dirac pulse.

Let F be the transfer function from e to y and D the transfer function from e to u in the closed loop system. Then

$$AF = BD + C. \tag{9.12}$$

It follows from theorem 7.2 that D is feedback realizable if and only if there are Λ-stable X and Y_1, such that

$$D = X\hat{C} \tag{9.13}$$

$$F = Y_1\hat{C}. \tag{9.14}$$

It was shown in section 8.4 that the requirement (9.4) can be fulfilled if and only if there is a Λ-stable Y_2, such that

$$F = Y_2R. \tag{9.15}$$

Since R commutes with any matrix the requirements (9.14) and (9.15) are fulfilled if and only if there is a Λ-stable Y_o, such that

$$F = Y_oR\hat{C}. \tag{9.16}$$

Insert (9.13) and (9.16) into (9.12). Then

$$C = -B X \hat{C} + A Y_o R \hat{C}, \tag{9.17}$$

which is equivalent to

$$I = -BX + AY_o R \tag{9.18}$$

because $C = \hat{C}$ and C has linearly independent rows. There are Λ-stable X and Y_o, satisfying (9.18), if and only if there are Λ-stable X and Y satisfying

$$I = -BX + RY, \tag{9.19}$$

since R commutes with any matrix and det A has no zeros in Λ. A solution to (9.19) is given by

$$X = \begin{pmatrix} -7+6\lambda & 6-4\lambda \\ 4.5+3\lambda & -5-2\lambda \end{pmatrix} \text{ and } Y = \begin{pmatrix} 1 & -4\lambda \\ -4.5\lambda & 1+3\lambda \end{pmatrix}. \tag{9.20}$$

A feedback realization of (9.13) is obtained as in the proof of theorem 7.2, i.e.

$$RY\xi = Ay \tag{9.21 a}$$

$$u = X\xi. \tag{9.21 b}$$

According to theorem 6.1 the input v should be determined so that the left Λ-structure matrix of the system is preserved. A general method is given in theorem 6.1. It is shown below that the left Λ-structure matrix is preserved, in this case, if v is applied in the following way.

$$RY\xi = Ay + v \tag{9.22 a}$$

$$u = X\xi. \tag{9.22 b}$$

The transfer function from v to y in the closed loop system then becomes

$$G_{vy} = A^{-1}BX. \tag{9.23}$$

The right Λ-structure matrix of G_{vy} is equal to the right Λ-structure matrix of

$$G = A^{-1}B, \tag{9.24}$$

since det X has no zeros in Λ. Consequently, the right Λ-structure matrix is preserved if v is applied as in (9.22).

A state space representation of (9.22) is given by

$$\dot{x} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 14 & -4 & 0 & -12 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 4.5 & 0 & 0 & -5 \end{pmatrix} x + \begin{pmatrix} 1 & 0 \\ 0 & 4 \\ 17 & -24 \\ 0 & 1 \\ 4.5 & -4 \end{pmatrix} y + \begin{pmatrix} 1 & 0 \\ 0 & 4 \\ 18 & -20 \\ 0 & 1 \\ 4.5 & -3 \end{pmatrix} v \tag{9.25 a}$$

$$u = \begin{pmatrix} -8 & -7 & 0 & 8 & 6 \\ 8 & 2.5 & 0 & -12 & -5 \end{pmatrix} x + \begin{pmatrix} -7 & 6 \\ 2.5 & -5 \end{pmatrix} y + \begin{pmatrix} -7 & 6 \\ 2.5 & -5 \end{pmatrix} v. \tag{9.25 b}$$

9.3. The feedforward controller

It was shown in example 5.5 that diagonal transfer functions from u_r to y of the form

$$H^* = \begin{pmatrix} a\,\frac{p-1}{m(p)} & 0 \\ 0 & b\,\frac{p-1}{n(p)} \end{pmatrix} \tag{9.26}$$

can be obtained. Here m(p) and n(p) are polynomials of degree 2 and with zeros in the open left half plane.

Two choices of H* will be examined. The first choice is

$$H_1^* = \begin{pmatrix} \frac{-p+1}{p^2+p+1} & 0 \\ 0 & \frac{-p+1}{p^2+p+1} \end{pmatrix}. \qquad (9.27)$$

The diagonal entries both have a steady state gain equal to 1 and a pole-zero configuration as is shown in figure 9.2.

