DIPLOMARBEIT

Changing Stability of Closed Loop Systems by Modifying the Cost Function

ausgeführt am

Institut für Analysis, Technische Mathematik und
Versicherungsmathematik der Technischen Universität Wien

unter Anleitung von
o.Univ.Prof.Dr. Heinz Langer

und am

Institut für Mathematik und Computer Wissenschaften
der Freien Universität Amsterdam

unter Anleitung von
dr. A.C.M. Ran

durch

Gerhard Reitmayr
Hagenauertalstraße 24
2372 Gießhübl

Datum

Unterschrift

Contents

List of Figures

Foreword

Mathematical systems theory is a widely used tool in todays engineering practice. Devices are commonly expressed in systems of linear ordinary differential equations and their properties and behaviour is addressed in the language of this theory. In this setting the eigenvalues of the matrix describing the system of equations ultimately define the properties of the system of interest. Usually the task of controlling a system is to change it in such a way that it attains some desired properties such as stability.

This work is concerned with one approach to control a system using Linear Quadratic Control. The control is a linear feedback of the system state and furthermore, satisfies a quadratic optimality constraint.

After a short introduction to the concepts and definitions used in mathematical systems theory, Chapter 2 gives a result on the influences of certain perturbations of the optimality condition on the resulting systems behaviour. Chapter 3 develops a pole-placement theorem for one dimensional input systems and Chapter 4 introduces an algorithm to accomplish the same for multi dimensional input systems. The last chapter contains examples showing the various results of this work.

This master thesis was developed to a great extend during my stay in Amsterdam in autumn 1998. While attending two lectures about mathematical systems theory, I was lucky to work with Prof. Ran from the Free University on the results presented in this work. During the spring and summer of 1999 the actual thesis was written under the supervision of Prof. Langer.

I enjoyed my stay in Amsterdam in many ways and had a very good time. Not only the city of Amsterdam itself was of interest, I also found a number of new friends with whom I hope to spend a lot more time in the years to come.

I would like to thank Prof. Langer for making my stay in Amsterdam possible and provide me with his experience during his supervision. Then Prof. Ran who introduced me to the subject and to mathematical research in general. He was a very good supervisor and friend who helped me through a lot of dark spots during the development of this work. Also to my friends in Amsterdam who usually provided me with a lot of distraction and good companionship after having worked too much. Finally I'm very grateful to my family who at all times supported my interests in mathematics and enabled me to reach this point.

1 Introduction

We will consider linear quadratic control problems. Such problems are composed of a system of linear differential equations together with a quadratic cost function. The aim of such a problem is to find an optimal control function such that the cost function is minimized for this control. To start we need some definitions and results from mathematical systems theory. A fairly complete discussion of these topics can be found in [4].

1.1 Linear Systems

Definition.

1. A linear continous-time time-invariant system $\Sigma = (A, B)$ consists of the matrices $A \in K^{n\times n}$ and $B \in K^{n\times m}$, where K is the field $\mathbb{C}$ or $\mathbb{R}$, and the equation

$$\dot{x}(t) = Ax(t) + Bu(t) \text{ for } t \in \mathbb{R}. \tag{1.1}$$

2. A linear continous-time time-invariant system $\Sigma = (A, B, C, D)$ with output consists of a system $\Sigma_0 = (A, B)$ and the matrices $C \in K^{p\times n}$ and $D \in K^{p\times n}$, where K is the field $\mathbb{C}$ or $\mathbb{R}$, and the equation

$$y(t) = Cx(t) + Du(t) \text{ for } t \in \mathbb{R}. \tag{1.2}$$

   The function $y(t)$ is called the output of the system Σ.

Remark 1.1. The system equation (1.1) defines a map

$$\phi : \begin{cases} \mathbb{R} \times \mathbb{R} \times K^n \times C(K^m) \to K^n \\ (\tau, \sigma, x_0, u) \to x(\tau; \sigma, x_0, u) \end{cases},$$

where $x(\tau; \sigma, x_0, u)$ is the solution of the differential equation (1.1) with initial value $x_0 = x(\sigma)$ evaluated at time τ. This map is called the *transition map*. The system with output defines a map $h : \mathbb{R} \times K^n \to K^p$ by (1.2), the so called *measurement map*. We will also write $y(\tau; \sigma, x_0, u)$ to denote the value of this map. The variation of constants formula gives

$$x(\tau; \sigma, x_0, u) = e^{(\tau-\sigma)A}x_0 + \int_\sigma^\tau e^{(\tau-\sigma)A}Bu(s)ds$$

for the transition map and likewise for the measurement map.

If $m = 1$ we call the system *single input*, if $p = 1$ *single output*. The elements $u(t) \in C(K^m)$ are called *controls*. Throughout the following we are only interested in real matrices and thus we will write $\mathbb{R}$ instead of the general field K. Furthermore we assume without loss of generality that the system will always start at time $\sigma = 0$ and therefore σ will be omitted in the following. This can be done because of the time-invariance of the system.

Definition. A system $\Sigma = (A, B)$ is said to be in *control canonical form*, if the matrices A and B have the following form :

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1r} \\ A_{21} & A_{22} & \dots & A_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ A_{r1} & A_{r2} & \dots & A_{rr} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & b_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots \\ 0 & 0 & \dots & b_r & 0 & \dots \end{pmatrix},$$

where the blocks A_{ij} have the form

$$A_{ij} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \\ * & * & \dots & * \end{pmatrix} \text{ for } i \neq j \text{ and } A_{ii} = \begin{pmatrix} 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ * & * & \dots & * \end{pmatrix},$$

and the blocks b_k are vectors of the form

$$b_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

The dimensions κ_k of the vectors b_k are called *controllability indices*. They are uniquely determined by (A, B).

Remark 1.2. For any controllable pair of matrices (A, B), there exists an invertible matrix T, such that $\widetilde{A} = T^{-1}AT, \widetilde{B} = T^{-1}B$ are in control canonical form. In the single input case $\widetilde{A}$ takes the form

$$\widetilde{A} = \begin{pmatrix} 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix},$$

and $\lambda^n - \alpha_n\lambda^{n-1} - \ldots - \alpha_2\lambda - \alpha_1$ is the characteristic polynomial of $\widetilde{A}$.

1.1.1 Controllability and Observability

Definition.

1. A system $\Sigma = (A, B)$ is called *controllable*, if for all x_0 and $x_1 \in \mathbb{R}^n$, there exists a control $u(t)$ such that the solution to (1.1) with initial state x_0 assumes the value x_1 for some $t \geq 0$.

2. A system with output $\Sigma = (A, B, C, D)$ is called *observable*, if for all pairs of initial values $x_0, x_1,\ x_0 \neq x_1$, there exists a control $u(t)$ such that for some time $t \geq 0$ the relation $y(t; x_0, u) \neq y(t; x_1, u)$ holds.

The following theorem gives a characterization of controllable and observable systems in terms of their matrices. A development of the theory and the proof can be found in [4].

Theorem 1.3.

(i) *The system $\Sigma = (A, B)$ is controllable, if and only if*

$$\operatorname{rank}(B, AB, \ldots, A^{n-1}B) = n$$

holds. The pair (A, B) is then also called controllable.

(ii) *The system with output $\Sigma = (A, B, C, D)$ is observable, if and only if*

$$\operatorname{rank}\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} = n.$$

holds. The pair (A, C) is then also called observable. This is equivalent to (A^*, C^*) being controllable.

There are some other characterizations of controllable pairs of matrices. The next theorem gives some equivalent conditions. The proofs can be found in [4].

Theorem 1.4. *Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$. Then the following statements are equivalent.*

(i) *(A, B) is controllable,*

(ii) $\operatorname{rank}(A - \lambda I, B) = n$ *for each $\lambda \in \mathbb{C}$,*

(iii) $\operatorname{rank}(A - \lambda I, B) = n$ *for each eigenvalue λ of A.*

1.1.2 Stability and Feedback Control

The differential equation

$$\dot{x} = Ax,$$

where $A \in \mathbb{R}^{n \times n}$, is called *asymptotically stable*, if for all initial values $x_0 = x(0)$

$$\lim_{t \to \infty} x(t) = 0$$

holds. This is equivalent to the fact that all eigenvalues of A are in the open left half of the complex plane. Furthermore, we can think of such an equation as a system $\Sigma = (A)$ without input.

For a system $\Sigma = (A, B)$, we are looking at a special kind of control. We are interested in a linear constant feedback control $u(t)$ of the form $u(t) = Fx(t)$, where $F \in \mathbb{R}^{m \times n}$. This is called feedback control, because the state of the system itself is used to compute the control. In such a case we also speak of a closed loop system.

The system Σ then becomes a system without input and with the system matrix $A + BF$. It is asymptotically stable, if the eigenvalues of $A + BF$ are in the open left half plane. The following theorem states that the matrix $A + BF$ can achieve any set of eigenvalues under certain conditions.

Theorem 1.5 (Pole-Shifting Theorem). *The pair (A, B) is controllable, if and only if, for each set of complex numbers $\lambda_1, \ldots, \lambda_n$, there exists a matrix $F \in \mathbb{R}^{m \times n}$, such that the eigenvalues of $A + BF$ are exactly the numbers $\lambda_1, \ldots, \lambda_n$.*

Hence, for a controllable system it is possible to achieve any asymptotic behaviour with a linear constant feedback control. This method is also called pole-placement.

1.1.3 Linear Quadratic Control

This section is concerned with linear constant feedback controls under other constraints than pole-placement. We want to find a control that results in a asymptotically stable system and in addition meets a minimality condition.

Let $\Sigma = (A, B)$ be a system without output. Then we associate with it the quadratic cost function

$$J_\infty(x_0, u) := \int_0^\infty (u(t)^* Ru(t) + x(t)^* Qx(t))dt,$$

where $Q \in \mathbb{R}^{n\times n}$ is a positive semidefinite and $R \in \mathbb{R}^{m\times m}$ is a positive definite matrix. The cost function is evaluated for an initial value x_0 and a control $u(t)$. It may be infinite.

The general problem of the linear quadratic control is to find a linear constant feedback $u(t) = Fx(t)$, such that $J_\infty(x_0, u)$ is minimized. The solution to this problem is given by the following theorem, the proof of which can be found in [4], Theorem 31 :

Theorem 1.6. *Let the system*

$$\dot{x}(t) = Ax(t) + Bu(t),$$

and the quadratic cost function be defined by Q and R as above. Furthermore let (A, B) be controllable and (A, Q) observable. Then for each $x_0 \in \mathbb{R}$ there exists a unique optimal control, given by

$$u(t) = -R^{-1}B^*X, \tag{1.3}$$

where X is the stabilizing solution of the algebraic Riccati equation

$$XBR^{-1}B^*X - XA - A^*X - Q = 0. \tag{1.4}$$

A solution X of the Riccati equation (1.4) is called *stabilizing*, if the resulting closed loop system given by the control (1.3) is stable. That is, if the so-called closed loop system matrix $A - BR^{-1}B^*X$ has only eigenvalues in the open left half plane. Under the assumptions of the theorem this solution is unique.