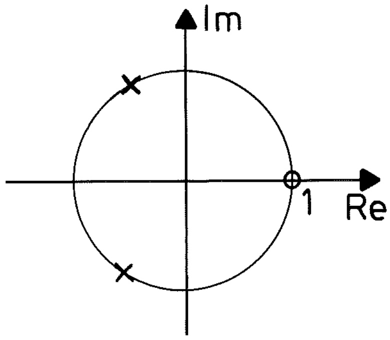


Figure 9.2. The pole-zero configuration of the diagonal entries of H_1^*.

The corresponding transfer function, expressed in λ, is given by

$$H_1 = \begin{pmatrix} \frac{-1+3\lambda}{1-3\lambda+3\lambda^2} & 0 \\ 0 & \frac{-1+3\lambda}{1-3\lambda+3\lambda^2} \end{pmatrix}. \qquad (9.28)$$

The equation

$$G_{vy}K_1 = H_1, \qquad (9.29)$$

where G_{vy} is given by (9.23), has a unique solution K since G_{vy} is square and has full rank. The solution is

$$K_1 = D_1^{-1}N_1, \qquad (9.30)$$

where

$$D_1 = (1-3\lambda+\lambda^2)\begin{pmatrix} 2 & 4-8\lambda \\ 0 & -10+12\lambda \end{pmatrix} \tag{9.31}$$

$$N_1 = \begin{pmatrix} (1-\lambda^2)(1-3\lambda) & 0 \\ (1-\lambda^2)(-2.25+4.5\lambda) & 1-\lambda \end{pmatrix} \tag{9.32}$$

A state space representation of (9.30) is given by

$$\dot{x} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -2 & -3 & -3 & 1.6 & 1.6 & 1.6 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -0.8 & -1.8 & -1.8 \end{pmatrix} x + \begin{pmatrix} -0.090 & 0.240 \\ -0.578 & -0.392 \\ 0.562 & 0.114 \\ 0.495 & -0.320 \\ -0.621 & -0.044 \\ 0.047 & 0.335 \end{pmatrix} u_r \tag{9.33 a}$$

$$v = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} x + \begin{pmatrix} 0.05 & 0.2 \\ 0.225 & -0.1 \end{pmatrix} u_r . \tag{9.33 b}$$

The second choice of H*-matrix is

$$H_2^* = \begin{pmatrix} \frac{3(-p+1)}{(p+1)(p+3)} & 0 \\ 0 & \frac{3(-p+1)}{(p+1)(p+3)} \end{pmatrix} . \tag{9.34}$$

The diagonal entries have a steady state gain equal to 1. The fact that H_2^* has the same poles as the transfer function G_{vy}^* from v to y makes it possible to find a feedforward controller of lower order.

The corresponding transfer function, expressed in λ, is given by

$$H_1 = \lambda \begin{pmatrix} \frac{3(-1+3\lambda)}{1-\lambda^2} & 0 \\ 0 & \frac{3(-1+3\lambda)}{1-\lambda^2} \end{pmatrix} . \tag{9.35}$$

The solution K_2 to

$$G_{vy}K_2 = H_2 \tag{9.36}$$

is

$$K_2 = D_2^{-1}N_2, \tag{9.37}$$

where

$$D_2 = \begin{pmatrix} 2 & 4-8\lambda \\ 0 & (1+\lambda)(-10+12\lambda) \end{pmatrix} \tag{9.38}$$

$$N_2 = 3\begin{pmatrix} 1-3\lambda & 0 \\ (1+\lambda)(-2.25+4.5\lambda) & 1 \end{pmatrix}. \tag{9.39}$$

A state space representation of (9.37) is given by

$$\dot{x} = \begin{pmatrix} -2 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & -2.4 & -3.8 \end{pmatrix} x + \begin{pmatrix} -0.240 & -0.360 \\ -0.180 & -0.020 \\ 0.144 & -0.084 \end{pmatrix} u_r \tag{9.40 a}$$

$$v = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} x + \begin{pmatrix} 0.15 & 0.6 \\ 0.675 & -0.3 \end{pmatrix} u_r. \tag{9.40 b}$$

9.4. Simulations

The system S has the following state space representation

$$\dot{x} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{pmatrix} x + \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 2 \end{pmatrix} w \tag{9.41 a}$$

$$y = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} x. \tag{9.41 b}$$

The closed loop system in figure 9.1 was simulated with (9.25) as R_{fb} and (9.33) as R_{ff}. Figure 9.3 shows the output y and control input u when a unit step in u_{r1} is applied at time t = 0 and a unit step in u_{r2} is applied at t = 10. The disturbance e is zero.

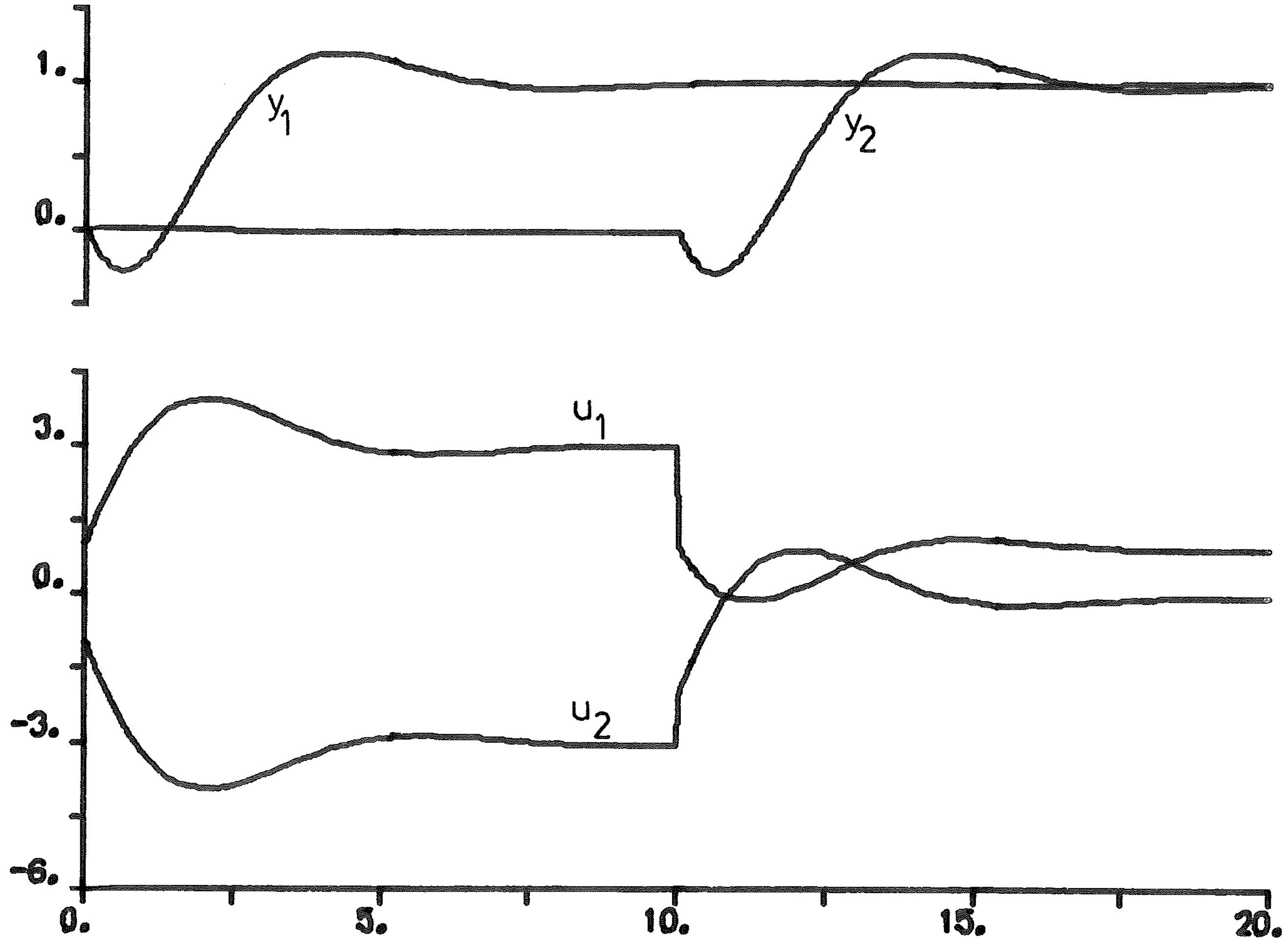


Figure 9.3. The output y and control input u, corresponding to step changes in the reference input u_r. R_{ff} is given by (9.33).

The system is completely decoupled. The non-minimum phase behaviour is due to the unavoidable zero at the point 1.

Figure 9.4 shows the corresponding responses with the feedforward controller R_{ff} given by (9.40).