To a given system a certain pair of matrices Q,R assigns a unique control function that yields a stable system matrix of the resulting closed loop system. However, we do not have information about the quality of the stability of this system. That is, the eigenvalues of the system can be close to the imaginary line leading to poor decay of solutions of the closed loop system, or have large imaginary part leading to large oscillations in the solutions.

The question we want to consider is, how can we influence the stability properties of the closed loop system, by changing or choosing appropriate matrices Q and R.

We will need some tools to tackle this problem. In the following sections we will introduce these tools.

1.1.4 The Hamiltonian

To solve the optimal feedback problem, one has to solve the algebraic Riccati equation (1.4). One way is by solving an eigenvalue problem. A complete development of these results can be found in [4].

Definition. A real matrix H of size $2n$ is called a *Hamiltonian matrix*, if and only if

$$H^*J + JH = 0, \qquad \text{where} \quad J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}.$$

Here I is the n-dimensional identity matrix.

The eigenvalues of a Hamiltonian matrix H are distributed symmetrically with respect to the imaginary axis.

Associated to the quadratic control problem as defined in Section 1.1.3, we introduce the Hamiltonian matrix

$$H = \begin{pmatrix} A & -BR^{-1}B^* \\ -Q & -A^* \end{pmatrix}.$$

It is well known, that under the assumptions that (A, B) is controllable and (Q, A) is observable, the subspace of $\mathbb{C}^{2n}$ spanned by eigenvectors and generalized eigenvectors of H corresponding to its eigenvalues in the open left half plane is of the form

$$M = \mathrm{Im} \begin{pmatrix} I \\ X \end{pmatrix},$$

where X is the stabilizing solution of the algebraic Riccati equation (1.4).

Furthermore, the restriction of H to M is similar to the closed loop system matrix $A - BR^{-1}B^*X$. Hence, the stable eigenvalues of the Hamiltonian H are exactly the eigenvalues of the closed loop system.

1.2 Realization Theory

Realization Theory is an important cornerstone of mathematical systems theory. However, we will only need a particular application of it related to matrix valued rational functions. The following definitions and results are taken from [6] for rational functions.

Definition. A function $r : \mathcal{D} \to K^{n\times m}$, where $\mathcal{D} \subset K$, is called *rational*, if the functions r_{ij} for the elements are rational functions. The field K can be $\mathbb{R}$ or $\mathbb{C}$. A rational matrix valued function is called *proper*, if $\lim_{|\lambda|\to\infty} r(\lambda)$ exists.

Definition. Let $r(\lambda)$ be a real rational function with values in $\mathbb{R}^{n\times m}$. A quadruple $\Phi = (A, B, C, D)$ of matrices, where $A \in \mathbb{R}^{q\times q}$, $B \in \mathbb{R}^{q\times m}$, $C \in \mathbb{R}^{n\times q}$ and $D \in \mathbb{R}^{n\times m}$, is called a *realization of* $r(\lambda)$, if

$$r(\lambda) = D + C(\lambda I - A)^{-1}B. \tag{1.5}$$

A realization is called *controllable* if (A, B) is controllable and *observable* if (A, C) is observable. The dimension of the square matrix A is called the state space dimension.

A realization of a fixed rational function is not unique. Two realizations $\Phi = (A, B, C, D)$ and $\Phi_1 = (A_1, B_1, C_1, D_1)$ are called similar, if $D = D_1$, the state space dimensions of Φ and Φ_1 are equal to n, and there exists an invertible transformation $S \in \mathbb{R}^{n\times n}$ such that the relations

$$A = SA_1S^{-1},\ B = SB_1,\ C = C_1S^{-1}$$

hold.

Furthermore Φ is called a *dilation* of Φ_1, or Φ_1 is called a *reduction* of Φ, if $D = D_1$ and for suitable matrices $A_2, A_3, A_4, A_5, A_6, B_2, C_2$, the following equalities hold

$$A = \begin{pmatrix} A_2 & A_3 & A_4 \\ 0 & A_1 & A_5 \\ 0 & 0 & A_6 \end{pmatrix}, B = \begin{pmatrix} B_2 \\ B_1 \\ 0 \end{pmatrix}, C = \begin{pmatrix} 0 & C_1 & C_2 \end{pmatrix}.$$

The following theorems give some results about realizations of rational functions and when we can expect them to be controllable and observable.

A system Φ and its reduction Φ_1 realize the same rational matrix valued function $r(\lambda)$, as the following calculation shows. Assume Φ to be given as above. Then it realizes the function

$$r(\lambda) = D + \begin{pmatrix} 0 & C_1 & C_2 \end{pmatrix} \left(\lambda I - \begin{pmatrix} A_2 & A_3 & A_4 \\ 0 & A_1 & A_5 \\ 0 & 0 & A_6 \end{pmatrix}\right)^{-1} \begin{pmatrix} B_2 \\ B_1 \\ 0 \end{pmatrix}.$$

Computing the inverse gives

$$
\begin{aligned}
r(\lambda) &= D + \begin{pmatrix} 0 & C_1 & C_2 \end{pmatrix} \begin{pmatrix} (\lambda I - A_2)^{-1} & \overline{A}_3 & \overline{A}_4 \\ 0 & (\lambda I - A_1)^{-1} & \overline{A}_5 \\ 0 & 0 & (\lambda I - A_6)^{-1} \end{pmatrix} \begin{pmatrix} B_2 \\ B_1 \\ 0 \end{pmatrix} \\
&= D + C_1 (\lambda I - A_1)^{-1} B_1,
\end{aligned}
$$

for some matrices $\overline{A}_3$, $\overline{A}_4$ and $\overline{A}_5$. The last line is the rational matrix valued function realized by the reduction Φ_1.

Theorem 1.7. *A proper rational matrix valued function $r(\lambda)$ allows a realization. Furthermore, if (A, B, C, D) is a realization of $r(\lambda)$, then*

$$\lim_{|\lambda| \to \infty} r(\lambda) = D.$$

Theorem 1.8. *Two controllable and observable realizations of the same rational matrix valued function are similar and the corresponding transformation is unique.*

A realization Φ is called *minimal*, if all realizations of the same rational matrix valued function have state space dimension greater than or equal to the state space dimension of Φ.

Theorem 1.9. *A realization is minimal, if and only if it is controllable and observable.*

1.2.1 A Realization using the Hamiltonian

Let be given a linear quadratic control problem with the system $\Sigma = (A, B)$ and the cost function matrices R and Q. Furthermore we suppose that the assumptions of Theorem 1.6 hold. Then we introduce the rational matrix valued function

$$
\begin{aligned}
Z(\lambda) &= R + \begin{pmatrix} 0 & B^* \end{pmatrix} \left(\lambda I - \begin{pmatrix} A & 0 \\ -Q & -A^* \end{pmatrix} \right)^{-1} \begin{pmatrix} B \\ 0 \end{pmatrix} \qquad (1.6) \\
&= 1 - B^*(\lambda + A^*)^{-1} Q (\lambda - A)^{-1} B.
\end{aligned}
$$

and consider some properties of it. A short calculation establishes its inverse to be the following function

$$Z(\lambda)^{-1} = R^{-1} - R^{-1} \begin{pmatrix} 0 & B^* \end{pmatrix} \left(\lambda I - \begin{pmatrix} A & -BR^{-1}B^* \\ -Q & -A^* \end{pmatrix} \right)^{-1} \begin{pmatrix} B \\ 0 \end{pmatrix} R^{-1}.$$

If the realization (1.6) is minimal, then also the realization of the inverse function $Z(\lambda)^{-1}$ is minimal. The poles of this inverse function are exactly the eigenvalues of the Hamiltonian H which is associated with the given problem. These poles are by definition the zeros of the function $Z(\lambda)$ given by (1.6).

Under the assumption that $\sigma(H) \cap \sigma(A) = \emptyset$ the realization in (1.6) is minimal. Indeed, assume that $(H, (0, B^*))$ is not observable, which is equivalent to $(H^*, (0, B^*)^*)$ not controllable. Then, by Theorem 1.4,

$$\operatorname{rank}\left(H^* - \lambda_0 I, \begin{pmatrix} 0 \\ B \end{pmatrix}\right) = n$$

for some eigenvalue λ_0 of H^*. Observe that then $\overline{\lambda_0}$ is an eigenvalue of H. Hence, there exists a row vector (x^*, y^*) such that

$$(x^*, y^*)\left(H^* - \lambda_0 I, \begin{pmatrix} 0 \\ B \end{pmatrix}\right) = 0.$$

This leads to $y^*B = 0$ and furthermore to $x^*A^* = x^*\lambda_0$. Thus $\overline{\lambda_0}$ is an eigenvalue of A. Since $\overline{\lambda_0}$ is also an eigenvalue of H by assumption, this leads to a contradiction to $\sigma(H) \cap \sigma(A) = \emptyset$. Thus $(H, (0\ B^*))$ is observable. But then the realization is also controllable, since

$$\operatorname{rank}\left(H - \lambda I, \begin{pmatrix} B \\ 0 \end{pmatrix}\right) = n.$$

2 Perturbations of the Cost Function

2.1 General Assumptions

From now on we will always consider a linear quadratic control problem, where the system is given in control canonical form and the cost function matrix R is the identity matrix. We can assume this without loss of generality because of the following calculation.

Assume a system $\Sigma = (A, B)$ with state space dimension n and input dimension m and the cost function matrices $R > 0$ and $Q \geq 0$ is given. Then there exists a positive matrix $P > 0$ such that $R = P^2$. Then P is invertible and we can transform the input space with P^{-1}. Set $A_1 = A$, $B_1 = BP^{-1}$, $R_1 = P^{-1}PPP^{-1} = I$ and $Q_1 = Q$. Now we have a new linear quadratic control problem.

The next step is to transform the pair (A, B) to control canonical form. To do so we need a state space similarity. From Remark 1.2 we know that there exists an invertible transformation T such that $A_2 = T^{-1}A_1T$ and $B_2 = T^{-1}B_1$ are in control canonical form. Set $Q_2 = T^*QT$ and $R_2 = R_1$. Then we get a new system $\Sigma_2 = (A_2, B_2)$ in control canonical form, and with the cost function matrices $R_2 = I$ and $Q_2 \geq 0$.