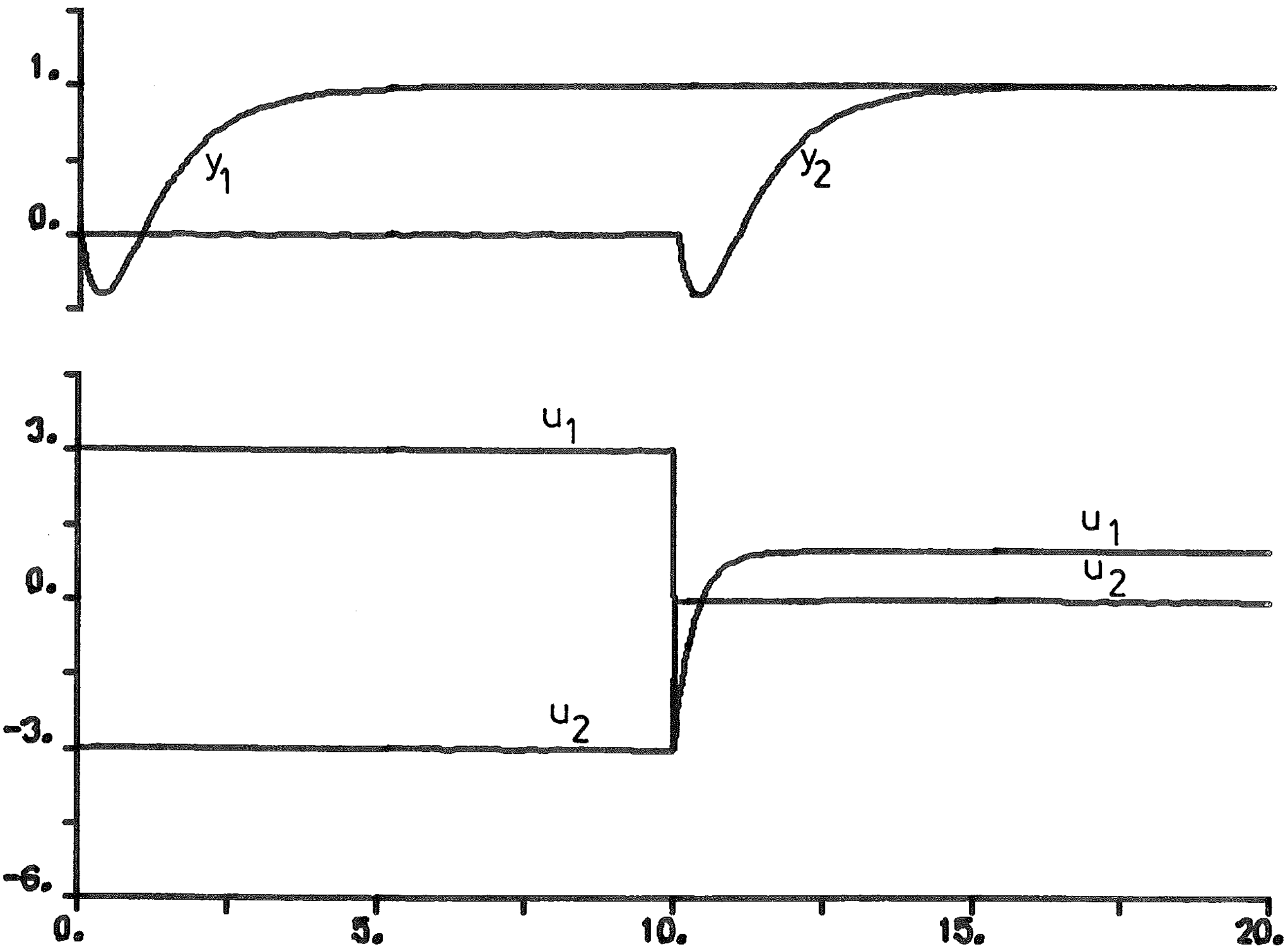


Figure 9.4. The output y and control input u, corresponding to step changes in the reference input u_r. R_{ff} is given by (9.40).

Even in this case the decoupling is complete. There is no overshoot, but the undershoot is larger. This feedforward controller is probably preferable because of its simplicity.

In figure 9.5 the output y and control input u are shown when a unit step in the disturbance e_1 is applied at t = 0 and a unit step in e_2 at t = 10. The reference input u_r is zero.

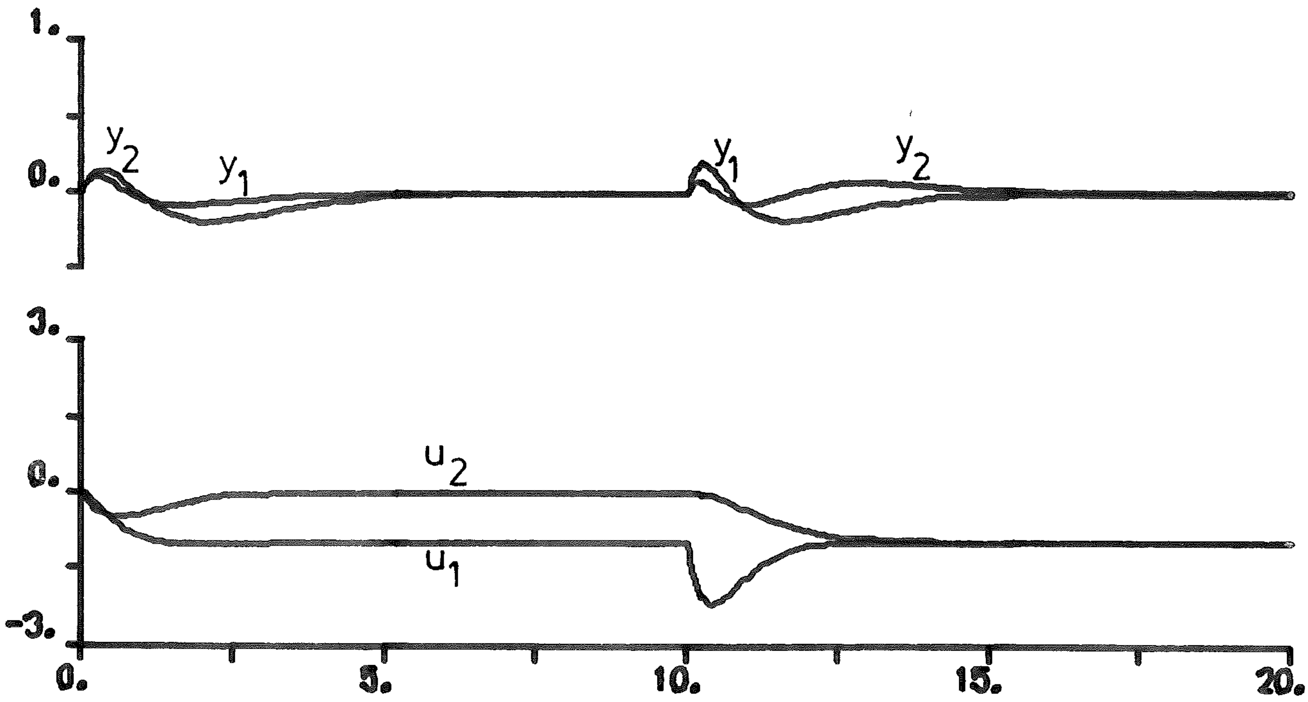


Figure 9.5. The output y and control input u, corresponding to step changes in the disturbance e.

There is no steady state error in y and the deviation from zero is small.

9.5. Robustness

The sensitivity to parameter changes in the system S was tested in the following way. The system (9.41) was substituted by the system

$$\dot{x} = \begin{pmatrix} -0.5 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -6 \end{pmatrix} x + \begin{pmatrix} 1.5 & 0 \\ 0 & 1.1 \\ 0 & 2.3 \end{pmatrix} w \tag{9.42 a}$$

$$y = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} x. \tag{9.42 b}$$

As in section 9.4 the closed loop system was simulated with a unit step in u_{r1} at t=0 and in u_{r2} at t=10. The disturbance e was zero. The result when the controller R_{ff} is given by (9.33) is shown in figure 9.6.