Now we assume that (A, B) is controllable and (A, Q) is observable. Then also (A_2, B_2) and (A_2, Q_2) are controllable and observable respectively. So we can apply Theorem 1.6 and find a stabilizing solution X_2 to the Riccati equation

$$X_2B_2R_2^{-1}B_2^*X_2 - X_2A_2 - A_2^*X_2 - Q_2 = 0.$$

Substituting the transformations T and P into this equation yields the following Riccati equation

$$\begin{aligned} X_2T^{-1}BP^{-1}(P^{-1}RP^{-1})^{-1}(T^{-1}BP^{-1})^*X_2 & \\ -X_2T^{-1}AT - (T^{-1}AT)^*X_2 - TQT &= 0. \end{aligned}$$

By expanding the parenthesis and multiplying from the left with $(T^{-1})^*$ and from the right with T^{-1} we get the following equation :

$$\begin{aligned} (T^{-1})^*X_2T^{-1}BR^{-1}B^*(T^{-1})^*X_2T^{-1} - (T^{-1})^*X_2T^{-1}A & \\ -A^*(T^{-1})^*X_2T^{-1} - Q &= 0. \end{aligned}$$

From this equation we see that $X = (T^{-1})^*X_2T^{-1}$ is a solution to the Riccati equation associated with the original linear quadratic control problem.

Because X_2 is the stabilizing solution, $A_2 - B_2 R_2^{-1} B_2^* X_2$ is stable. Substituting the transformations T and P into this term and multiplying from the left with T and from the right with T^{-1} yields

$$A - BR^{-1}B^*(T^{-1})^* X_2 T^{-1} = A - BR^{-1}B^* X.$$

Since the last matrix was derived by a similarity transformation from $A_2 - B_2 R_2^{-1} B_2^* X_2$, it has the same eigenvalues as the latter one. Hence X is the unique stabilizing solution to the Riccati equation of the original problem.

This shows that it is enough to investigate solutions of linear quadratic control problems, where the system is given in control canonical form and the cost function matrix R is the identity matrix.

Furthermore we will also assume that the rank of $B \in \mathbb{R}^{n\times m}$ is equal to m and hence $m \leq n$. To see this, assume (A, B) is given in control canonical form. If the rank of B is less then m the system has the form

$$\dot{x} = Ax + Bu = Ax + \begin{pmatrix} B_1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = Ax + B_1 u_1.$$

So we may assume from the beginning that $B = B_1$.

2.2 Perturbations of the Cost Function

In this subsection we study the asymptotic behaviour of the solution of the linear quadratic problem, abbreviated with LQ problem, for a certain perturbation of the cost function Q. This will be done for one dimensional input systems.

We assume a controllable system $\Sigma = (A, B)$ where $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times 1}$. The matrix Q of the cost function is a positive semi-definite matrix and R is a positive real number, which we scale so that $R = 1$. Furthermore, by Section 2.1 we can assume without loss of generality that the matrices are given in the control canonical form

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & 0 \\ 0 & & & 0 & 1 \\ -a_1 & -a_2 & \dots & -a_{n-1} & -a_n \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

We perturb Q with a positive semi-definite matrix Q_1 as

$$Q(\varepsilon) = Q + \varepsilon Q_1 \ , \ \varepsilon \geq 0.$$

The perturbation matrix Q_1 is chosen such that $Q_1 B A^j = 0$ for $j = 0, \ldots, n-2$. A short calculation yields that the only non-zero entry in Q_1 is the entry in the left upper corner, which we choose to be 1, so that

$$Q_1 = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \\ 0 & \ldots & & 0 \end{pmatrix}.$$

By Theorem 1.6, there exists an unique solution X to the Riccati equation, such that $A - BB^*X$ is stable and the cost function is minimized for this feedback control. We are interested in the behaviour of the closed loop eigenvalues for $\varepsilon \to \infty$. The closed loop eigenvalues are precisely the stable eigenvalues of the Hamiltonian associated with the LQ problem. Hence we investigate the behaviour of the eigenvalues of the perturbed Hamiltonian H_ε. We will use the characteristic polynomial to do that.

2.2.1 The Characteristic Polynomial of the Hamiltonian

The Hamiltonian matrix H_ε associated with the perturbed problem takes the following form

$$H_\varepsilon = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 & & \ldots & & 0 \\ \vdots & \ddots & \ddots & & \vdots & & & & & \vdots \\ & & \ddots & \ddots & 0 & \vdots & & \ddots & & \\ 0 & & & 0 & 1 & & & & 0 & 0 \\ -a_1 & -a_2 & \ldots & -a_{n-1} & -a_n & 0 & & \ldots & 0 & -1 \\ -q_{1,1} - \varepsilon & -q_{1,2} & \ldots & -q_{1,n-1} & -q_{1,n} & 0 & & \ldots & 0 & a_1 \\ & & & & & -1 & \ddots & & 0 & a_2 \\ \vdots & & \ddots & & \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\ & & & & & \vdots & & \ddots & 0 & a_{n-1} \\ -q_{1,n} & & \ldots & & -q_{n,n} & 0 & \ldots & 0 & -1 & a_n \end{pmatrix}.$$

The Hamiltonian of the original problem will be denoted by H. Furthermore, we denote the characteristic polynomial of a matrix A by $p_A(\lambda)$, where

$$p_A(\lambda) = \det(A - \lambda I) = |A - \lambda I|.$$

We can split up the determinant into a sum of two by splitting the first column into the sum of two vectors. Thus the characteristic polynomial of the perturbed Hamiltonian H_ε is the sum of the determinants of these matrices:

$$p_{H_\varepsilon}(\lambda) = |H - \lambda I| + \begin{vmatrix}
0 & 1 & 0 & \dots & 0 & 0 & & \dots & & 0 \\
\vdots & -\lambda & \ddots & & \vdots & & & & & \vdots \\
 & & \ddots & \ddots & 0 & \vdots & & \ddots & & \\
0 & & & -\lambda & 1 & & & & 0 & 0 \\
0 & -a_2 & \dots & -a_{n-1} & -a_n - \lambda & 0 & & \dots & 0 & -1 \\
-\varepsilon & -q_{1,2} & \dots & -q_{1,n-1} & -q_{1,n} & -\lambda & & \dots & 0 & a_1 \\
0 & & & & & -1 & -\lambda & & 0 & a_2 \\
\vdots & & \ddots & & \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\
 & & & & & \vdots & & \ddots & -\lambda & a_{n-1} \\
0 & -q_{n,2} & \dots & & -q_{n,n} & 0 & \dots & 0 & -1 & a_n - \lambda
\end{vmatrix}.$$

Next we develop the second determinant with respect to the first column. Then it becomes

$$(-1)^{n+1}\varepsilon \begin{vmatrix}
1 & 0 & \dots & 0 & 0 & & \dots & & 0 \\
-\lambda & \ddots & & \vdots & & & & & \vdots \\
0 & \ddots & \ddots & 0 & \vdots & & \ddots & & \\
\vdots & & -\lambda & 1 & & & & 0 & 0 \\
-a_2 & \dots & -a_{n-1} & -a_n - \lambda & 0 & & \dots & 0 & -1 \\
-q_{2,2} & & \dots & -q_{2,n} & -1 & -\lambda & & 0 & a_2 \\
\vdots & \ddots & & \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\
 & & & & \vdots & & -1 & -\lambda & a_{n-1} \\
-q_{n,2} & & \dots & -q_{n,n} & 0 & \dots & 0 & -1 & a_n - \lambda
\end{vmatrix}.$$

We develop the determinant further with respect to the first row, which in turn yields a matrix where the first row has only one entry. By repeating this calculation we can eliminate the left half of the matrix and arrive at the following expression

$$(-1)^{n+1}\varepsilon \begin{vmatrix} 0 & \dots & & 0 & -1 \\ -1 & -\lambda & & 0 & a_2 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & & -1 & -\lambda & a_{n-1} \\ 0 & \dots & 0 & -1 & a_n - \lambda \end{vmatrix}.$$

Now, following the same procedure with the columns we can reduce the matrix to the scalar -1 in the upper right corner without changing the sign of the determinant. Thus this term becomes $(-1)^n\varepsilon$ and we can write the characteristic polynomial of H_ε as

$$p_{H_\varepsilon}(\lambda) = p_H(\lambda) + (-1)^n\varepsilon.$$

The constant term of the characteristic polynomial of H is the product of the eigenvalues of H which are distributed symmetrically with respect to the imaginary axis. Denote the eigenvalues of H by $\lambda_k,\ k = 1, \dots, 2n$. Thus the constant term p_0 becomes

$$p_0 = \prod_{k=1}^{2n} \lambda_k = \prod_{j=1}^{n} (\lambda_j)(-\overline{\lambda}_j) = (-1)^n \prod_{j=1}^{n} |\lambda_j|^2.$$

and has the same sign as the perturbation ε. Thus we can write the characteristic polynomial of the perturbed matrix as

$$p_{H_\varepsilon}(\lambda) = \lambda^{2n} + p_{n-1}\lambda^{2(n-1)} + \dots + (-1)^n(p+\varepsilon) \tag{2.1}$$

for some positive $p = \prod_{j=1}^{n} |\lambda_j|^2$ and constants $p_i,\ i = 1, \dots, n$. The fact that only powers of λ^2 appear in the polynomial follows from the symmetric distribution of the eigenvalues of H. Note that for $\varepsilon \to \infty$ the constant term is always nonzero, thus for all positive ε no eigenvalue is equal to zero.

2.2.2 Zeros of the Perturbed Polynomials

The following lemma describes the effect of the perturbation ε on the zeros of the perturbed polynomial.

Lemma 2.1. *For every real valued polynomial* $p(\lambda) = \lambda^{2n} + p_{n-1}\lambda^{2(n-1)} + \ldots + (-1)^n p_0$, $p_0 > 0$ *and every real* δ *with* $0 < \delta \leq 1$ *the following holds. There exists a constant* N *such that for all* $p_0 > N$ *there is a one-to-one correspondence between the zeros* λ_k *of* $p(\lambda)$ *and the zeros* ζ_k *of* $q(\lambda) = \lambda^{2n} + (-1)^n p_0$ *such that for corresponding zeros the relation* $|\lambda_k - \zeta_k| < \delta R$ *holds, where* $R = p_0^{\frac{1}{2n}} \sin\left(\frac{\pi}{2n}\right)$.

Proof. We apply Rouche's Theorem to show that in a certain disk around a zero ζ_k of $q(\lambda)$ there is exactly one zero of $p(\lambda)$.