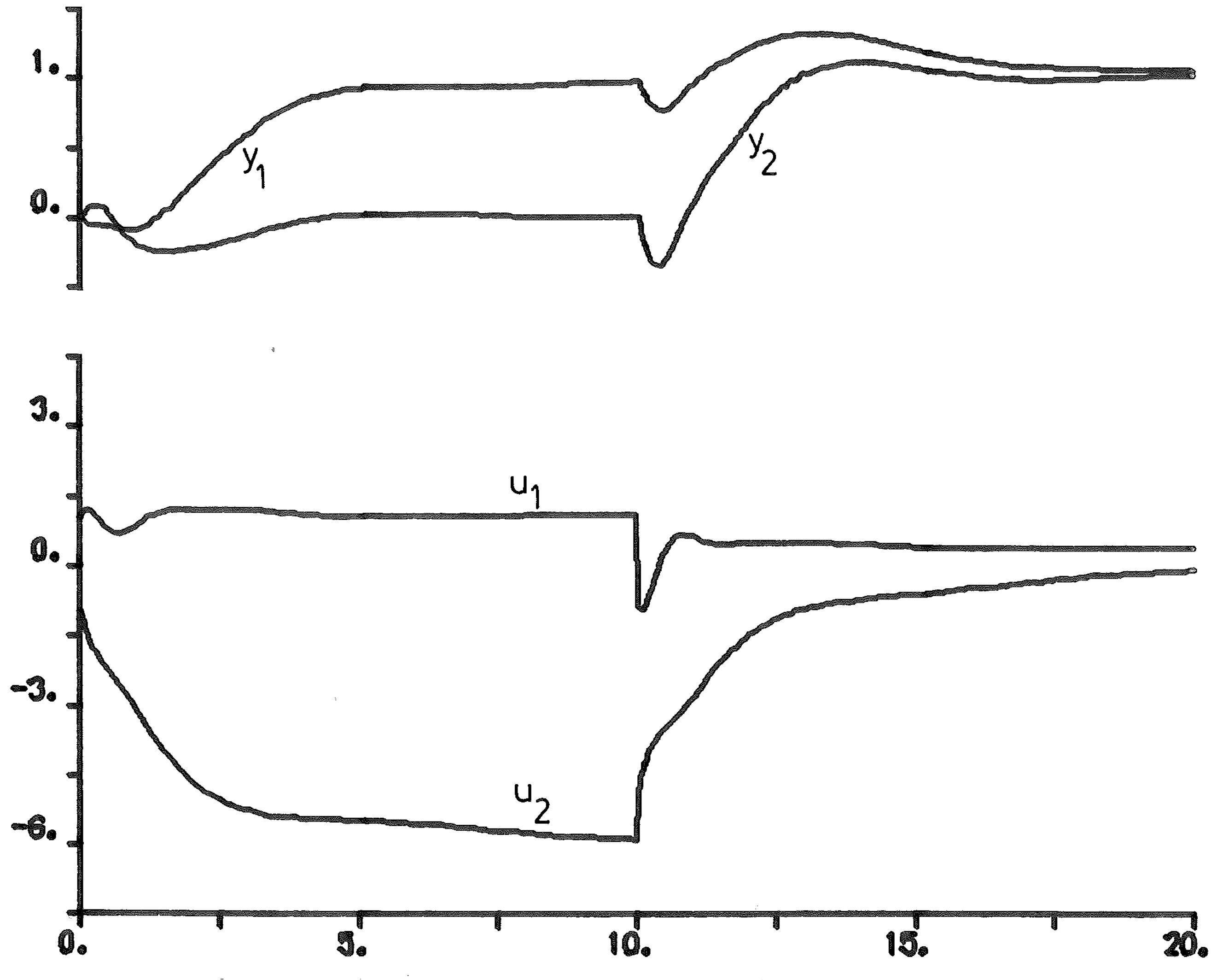


Figure 9.6. The output y and control input u, corresponding to step changes in the reference input u_r. R_{ff} is given by (9.33).

The corresponding result when the feedforward controller R_{ff} is given by (9.40) is shown in figure 9.7.

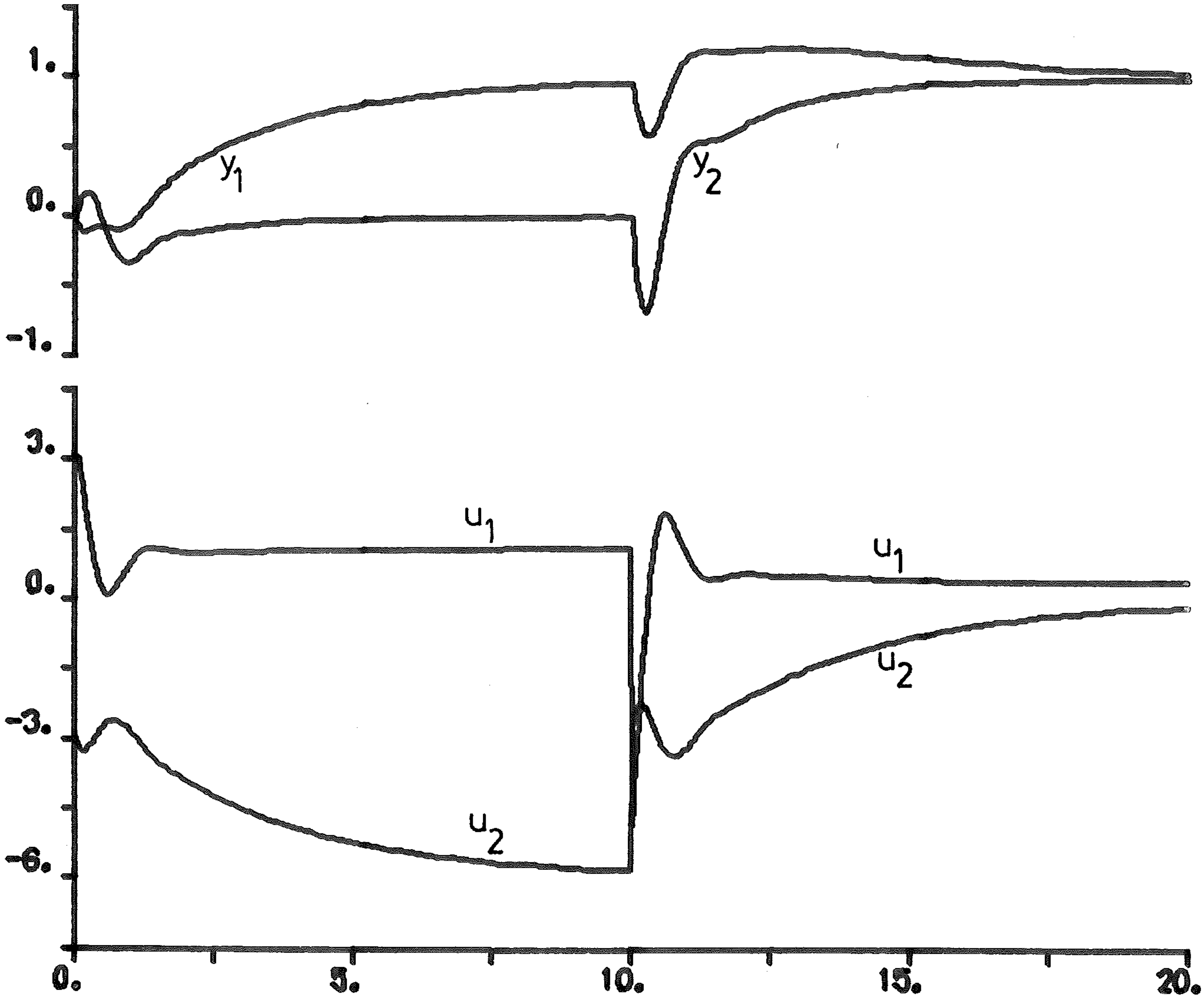


Figure 9.7. The output y and control input u, corresponding to step changes in the reference input u_r. R_{ff} is given by (9.40).

The system is no longer completely decoupled, but in steady state it is decoupled for both feedforward controllers.

In figure 9.8 a unit step is applied in the disturbance e_1 at $t = 0$ and in e_2 at $t = 10$. The reference input u_r is zero.