The zeros of $q(\lambda)$ are $\zeta_k = p_0^{\frac{1}{2n}} e^{\frac{(2k+n+1)i\pi}{2n}}$. Fix R such that

$$R = \frac{1}{2} \cdot \min_{i \neq j} |\zeta_i - \zeta_j|.$$

The zeros ζ_k are evenly distributed on a circle with center 0 and the minimum is equal to the distance of two consecutive zeros.

$$\min_{i \neq j} |\zeta_i - \zeta_j| = |p_0|^{\frac{1}{2n}} |e^{\frac{(2k+n+1)i\pi}{2n}}||1 - e^{\frac{i\pi}{n}}| = 2p_0^{\frac{1}{2n}} \sin\left(\frac{\pi}{2n}\right)$$

So $R = p_0^{\frac{1}{2n}} \sin\left(\frac{\pi}{2n}\right)$.

Fix a zero ζ_k of $q(\lambda)$. We compare $|p(\lambda) - q(\lambda)|$ and $|q(\lambda)|$ on a circle with center ζ_k and radius δR. For λ on this circle, $|\zeta_k| - \delta R \leq |\lambda| \leq |\zeta_k| + \delta R$ holds. Moreover the distance to the fixed zero ζ_k is δR and to any other zero it is less or equal to R. Hence, we can estimate $|q(\lambda)|$ as follows

$$|q(\lambda)| = \prod_{i=1}^{2n} |\lambda - \zeta_i| \geq \delta (R)^{2n} = p_0 \delta \left(\sin\left(\frac{\pi}{2n}\right)\right)^{2n}.$$

First we take p_0 large enough so that on the circle with center ζ_k and radius δR there are no values λ with $|\lambda| < 1$. This condition yields a lower bound for p_0 since

$$|\lambda| \geq |\zeta_k| - \delta R = p_0^{\frac{1}{2n}} - \delta p_0^{\frac{1}{2n}} \sin\left(\frac{\pi}{2n}\right) = p_0^{\frac{1}{2n}} \left(1 - \delta \sin\left(\frac{\pi}{2n}\right)\right).$$

The last term is greater than or equal to one, if

$$p_0 \geq \frac{1}{\left(1 - \delta \sin\left(\frac{\pi}{2n}\right)\right)^{2n}} =: N_1$$

holds. We can estimate the modulus of the difference for $|\lambda| \geq 1$ with the following chain of inequalities.

$$
\begin{aligned}
|p(\lambda) - q(\lambda)| = |p_{n-1}\lambda^{2(n-1)} + \ldots + p_1\lambda^2| &\leq |\lambda^{2(n-1)}| \sum_{i=1}^{n-1} |p_i| \\
&\leq (|\zeta_k| + \delta R)^{2(n-1)} \sum_{i=1}^{n-1} |p_i| = \left(1 + \delta \sin\left(\frac{\pi}{2n}\right)\right)^{\frac{n-1}{n}} p_0^{\frac{n-1}{n}} \sum_{i=1}^{n-1} |p_i|
\end{aligned}
$$

Hence $|p(\lambda) - q(\lambda)| < |q(\lambda)|$ holds, if

$$
\left(1 + \delta \sin\left(\frac{\pi}{2n}\right)\right)^{\frac{n-1}{n}} p_0^{\frac{n-1}{n}} \sum_{i=1}^{n-1} |p_i| < p_0 \delta \left(\sin\left(\frac{\pi}{2n}\right)\right)^{2n}.
$$

This is true for

$$
p_0 > \frac{\left(1 + \delta \sin\left(\frac{\pi}{2n}\right)\right)^{n-1} \left(\sum_{i=1}^{n-1} |p_i|\right)^n}{\delta^n \left(\sin\left(\frac{\pi}{2n}\right)\right)^{3n}} =: N_2. \tag{2.2}
$$

The estimation of the modulus of the difference of the two polynomials requires that $|\lambda| \geq 1$, which is equal to $p_0 \geq N_1$. Set

$$
p_0 > N := \max\{N_1, N_2\}.
$$

Then the last inequality is true for all λ on the circle with center ζ_k and radius δR.

Now we apply Rouche's Theorem and can conclude that $p(\lambda)$ has the same number of zeros as $q(\lambda)$ in the circle with center ζ_k and radius R. R was chosen such that $q(\lambda)$ has ζ_k as its only zero in this circle and therefore $p(\lambda)$ has one zero there too. The zero ζ_k of $q(\lambda)$ was chosen arbitrarily, therefore for $p_0 > N$ we have a bijective mapping between the zeros λ_k of $p(\lambda)$ and ζ_k of $q(\lambda)$ such that $|\lambda_k - \zeta_k| < \delta R$ holds. $\square$

Note that the first constant N_1 tends to 1 for $\delta \to 0$, whereas the constant N_2 grows rather fast. So for small δ or large p_0 the second bound will be important.

2.2.3 Asymptotics of the Eigenvalues of the Closed Loop System

We want to conclude with a Theorem giving more general information about the behaviour of the closed loop eigenvalues.

Theorem 2.2. *Let $\Sigma = (A, B)$ be a controllable system with state space dimension n and one dimensional input. Let Q be a positive semi-definite matrix, let (A, Q) be observable and $R = 1$. Furthermore we assume without loss of generality that Σ is given in control canonical form. Perturb the matrix Q with a matrix $Q_1(\varepsilon)$ that has the form*

$$Q_1(\varepsilon) = \begin{pmatrix} \varepsilon & 0 & \dots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & 0 \end{pmatrix}.$$

Let $\lambda_k,\ k = 0, \dots, n-1$, be the eigenvalues of the closed loop system given by the solution to the LQ problem defined by Σ, R and $Q + Q_1(\varepsilon)$. Then for all real values δ with $0 < \delta \le 1$ and for ε larger than some constant N, the following statements hold.

(i) *Denote by $\zeta_k,\ k = 0, \dots, n-1$ the zeros of the polynomial $\lambda^{2n} + (-1)^n(p + \varepsilon)$ in the open left half plane. Then*

$$\lim_{\varepsilon \to \infty} |\zeta_k - \lambda_k| = 0$$

for an appropriate numbering of the eigenvalues of the closed loop system.

(ii) *Define α by $\sin(\alpha) = \delta \sin(\frac{\pi}{2n})$ and $0 \ge \alpha \ge \frac{\pi}{2}$. Furthermore, let $\alpha_1 = \alpha + \pi \frac{n-1}{2n}$. Then all λ_k are in the sector*

$$\{z \in \mathbb{C} \mid |\arg z| \le \alpha_1\} \cup \{z \mid |\arg z - \pi| \le \alpha_1\}. \tag{2.3}$$

(iii) *The real parts of the eigenvalues of the closed loop system are bounded by*

$$-C\left(1 + \delta \sin\left(\frac{\pi}{2n}\right)\right) \le \Re(\lambda_k) \le -C \sin\left(\frac{\pi}{2n}\right)(1 - \delta), \tag{2.4}$$

where $C = (p + \varepsilon)^{\frac{1}{2n}}$.

Proof. Let H_ε denote the Hamiltonian of the perturbed LQ problem. By Theorem 1.6 and Section 2.2 we have to study the eigenvalues of H_ε. Section 2.2.1 yields that (2.1) is the characteristic polynomial $p_{H_\varepsilon}(\lambda)$ of H_ε.

Now we can apply Lemma 2.1 and can compare the zeros of the characteristic polynomial (2.1) of H_ε with the polynomial $q(\lambda) = \lambda^{2n} + (-1)^n(p + \varepsilon)$. We denote the zeros of $p_{H_\varepsilon}(\lambda)$ by λ_k and the zeros of $q(\lambda)$ by ζ_k. By Lemma 2.1, the inequality

$$|\lambda_k - \zeta_k| < \delta\,(p+\varepsilon)^{\frac{1}{2n}} \sin\left(\frac{\pi}{2n}\right) = \delta R \tag{2.5}$$

holds, for an appropriate numbering of the zeros and if $p+\varepsilon$ is greater some constant N. The number p equals the modulus of the determinant of H, the unperturbed Hamiltonian.

(i) Set

$$\delta = \frac{\alpha}{(p+\varepsilon)^{\frac{1}{2n}} \sin(\frac{\pi}{2n})},$$

where $1 > \alpha > 0$. Then the lower bound N_2 of equation (2.2) for $(p+\varepsilon)$ becomes

$$N_2 = \frac{(p+\varepsilon)^{\frac{1}{n}} \left(1 + \frac{\alpha}{(p+\varepsilon)^{\frac{1}{2n}}}\right)^{n-1} \left(\sum_{i=1}^{n-1} |p_i|\right)^n}{\alpha^n \left(\sin\left(\frac{\pi}{2n}\right)\right)^{2n}} \leq \frac{(p+\varepsilon)^{\frac{1}{n}} 2^{n-1} \left(\sum_{i=1}^{n-1} |p_i|\right)^n}{\alpha^n \left(\sin\left(\frac{\pi}{2n}\right)\right)^{2n}}.$$

The last estimate holds for $p+\varepsilon > 1$. Comparing the last term with $p+\varepsilon$, yields a new lower bound $\overline{N}_2$ for $p+\varepsilon$:

$$\frac{(p+\varepsilon)^{\frac{1}{n}} 2^{n-1} \left(\sum_{i=1}^{n-1} |p_i|\right)^n}{\alpha^n \left(\sin\left(\frac{\pi}{2n}\right)\right)^{2n}} \leq (p+\varepsilon)$$

and hence

$$\overline{N}_2 := \frac{2^{n-1} \left(\sum_{i=1}^{n-1} |p_i|\right)^n}{\alpha^n \left(\sin\left(\frac{\pi}{2n}\right)\right)^{2n}} \leq (p+\varepsilon)^{\frac{n-1}{n}}.$$

Using inequality (2.5) the distance between the two zeros is

$$|\lambda_k - \zeta_k| < \alpha,$$

if $p+\varepsilon$ is greater or equal to $\overline{N}_2^{\frac{n}{n-1}}$. Hence statement *(i)* holds.

(ii) Fix a zero $\zeta_k = (p+\varepsilon)^{\frac{1}{2n}}\, e^{\frac{(2k+n+1)i\pi}{2n}}$ of $q(\lambda)$ in the left half plane. Define α by $\sin\alpha = \delta \sin\left(\frac{\pi}{2n}\right)$ and $0 \leq \alpha \leq \frac{\pi}{2}$. From (2.5) it follows that the zero λ_k is in the sector

$$\left\{ z \in \mathbb{C} \mid \alpha \geq \left| \arg z - \frac{(2k+n+1)\,\pi}{2n} \right| \right\}.$$

This can easily be deduced from some simple geometric considerations. Then this sector is a subset of the sector (2.3). Thus all stable eigenvalues of H_ε are in the sector (2.3).

(iii) The real part of a zero ζ_k of $q(\lambda)$ is

$$\Re(\zeta_k) = (p+\varepsilon)^{\frac{1}{2n}} \cos\left(\frac{(2k+n+1)\,\pi}{2n}\right) = (p+\varepsilon)^{\frac{1}{2n}} \sin\left(\frac{(2k+1)\,\pi}{2n}\right)$$

From (2.5) we get the following bounds for the real parts of the closed loop eigenvalues λ_k, because they lie inside of circles around the numbers ζ_k :

$$\Re(\zeta_k) - \delta R \leq \Re(\lambda_k) \leq \Re(\zeta_k) + \delta.$$

Substituting for δR and ζ_k the left inequality becomes

$$(p+\varepsilon)^{\frac{1}{2n}} \left(\left|\sin\left(\frac{(2k+1)\,\pi}{2n}\right)\right| - \delta \sin\left(\frac{\pi}{2n}\right)\right) \leq \Re(\lambda_k)$$

and the right inequality becomes

$$\Re(\lambda_k) \leq (p+\varepsilon)^{\frac{1}{2n}} \left(\left|\sin\left(\frac{(2k+1)\,\pi}{2n}\right)\right| + \delta \sin\left(\frac{\pi}{2n}\right)\right).$$

To find simultaneous bounds for all stable eigenvalues, we maximize the upper bound and minimize the lower bound by looking at the zeros ζ_k.

The lower bound is minimized by zeros closest to the real axis. For n even these zeros are $\zeta_{\frac{n-2}{2}}$ and $\zeta_{\frac{n}{2}}$, and the bound becomes

$$-(p+\varepsilon)^{\frac{1}{2n}} \left(\cos\left(\frac{\pi}{2n}\right) + \delta \sin\left(\frac{\pi}{2n}\right)\right).$$

For n odd there is one zero $\zeta_{\frac{n-1}{2}}$ in the open left half plane that lies on the real axis. The bound computes to

$$-(p+\varepsilon)^{\frac{1}{2n}} \left(1 + \delta \sin\left(\frac{\pi}{2n}\right)\right).$$

This value is smaller than the value for n even and thus yields the simultaneous bound.

The upper bound is maximized for zeros ζ_k closest to the imaginary axis. These are ζ_0 and ζ_{n-1} and the upper bound has the value

$$-(p+\varepsilon)^{\frac{1}{2n}} \sin\left(\frac{\pi}{2n}\right)(1-\delta).$$

These are the bounds of the strip (2.4). □

3 Pole Placement for One Dimensional Input

In this section we prove a theorem for pole placement in the case of a system with a one dimensional input. This theorem will yield a cost function matrix Q which solves the pole placement problem when possible. Furthermore we characterize all possible matrices Q that solve the pole placement problem for a fixed set of eigenvalues, if such a Q exists.

3.1 Main Theorem

Throughout this chapter we will assume that $\Sigma = (A, B)$ is a controllable system with one dimensional input and state space dimension n. The system is given in control canonical form. The numbers a_i, $i = 1, \dots, n$, denote the elements of the last row of the system matrix A. The matrix R of the cost function is $R = (1)$. This can be assumed without loss of generality, see Section 2.1. Then the following theorem gives conditions for pole placement by choosing a cost function matrix Q and describes an algorithm to find Q.

Theorem 3.1. *Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be n complex numbers in the open left half plane such that neither λ_i nor $-\overline{\lambda_i}$ are eigenvalues of the system matrix A and that $\prod_{i=1}^{n}(\lambda - \lambda_i)$ is a real polynomial. Then there exists a positive semidefinite matrix Q such that the closed loop eigenvalues of the LQ-optimal system with respect to the matrices $R = (1)$ and Q are $\lambda_1, \lambda_2, \dots, \lambda_n$, if and only if*

$$r(\lambda) = (-1)^n \prod_{i=1}^{n}(\lambda - \lambda_i)(\lambda + \overline{\lambda_i}) - p_A(\lambda)\overline{p_A(-\overline{\lambda})} \tag{3.1}$$

is nonnegative on $i\mathbb{R}$.

Proof. Recall that the real rational function $Z(\lambda)$ is given by

$$Z(\lambda) = 1 - B^*(\lambda + A^*)^{-1}Q(\lambda - A)^{-1}B.$$

Using Cramers rule we rewrite this as

$$Z(\lambda) = \frac{p_A(\lambda)\overline{p_A(-\overline{\lambda})} + B^*(-\lambda - A^*)^{\dagger}Q(\lambda - A)^{\dagger}B}{p_A(\lambda)\overline{p_A(-\overline{\lambda})}},$$

where $Y^{\dagger}$ denotes the matrix such that $YY^{\dagger} = det(Y)I$ holds.

Now assume that there exists a positive semidefinite matrix Q such that the eigenvalues of the closed loop system are $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then the Hamiltonian H has the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n, -\overline{\lambda}_1, -\overline{\lambda}_2, \ldots, -\overline{\lambda}_n$. By assumption $\sigma(H) \cap \sigma(A) = \emptyset$ holds and the eigenvalues are precisely the zeros of the function $Z(\lambda)$.

The leading coefficient of the polynomial in the numerator of $Z(\lambda)$ is $(-1)^n$. Thus the numerator is of the form $(-1)^n \prod_{i=1}^n (\lambda - \lambda_i)(\lambda + \overline{\lambda_i})$. Then it follows from the definition of $Z(\lambda)$ that

$$r(\lambda) = B^*(-\lambda - A^*)^\dagger Q(\lambda - A)^\dagger B$$

holds. We choose $\lambda = i\omega$, where ω is a real number, and obtain

$$r(i\omega) = B^*(-i\omega - A^*)^\dagger Q(i\omega - A)^\dagger B = ((i\omega - A)^\dagger B)^* Q(i\omega - A)^\dagger B.$$

Since Q is positive semidefinite, the last term is also positive semidefinite. Hence $r(\lambda)$ is nonnegative on $i\mathbb{R}$.

Conversely assume that $r(\lambda)$ is nonnegative on $i\mathbb{R}$. Because its values are real for all imaginary numbers, its zeros are symmetric with respect to the imaginary axis. Furthermore it is nonnegative and therefore allows the decomposition

$$r(\lambda) = v(\lambda)\overline{v(-\overline{\lambda})}.$$

where $v(\lambda)$ is a real polynomial. This is true, since $r(\lambda)$ is a real polynomial and thus its zeros also are symmetric with respect to the real axis. Furthermore, $r(\lambda)$ and hence also $v(\lambda)$ have no common zeros with $p_A(\lambda)$. Therefore no pole zero cancellation occurs in the real rational function $\frac{v(\lambda)}{p_A(\lambda)}$, and it admits a minimal realization

$$\frac{v(\lambda)}{p_A(\lambda)} = \widetilde{S}(\lambda - \widetilde{A})^{-1}\widetilde{B},$$

where $\widetilde{A}$ is similar to A. Since this is a minimal realization, $\widetilde{A}$ and $\widetilde{B}$ are controllable. Thus we can transform this realization to control canonical form. This yields a new realization in terms of the matrices A and B, because we assumed that the system is given in control canonical form. So there exists a regular matrix T such that

$$\widetilde{A} = TAT^{-1}, \quad \widetilde{B} = TB, \quad \widetilde{S} = ST^{-1}$$

hold.

The matrix T can be computed in the following way : Let $T_1, T_2, \ldots, T_n$ be the column vectors of T; a similar notation is used for the other matrices. The second equation shows that $T_n = \widetilde{B}$ holds. Furthermore the first equation can be written as $\widetilde{A}T = TA$. From these two equations it follows that $T_{n-1} = (\widetilde{A} - Ia_n)\widetilde{B} = \widetilde{A}T_n - a_n\widetilde{B}$. Looking consecutively at the columns of T, we find the recursion $T_{n-k-1} = \widetilde{A}T_{n-k} - a_{n-k}\widetilde{B}$ for the column vectors of T. The transformation matrix is regular because $\widetilde{A}$ and $\widetilde{B}$ are controllable.

Transforming the state space of the realization with T yields a new minimal realization in terms of A, B and S. Furthermore, if we compute the Hermitian transpose of the rational function $\frac{v(\lambda)}{p_A(\lambda)}$ and its realization we get

$$\overline{\frac{\overline{v(\lambda)}}{p_A(\lambda)}} = B^*(\overline{\lambda} - A^*)^{-1}S^*.$$

Replacing λ by $-\overline{\lambda}$ gives

$$\overline{\frac{\overline{v(-\overline{\lambda})}}{p_A(-\overline{\lambda})}} = B^*(-\lambda - A^*)^{-1}S^*.$$

Computing the product of these two realizations and adding 1 gives the real rational function associated with an LQ-problem for the system (A, B). This function has exactly the zeros $\lambda_1, \lambda_2, \ldots, \lambda_n, -\overline{\lambda}_1, -\overline{\lambda}_2, \ldots, -\overline{\lambda}_n$, since

$$\begin{aligned}
1 + B^*(-\lambda - A^*)^{-1}S^*S(\lambda - A)^{-1}B &= 1 + \frac{v(\lambda)\overline{v(-\overline{\lambda})}}{p_A(\lambda)\overline{p_A(-\overline{\lambda})}} \\
&= \frac{p_A(\lambda)\overline{p_A(-\overline{\lambda})} + v(\lambda)\overline{v(-\overline{\lambda})}}{p_A(\lambda)\overline{p_A(-\overline{\lambda})}} = \frac{p_A(\lambda)\overline{p_A(-\overline{\lambda})} + r(\lambda)}{p_A(\lambda)\overline{p_A(-\overline{\lambda})}} \\
&= \frac{(-1)^n \prod_{i=1}^n (\lambda - \lambda_i)(\lambda + \overline{\lambda_i})}{p_A(\lambda)\overline{p_A(-\overline{\lambda})}}
\end{aligned}$$

holds.

Setting $Q = S^*S$, we have found a positive semidefinite matrix such that the eigenvalues of the Hamiltonian of the LQ problem are exactly the zeros of the above real rational function. Furthermore (A, Q) is observable because (A, S) is observable which follows from the fact that A,B and S give a minimal realization. Thus the eigenvalues of the resulting closed loop system are precisely $\lambda_1, \lambda_2, \ldots, \lambda_n$. $\square$

This theorem shows the existence of a cost function matrix Q or rank 1, that solves the pole placement problem under certain conditions. However, the computed Q is not unique because it depends on a factorization of a certain rational function and, in the general case, this factorization is not unique. We are interested in a description of all possible matrices Q.

3.2 Characterizations of Solutions

We will start with a result about the relation between two different solutions to the pole placement problem and then derive from that a complete characterization.

Theorem 3.2.