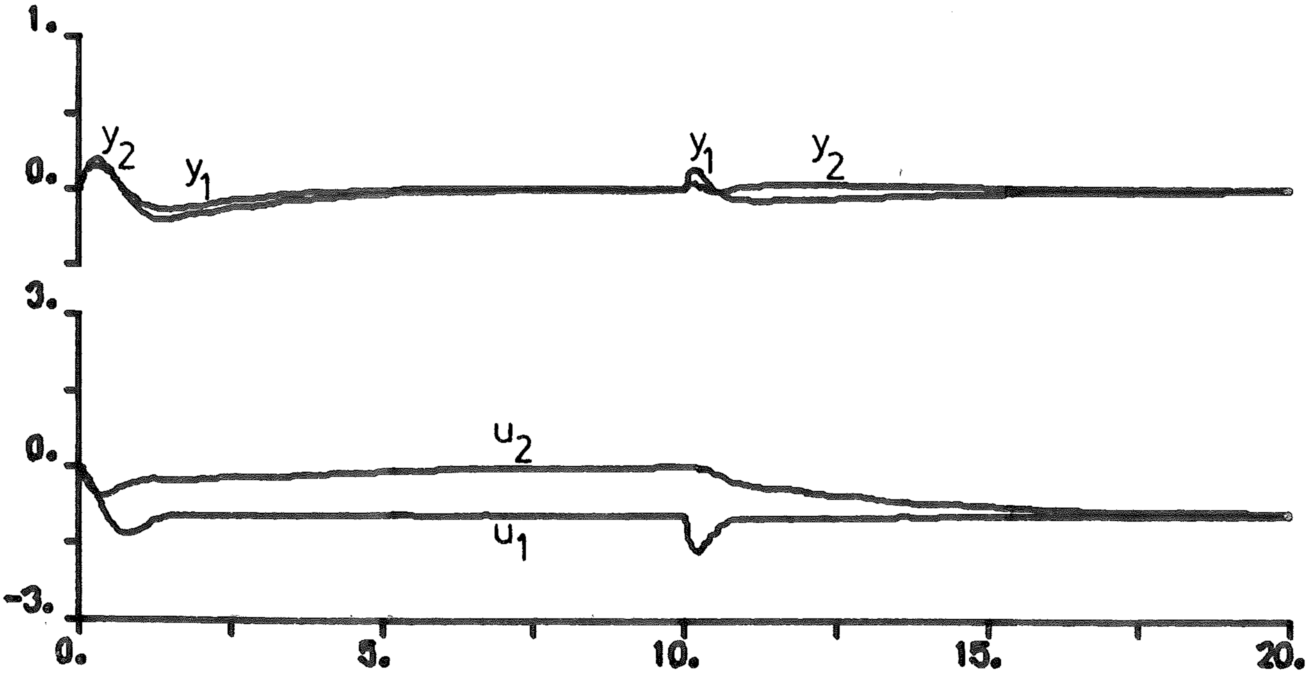


Figure 9.8. The output y and control input u, corresponding to step changes in the disturbance e.

Even in this case there is no steady state error in the output y and the deviation from zero is small.

10. REFERENCES

Åström K J (1970): Introduction to Stochastic Control Theory. Academic Press, New York.

Bengtsson G (1974): A Theory for Control of Linear Multivariable Systems. TFRT-1006, Dept of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Bengtsson G (1977a): Output Regulation and Internal Model - A Frequency Domain Approach. Automatica 13, 333-345.

Bengtsson G (1977b): Feedback Realizations in Linear Multivariable Systems. IEEE Transactions on Automatic Control AC-22, 576-585.

Bengtsson G, Wonham W M (1976): On Causal and Stable Solutions to Rational Matrix Equations with Applications to Control Theory. TFRT-7107, Dept of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Borisson U (1975): Self-Tuning Regulators - Industrial Application and Multivariable Theory. TFRT-1010, Dept of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Forney G D (1975): Minimal Basis of Rational Vector Spaces, with Applications to Multivariable Linear Systems. Siam Journal on Control 13, 493-520.

Mac Duffee C C (1946): The Theory of Matrices. Chelsea Publishing Company, New York.

Mac Farlane A G J, Postlethwaite I (1977): The Generalized Nyquist Stability Criterion and Multivariable Root Loci. International Journal of Control 25, 81-127.

Morse A S, Wonham W M (1971): Status of Noninteracting Control. IEEE Transactions on Automatic Control AC-16, 568-581.

Pernebo L (1978): Numerical Algorithms for Polynomial Matrix Systems. To appear.

Rosenbrock H H (1966): On the Design of Linear Multivariable Control Systems. Proceedings of the 3rd IFAC Congress, London.

Rosenbrock H H (1970): State-space and Multivariable Theory. Nelson, London.

Roth W E (1952): The Equations AX-YB=C and AX-XB=C in Matrices. Proceedings of the American Mathematical Society 3, 392-396.

Sain M K, Massey J L (1969): Invertibility of Linear Time--Invariant Dynamical Systems. IEEE Transactions on Automatic Control AC-14, 141-149.

Wang S H, Davison E J (1973a): A Minimization Algorithm for the Design of Linear Multivariable Systems. IEEE Transactions on Automatic Control AC-18, 220-225.

Wang S H, Davison E J (1973b): A Minimization Algorithm for the Design of Linear Multivariable Systems. Preprints JACC, AACC, Columbus, Ohio.

Wolovich W A (1974): Linear Multivariable Systems. Springer--Verlag, New York.

Wonham W M, Morse A S (1970): Decoupling and Pole Assignment in Linear Multivariable Systems: A Geometric Approach. Siam Journal on Control 8, 1-18.

SIGMA·TRYCK
TLTH · Lund