(a) *Let Q_1, Q_2 be two different solutions to the pole placement problem as in Theorem* 3.1. *Then $Q = Q_1 - Q_2$ satisfies*

$$-B^*(\lambda + A^*)^{-1}Q(\lambda - A)^{-1}B \equiv 0. \tag{3.2}$$

(b) *Let Q_1 be a solution to the pole placement problem of Theorem* 3.1. *Let Q be a hermitian matrix of same size as Q_1 such that $Q_1 + Q$ is positive semidefinite and that Q meets the equation* (3.2). *Then $Q_2 = Q_1 + Q$ is also a solution to the pole placement problem.*

Proof. First let Q_1, Q_2 be two different solutions to the same pole placement problem. From Theorem 3.1 it follows that the Hamiltonians for both matrices have the same eigenvalues and, furthermore, that these eigenvalues are not eigenvalues of A or A^*. So we see that the two rational functions

$$\begin{aligned} Z_1(\lambda) &= 1 - B^*(\lambda + A^*)^{-1}Q_1(\lambda - A)^{-1}B, \\ Z_2(\lambda) &= 1 - B^*(\lambda + A^*)^{-1}Q_2(\lambda - A)^{-1}B \end{aligned}$$

have the same zeros and poles because these are exactly the eigenvalues of the Hamiltonians and of A or A^*, respectively. Furthermore the leading coefficients of both functions are equal. Hence these two rational functions are equal and satisfy the relation

$$Z_1(\lambda) - Z_2(\lambda) = -B^*(\lambda + A^*)^{-1}(Q_1 - Q_2)(\lambda - A)^{-1}B \equiv 0.$$

Set $Q = Q_1 - Q_2$ and the first statement follows.

Now let Q_1 be a solution of the pole placement problem. Furthermore, let Q be as required. Set $Q_2 = Q_1 + Q$ which is positive semidefinite. Then

$$
\begin{aligned}
&1 - B^*(\lambda + A^*)^{-1}Q_2(\lambda - A)^{-1}B = 1 - B^*(\lambda + A^*)^{-1}(Q_1 + Q)(\lambda - A)^{-1}B \\
&= 1 - B^*(\lambda + A^*)^{-1}Q_1(\lambda - A)^{-1}B \,-\, B^*(\lambda + A^*)^{-1}Q(\lambda - A)^{-1}B \\
&= 1 - B^*(\lambda + A^*)^{-1}Q_1(\lambda - A)^{-1}B
\end{aligned}
$$

and Q_2 is also a solution to the pole placement problem, because it yields the same rational function as Q_1. □

The next theorem gives a characterization of all matrices which meet the equation (3.2) in terms of their elements.

Theorem 3.3. *A matrix $Q \in \mathbb{R}^{n\times n}$ satisfies the equation*

$$-B^*(\lambda + A^*)^{-1}Q(\lambda - A)^{-1}B \equiv 0,$$

if and only if its elements $q_{i,j}$, $i, j = 1, \ldots, n$, satisfy the conditions :

$$
\begin{aligned}
\sum_{i=1}^{k}(-1)^{k-i}q_{k-i+1,i} &= 0 \text{ for } k = 1, \ldots, n, \\
\sum_{i=k-n+1}^{n}(-1)^{i+n}q_{i,k-i+1} &= 0 \text{ for } k = n+1, \ldots, 2n-1.
\end{aligned}
$$

Proof. To prove this theorem we will compute first the matrices $(\lambda - A)^{-1}B$ and $-B^*(\lambda + A^*)^{-1}$. Observe that $-B^*(\lambda + A^*)^{-1} = ((-\lambda - A)^{-1}B)^*$ for real λ. Let $V = (\lambda - A)^{-1}B$ and $\widetilde{V} = -B^*(\lambda + A^*)^{-1}$.

Since the matrices A and B are in control canonical form, the product $(\lambda - A)^{-1}B$ yields the last column vector of $(\lambda - A)^{-1}$. Then the following relation holds :

$$B = (\lambda - A)V.$$

This leads to a set of equations for the elements v_i of V :

$$
\begin{aligned}
\lambda v_1 &= v_2 \\
\lambda v_2 &= v_3 \\
&\vdots \\
\lambda v_{n-1} &= v_n \\
a_1v_1 + a_2v_2 + \ldots + a_nv_n + \lambda v_n &= 1.
\end{aligned}
$$

Recursively substituting v_1 into these equations, we find

$$v_1(a_1 + \lambda a_2 + \lambda^2 a_3 + \ldots + \lambda^{n-1} a_n + \lambda^n) = 1.$$

Thus the vector V takes the form

$$V = \frac{(-1)^n}{p_A(\lambda)} \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-1} \end{pmatrix}.$$

Further, the second matrix $-B^*(\lambda + A^*)^{-1}$, denoted by $\widetilde{V}$, becomes

$$\widetilde{V} = \frac{(-1)^n}{p_A(-\lambda)} \left(1,\ -\lambda,\ \lambda^2,\ \ldots\ ,\ (-1)^{n-1}\lambda^{n-1}\right).$$

The condition on Q is now

$$\frac{(-1)^{2n}}{p_A(-\lambda)p_A(\lambda)} \left(1,\ -\lambda,\ \lambda^2,\ \ldots\ ,\ (-1)^{n-1}\lambda^{n-1}\right) Q \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-1} \end{pmatrix} \equiv 0.$$

The left hand side is a rational function and this relation holds if and only if the numerator of this function is equal to 0. The numerator equals the matrix product, because all terms of the denominator are in the product of the characteristic polynomials. Computing this product and ordering for powers of λ yields the conditions of the theorem. $\square$

Remark 3.4. The conditions of Theorem 3.3 state that the sums of elements with alternating signs of the antidiagonals of Q are equal to 0. Here we call diagonal lines running from the lower left to the upper right antidiagonals, contrary to the ordinary diagonals of a matrix. By definition, hermitian matrices meet these conditions for all antidiagonals that do not contain a diagonal element of the matrix. Furthermore the elements in the upper left and in the lower right corner are always equal to 0. Hence, a hermitian matrix which meets these conditions can never be definite.

The Theorems 3.2 and 3.3 yield a complete characterization of the solutions to the pole placement problem. Assume we have computed a solution Q_1 by Theorem 3.1. Denote by $\mathbb{S}$ the set of all matrices $Q_1 + Q$, where Q meets (3.2). Then $\mathbb{S}$ contains all possible solutions. To see this, consider another solution Q_2. Then by Theorem 3.2, we conclude that the difference $Q_1 - Q_2$ meets (3.2). Hence $Q_2 \in \mathbb{S}$.

3.3 Eigenvalue Conditions

The question arises if the condition on the polynomial $r(\lambda)$ defined by (3.1) can be reformulated in terms of the numbers chosen for the pole placement problem. In other words, whether conditions on the eigenvalues of the resulting closed loop system can be stated. The following corollaries give some answers to this question. In the following we use the notation from Theorem 3.1.

Corollary 3.5.

(a) If the coefficients of the polynomial $r(\lambda)$ *defined by* (3.1) *have alternating sign, with the constant coefficient being positive, then it is positive semidefinite on* $i\mathbb{R}$*.*

(b) If the condition of Theorem 3.1 *holds, the eigenvalues of the optimal system satisfy the inequality*

$$\prod_{i=1}^{n} |\lambda_i|^2 \geq \det(A)^2. \tag{3.3}$$

Proof. The assumption of statement (a) is that $r(\lambda)$ has the form

$$r(\lambda) = (-1)^n r_n \lambda^{2n} + (-1)^{n-1} r_{n-1} \lambda^{2n-2} + \ldots - r_1 \lambda^2 + r_0,$$

where $r_i,\ i = 0, \ldots, n$, are nonnegative real numbers. Evaluating this for an imaginary number $i\omega$, where ω is real, yields

$$\begin{aligned} r(i\omega) &= (-1)^n r_n i^{2n} \omega^{2n} + (-1)^{n-1} r_{n-1} i^{2n-2} \omega^{2n-2} + \ldots - r_1 i^2 \omega^2 + r_0 = \\ &= (-1)^n r_n (-1)^n \omega^{2n} + (-1)^{n-1} r_{n-1} (-1)^{n-1} \omega^{2n-2} + \ldots + r_1 \omega^2 + r_0, \end{aligned}$$

which is clearly nonnegative for all real numbers ω.

To prove statement (b) observe that the constant terms of both $p_A(\lambda)$ and $\overline{p_A(-\overline{\lambda})}$ are $\det(A)$. Thus the constant term of $r(\lambda)$ is

$$(-1)^n \prod_{i=1}^{n} (-\lambda_i)\overline{\lambda_i} - \det(A)^2 = \prod_{i=1}^{n} |\lambda_i|^2 - \det(A)^2.$$

If the condition of Theorem 3.1 holds, $r(\lambda)$ is positive semidefinite on $i\mathbb{R}$. Evaluating $r(\lambda)$ at 0 yields

$$r(0) = \prod_{i=1}^{n} |\lambda_i|^2 - \det(A)^2 \geq 0,$$

from which the inequality (3.3) follows. □

Corollary 3.6. *Let the dimension of the system* Σ *be* $n = 2$. *Let* λ_1, λ_2 *be either two complex conjugate numbers in the open left half plane, or two negative real numbers. Then the condition of Theorem* 3.1 *holds, if and only if the following inequalities hold :*

$$\lambda_1\lambda_2 \geq |\det(A)|$$
$$\lambda_1^2 + \lambda_2^2 \geq \operatorname{tr}(A)^2 - 2\det(A)$$

Proof. Denote $c_1 = \det(A)$ and $c_2 = -\operatorname{tr}(A)$. Then the characteristic polynomial equals $p_A(\lambda) = \lambda^2 + c_2\lambda + c_1$.

First assume that λ_1, λ_2 are two complex conjugate numbers. We denote $a = \Re(\lambda_1)$ and $b = \Im(\lambda_1)$. Observe that $\lambda_1\lambda_2 = a^2+b^2$ and $\lambda_1^2+\lambda_2^2 = 2(a^2-b^2)$ hold.

Now $r(\lambda)$ becomes

$$\begin{aligned} r(\lambda) &= (\lambda - a - ib)(\lambda + a - ib)(\lambda - a + ib)(\lambda + a + ib) - p_A(\lambda)\overline{p_A(-\overline{\lambda})} \\ &= \lambda^4 - 2(a^2 - b^2)\lambda^2 + (a^2 + b^2)^2 - \lambda^4 - (2c_1 - c_2^2)\lambda^2 - c_1^2 \\ &= -\lambda^2(2(a^2 - b^2) + 2c_1 - c_2^2) + (a^2 + b^2)^2 - c_1^2. \end{aligned}$$

The condition of Theorem 3.1 states that $r(\lambda) \geq 0$ for $\lambda = ix$ where $x \in \mathbb{R}$. This computes to

$$x^2(2(a^2 - b^2) + 2c_1 - c_2^2) + (a^2 + b^2)^2 - c_1^2 \geq 0 \text{ for } \forall x \in \mathbb{R}.$$

Evaluating this inequality for $x = 0$ and x large enough yields that the last inequality holds if and only if

$$(a^2 + b^2)^2 \geq c_1^2 \text{ and } 2(a^2 - b^2) \geq c_2^2 - 2c_1$$

hold. Following the observation from the beginning, these inequalities are equivalent to the inequalities of the proposition.

If λ_1, λ_2 are two negative real numbers, $r(\lambda)$ becomes

$$\begin{aligned} r(\lambda) &= (\lambda - \lambda_1)(\lambda + \lambda_1)(\lambda - \lambda_2)(\lambda + \lambda_2) - p_A(\lambda)\overline{p_A(-\overline{\lambda})} \\ &= (-2c_1 + c_2^2 - \lambda_1^2 - \lambda_2^2)\lambda^2 + \lambda_1^2\lambda_2^2 - c_1^2 \end{aligned}$$

Again for $\lambda = ix$ where $x \in \mathbb{R}$ this computes to

$$x^2(+2c_1 - c_2^2 + \lambda_1^2 + \lambda_2^2) + \lambda_1^2\lambda_2^2 - c_1^2.$$

With the same argument as before this is greater or equal to zero, if and only if

$$\lambda_1^2 + \lambda_2^2 \geq c_2^2 - 2c_1 \text{ and } \lambda_1\lambda_2 \geq |c_1|$$

hold. $\square$

4 Pole Placement for Multi Dimensional Input

In this section we will establish an algorithm for a cost function matrix Q that solves the pole placement problem for multi dimensional input systems. The algorithm is based on an algorithm for pole placement by use of a Schur method, which originates from A. Varga, see [5]. That algorithm is modified to fit our requirements for linear quadratic optimal control.

4.1 Pole Placement

Given a controllable system $\Sigma = (A, B)$, where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $\mathrm{rank}(B) = m > 1$, and a set of n complex numbers $\lambda_i,\ i = 1,\ldots,n$ in the open left half plane, we ask for a cost function matrix Q such that the eigenvalues of the closed-loop system that minimizes the cost function are exactly the given complex numbers. Without loss of generality, let R be the identity matrix to define the linear quadratic control problem properly. It follows from the results of Section 2.1 that this is possible.

By Theorem 1.6, we have to find a pair of matrices Q and X such that X meets the Riccati equation $XBB^*X - XA - A^*X = Q$ and the eigenvalues of $A - BB^*X$ are exactly the numbers $\lambda_i,\ i = 1,\ldots,n$. The cost function matrix Q has to be positive semi-definite and X hermitian. Furthermore (A, Q) must be observable.

The algorithm decomposes the system matrix into smaller 1×1 or 2×2 blocks and solves the pole placement problem for each of these blocks. This yields a cost function matrix and feedback control matrix in each step, which sum up to the matrices of the overall solution.

4.2 The Real Schur Form

The algorithm uses a special canonical form for real matrices, the real Schur form. We give a short description of this form and develop the consequences for our problem.

Theorem 4.1 (Real Schur Form). *Let $A \in \mathbb{R}^{n\times n}$. Then there exists an orthogonal similarity transformation U such that UAU^T is quasi-upper triangular and has only 1×1 or 2×2 blocks on the diagonal, corresponding to real or to complex conjugate eigenvalues, respectively. Moreover, U can be* *chosen so that the 1×1 or 2×2 blocks on the diagonal appear in any desired order.*

The proof of this theorem together with an algorithm to compute it can be found in [2].

Now we establish relations between the properties of the system matrices and the blocks of the Schur form. The first lemma contains a result on controllability of blocks of the system matrix. Its proof is simple.

Lemma 4.2. *If the pair (A, B) is controllable, where*

$$A = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix} \quad B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix},$$

then (A_3, B_2) is controllable.

The next step is to look at the Riccati equation for a single block of the system matrix and its relation to the Riccati equation for the overall system.

Remark 4.3. Let A and B be defined as in Lemma 4.2. If we take a hermitian matrix X of the form

$$X = \begin{pmatrix} 0 & 0 \\ 0 & X_1 \end{pmatrix}$$

where X_1 has the same size as A_3, then the conditions on X, outlined in Section 4.1, take the following form. The matrix of the closed loop system is

$$\begin{aligned} A - BB^*X &= \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix} - \begin{pmatrix} B_1B_1^* & B_1B_2^* \\ B_2B_1^* & B_2B_2^* \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & X_1 \end{pmatrix} \\ &= \begin{pmatrix} A_1 & A_2 - B_1B_2^*X_1 \\ 0 & A_3 - B_2B_2^*X_1 \end{pmatrix} \end{aligned}$$

and the algebraic Riccati equation is

$$\begin{aligned} &XBB^*X - XA - A^*X = \\ &= \begin{pmatrix} 0 & 0 \\ 0 & X_1B_2B_2^*X_1 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & X_1A_3 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & A_3^*X_1 \end{pmatrix} = Q. \end{aligned}$$

If there is an X such that the Riccati equation holds for the positive semi-definite matrix Q, then this matrix X shifts only eigenvalues of the block A_3. Moreover the problem of finding such a matrix X is reduced to find a matrix X_1 such that $A_3 - B_2B_2^*X_1$ has the required eigenvalues and $X_1B_2B_2^*X_1 - X_1A_3 - A_3^*X_1 = Q_1$ holds, where Q_1 is the corresponding block of Q.

Then we have to consider how two such solutions for two blocks can be added to yield a common solution for both blocks.

Remark 4.4. Let (A, B) be controllable, but no special structure is assumed. Furthermore let X_1 be a hermitian matrix such that $A_1 := A - BB^*X_1$ shifts some eigenvalues of A and $X_1BB^*X_1 - X_1A - A^*X_1 = Q_1$ holds, and let X_2 be another hermitian matrix such that $A_2 := A_1 - BB^*X_2$ shifts again some eigenvalues of A_1 and $X_2BB^*X_2 - X_2A_1 - A_1^*X_2 = Q_2$ holds. Both Q_1 and Q_2 are positive semi-definite matrices. Let $X := X_1 + X_2$. Then X itself shifts the eigenvalues of A to those of A_2 because

$$\begin{aligned} A_2 = A_1 - BB^*X_2 &= A - BB^*X_1 - BB^*X_2 \\ &= A - BB^*(X_1 + X_2) = A - BB^*X \end{aligned}$$

holds. Furthermore the Riccati equation for X holds with $Q = Q_1 + Q_2$:

$$\begin{aligned} & X_2\ BB^*X_2 - X_2A_1 - A_1^*X_2 + X_1BB^*X_1 - X_1A - A^*X_1 \\ = \ & X_2BB^*X_2 - X_2(A - BB^*X_1) - (A^* - X_1BB^*)X_2 \\ & +X_1BB^*X_1 - X_1A - A^*X_1 \\ = \ & (X_1 + X_2)BB^*(X_1 + X_2) - (X_1 + X_2)A - A^*(X_1 + X_2) = Q_1 + Q_2. \end{aligned}$$

4.3 The Reduced Pole Placement Problem

Now we deal with the problem reduced to single blocks of the system matrix. We want to find matrices X and Q for systems of state space dimension one or two. Let p be the dimension of the state space and $p = 1$ or $p = 2$. Let (A, B) be controllable, where $A \in \mathbb{R}^{p\times p}$ and $B \in \mathbb{R}^{p\times m}$ and let $\lambda_i,\ i = 1, \dots, p$, be complex numbers symmetrically distributed with respect to the real axis, and $r = \text{rank}(B)$.

First we compute the orthogonal factorization

$$B = U \begin{pmatrix} \widetilde{B} & 0 \end{pmatrix} V^T$$

where $\widetilde{B} \in \mathbb{R}^{p\times r}$ is upper right triangular, and $U \in \mathbb{R}^{p\times p}$ and $V \in \mathbb{R}^{m\times m}$ are orthogonal matrices. Furthermore we compute $\widetilde{A} = U^TAU$.

It is now possible to reduce the problem to the controllable pair $(\widetilde{A}, \widetilde{B})$ because

$$\begin{pmatrix} \widetilde{B} & 0 \end{pmatrix} \begin{pmatrix} \widetilde{B} & 0 \end{pmatrix}^* = \widetilde{B}\widetilde{B}^*$$

and thus the pole shifting equation and the Riccati equation are the same for the matrices $(\widetilde{B}\ 0)$ and $\widetilde{B}$ and any solution $\widetilde{X}$ for one of them is also a solution for the other. Now we have to consider the cases of the input matrix $\widetilde{B}$ being invertible or not.

4.3.1 Invertible Input

Assume we have found a hermitian matrix $\widetilde{X}$ such that $\widetilde{A} - \widetilde{B}\widetilde{B}^*\widetilde{X} = C$ where C has the desired eigenvalues and $\widetilde{X}\widetilde{B}\widetilde{B}^*\widetilde{X} - \widetilde{X}\widetilde{A} - \widetilde{A}^*\widetilde{X} = \widetilde{Q} > 0$. The strict inequality is used to achieve observability of $(\widetilde{A}, \widetilde{Q})$. Then $\widetilde{X}$ is the solution of the Riccati equation associated to the linear quadratic control problem with the cost functions $\widetilde{R} = I$ and $\widetilde{Q}$.

Since $\widetilde{B}$ is invertible, these two assumptions are equivalent to

$$\widetilde{X} = (\widetilde{B}\widetilde{B}^*)^{-1}(\widetilde{A} - C) \tag{4.1}$$

and, by substituting into this the Riccati equation,

$$\begin{aligned}
& \widetilde{X}^*\widetilde{B}\widetilde{B}^*\widetilde{X} - \widetilde{X}^*\widetilde{A} - \widetilde{A}^*\widetilde{X} = \widetilde{Q} > 0 \\
\Longleftrightarrow\ & (\widetilde{A}^* - C^*)(\widetilde{B}\widetilde{B}^*)^{-1}\widetilde{B}\widetilde{B}^*(\widetilde{B}\widetilde{B}^*)^{-1}(\widetilde{A} - C) \\
& -(\widetilde{A}^* - C^*)(\widetilde{B}\widetilde{B}^*)^{-1}\widetilde{A} - \widetilde{A}^*(\widetilde{B}\widetilde{B}^*)^{-1}(\widetilde{A} - C) > 0 \\
\Longleftrightarrow\ & C^*(\widetilde{B}\widetilde{B}^*)^{-1}C > \widetilde{A}^*(\widetilde{B}\widetilde{B}^*)^{-1}\widetilde{A}
\end{aligned}$$

Hence the following two conditions are sufficient and necessary for the existence of $\widetilde{X}$:

$$(\widetilde{B}\widetilde{B}^*)^{-1}(\widetilde{A} - C) \text{ is hermitian} \tag{4.2}$$

$$C^*(\widetilde{B}\widetilde{B}^*)^{-1}C > \widetilde{A}^*(\widetilde{B}\widetilde{B}^*)^{-1}\widetilde{A} \tag{4.3}$$

In the case $p = 1$ this is reduced to $\widetilde{A} = (\alpha)$, $\widetilde{B} = (\beta)$, $C = (\gamma)$ and the condition becomes

$$\frac{\gamma^2}{\beta^2} > \frac{\alpha^2}{\beta^2} \Leftrightarrow \gamma^2 > \alpha^2 \Leftrightarrow |\gamma| > |\alpha|.$$

The case $p = 2$ is not so straightforward. Two problems have to be solved. Firstly we have to decide whether a matrix C exists that meets the two conditions for a given pair of eigenvalues. Secondly we have to find such a matrix.

For given matrices A and B, and for two complex numbers λ_1, λ_2, the matrix C has to meet the following conditions. The eigenvalues of C are λ_1, λ_2 and C meets the two conditions (4.2) and (4.3), where $\widetilde{A}$ and $\widetilde{B}$ are replaced by A and B.

The eigenvalue condition (4.1) is equivalent to the following two equations :

$$c_{1,1} + c_{2,2} = \lambda_1 + \lambda_2,$$
$$c_{1,1}c_{2,2} - c_{1,2}c_{2,1} = \lambda_1\lambda_2.$$

Condition (4.2) yields the following equation :

$$b_{1,1}(a_{1,2} - c_{1,2}) + b_{1,2}(a_{2,2} - c_{2,2}) = b_{1,2}(a_{1,1} - c_{1,1}) + b_{2,2}(a_{2,1} - c_{2,1}).$$

We can use these three equations to parameterize all matrices C meeting these three equations. Substituting such a parametrization into condition (4.3) can be used to decide whether a matrix C meeting all conditions exists and to compute it, if possible. By computing the main minors of the matrix $C^*(BB^*)^{-1}C - A^*(BB^*)^{-1}A$, one can compute whether condition (4.3) is met or not.

4.3.2 Non Invertible Input

In this case $p = 2$ and $\widetilde{B} = \begin{pmatrix} \beta \\ 0 \end{pmatrix}$ holds because $(\widetilde{A}, \widetilde{B})$ are controllable and therefore $p = 1$ is impossible. Then there exists a transformation T such that

$$\overline{A} := T\widetilde{A}T^{-1} \quad \overline{B} := TB$$

are in control canonical form. Now we apply Theorem 3.1 and get a positive semi-definite matrix $\overline{Q}$ such that the stabilising solution $\overline{X}$ places the poles of the closed loop system in the desired places. Conditions on the eigenvalues and locations of the poles are discussed in Corollary 3.6 which gives us a complete characterization of the possible solutions.

By setting $\widetilde{X} = T^*\overline{X}T$ we find a solution for this case. Furthermore, the cost function matrix equals $\widetilde{Q} = T^*\overline{Q}T$.

4.4 An Algorithm

The matrices A, B and R are given as in the beginning of this section. Let $\lambda_j,\ j = 1, \ldots, n$ be complex numbers symmetrically distributed with respect to the real axis in the open left half plane. Furthermore, neither λ_j nor $-\overline{\lambda}_j$ are eigenvalues of A for $j = 1, \dots, n$.

The syntax $A \leftarrow B$ denotes that the value of B is assigned to the variable A. Then any further use of A references this new value.

Algorithm 4.5. Sequential Computing of a cost function matrix Q which achieves Pole Placing

1. Reduce A to the real Schur form, accumulating the transformations in K, set $A \leftarrow KAK^*$ and $B \leftarrow KB$.

2. Set $X = 0$ and $Q = 0$ where $X, Q \in \mathbb{R}^{n\times n}$, and $k = 1$.

3. Set $\overline{A}$ equal to the last block in A of order p ($p = 1$ or 2) and set $\overline{B}$ equal to the last p rows of B.

4. Compute a hermitian matrix $\overline{X}$ such that $\overline{BB}^*\overline{X}$ shifts the eigenvalues of $\overline{A}$ to the numbers $\lambda_j,\ j = k, \dots, k+p-1$ and
$$\overline{XBB}^*\overline{X} - \overline{XA} - \overline{A}^*\overline{X} =: \overline{Q} \geq 0$$
holds. The new eigenvalues must meet some conditions outlined in Section 4.3 for such an $\overline{X}$ to exist. If there is no solution $\overline{X}$ the algorithm stops without solving the problem.

5. Extend $\overline{X}$ and $\overline{Q}$ to $n \times n$ matrices by putting $\overline{X}$ and $\overline{Q}$ into the lower right corner and choosing the other elements equal to zero. Compute $X \leftarrow X + \overline{X}$, $Q \leftarrow Q + \overline{Q}$ and $A \leftarrow A - BB^*\overline{X}$.

6. Move the last block of A to position (k, k) accumulating the transformations in K_1 and compute $K \leftarrow K_1K$, $B \leftarrow K_1B$, $X \leftarrow K_1XK_1^*$ and $Q \leftarrow K_1QK_1^*$.

7. Set $k \leftarrow k + p$. If $k \leq n$ go to step 3, otherwise continue to step 8.

8. Set $X \leftarrow K^*XK$ and $Q \leftarrow K^*QK$ and stop.

The algorithm computes a matrix Q and the corresponding feedback matrix X for the given system. The matrix K accumulates the orthogonal transformations and is used in Step 8 to transform X and Q back to the original system. In Step 5 Remark 4.4 is used to combine the feedback and cost function matrices yielding new ones shifting p more eigenvalues. Remark 4.3 and the conclusions of Sections 4.3.1 and 4.3.2 about systems of dimension 1 or 2 are used in Step 4 to compute feedback and cost function matrices that shift only one or two eigenvalues.

Remark 4.6. The locations of the complex numbers λ_j is important in Step 4, where the existence of a stabilizing feedback matrix depends on the relation between the original eigenvalues and these numbers. Thus the ordering of the numbers λ_j and of the blocks in the real Schur form can have an influence on the outcome of the algorithm. It is possible that the algorithm stops in Step 4 without completing, if a certain order is used, whereas for a different order it may well finish and compute a solution. Thus the algorithm might be more suitable for an interactive execution, where the user can influence the choice of the new eigenvalues.

4.5 Observability

In Theorem 1.6 it is assumed that the pair (A, Q) is observable. We will show now that the cost function matrix Q computed by the above algorithm meets this requirement. By Theorem 1.4 and the duality of controllability and observability we have to show that for each eigenvalue μ_i, $i = 1, \dots, n$, of A the relation $\operatorname{rank}((A - \lambda I, Q)^*) = n$ holds. This is equivalent to

$$\begin{pmatrix} A - \lambda I \\ Q \end{pmatrix} x \neq 0, \quad x \in \mathbb{R}^n, \quad x \neq 0. \tag{4.4}$$

Assume that

$$\begin{pmatrix} A - \lambda I \\ Q \end{pmatrix} x = 0$$

for some vector x. Then if follows that x is an eigenvector to some eigenvalue μ_i of A. Furthermore $Qx = 0$ holds.

The transformations of the algorithm do not change the eigenvector and its eigenvalue, if the block corresponding to the eigenvalue is not changed. Assume A is in Real Schur Form and the block A_{jj} corresponding to the eigenvalue μ_i is not the last one. Then the eigenvector x has only nonzero entries up to the row corresponding to the lower right corner of the block A_{jj}.

$$\begin{pmatrix} \ddots & \vdots & \vdots \\ 0 & A_{jj} & A_{jj+1} \\ 0 & 0 & A_{j+1j+1} \end{pmatrix} \begin{pmatrix} \vdots \\ x_j \\ 0 \end{pmatrix} = \mu_i \begin{pmatrix} \vdots \\ x_j \\ 0 \end{pmatrix}$$

The pole placement transformation $A + BF$ concerning the block A_{j+1j+1} does not affect the eigenvector x.

$$\begin{pmatrix} \ddots & \vdots & \vdots \\ 0 & A_{jj} & A_{jj+1} + B_j F_j \\ 0 & 0 & A_{j+1j+1} + B_{j+1} F_{j+1} \end{pmatrix} \begin{pmatrix} \vdots \\ x_j \\ 0 \end{pmatrix} =$$

$$\begin{pmatrix} \ddots & \vdots & \vdots \\ 0 & A_{jj} & A_{jj+1} \\ 0 & 0 & A_{j+1j+1} \end{pmatrix} \begin{pmatrix} \vdots \\ x_j \\ 0 \end{pmatrix} + \begin{pmatrix} \ddots & \vdots & \vdots \\ 0 & 0 & B_j F_j \\ 0 & 0 & B_{j+1} F_{j+1} \end{pmatrix} \begin{pmatrix} \vdots \\ x_j \\ 0 \end{pmatrix} = \mu_i \begin{pmatrix} \vdots \\ x_j \\ 0 \end{pmatrix}$$

Thus the transformations of the algorithm leave the eigenvector x and the eigenvalue μ_i invariant as long as the block corresponding to the eigenvalue μ_i is not changed. Also the partial cost function matrix Q_{j+1} maps x to 0.

Since we assume that the eigenvalues λ_i of the resulting closed loop system are not equal to the original eigenvalues or the eigenvalues mirrored at the imaginary axis, there is a step in the algorithm that changes the eigenvalue μ_i. At this step the block A_{jj} corresponding to μ_i is in the lower right corner of the system matrix and the algorithm yields a partial cost function matrix Q_j such that (A_{jj}, Q_j) is observable.

Since x is an eigenvector of A, the restriction x_j of x is an eigenvector of A_{jj}. Thus, by observability, $Q_j x_j \neq 0$ and, furthermore, $Qx \neq 0$. But this is a contradiction to $Qx = 0$. Thus (A, Q) is observable and we can apply Theorem 1.6.

5 Examples

In this section examples for the results obtained are given. These examples were calculated using MapleV Release 4 with the precision set to twenty digits. The numbers given are rounded to four decimals.

5.1 Perturbations of the Cost Function

We will consider the following controllable system $\Sigma = (A, B)$ given in control canonical form

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -8 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

The system matrix A has the eigenvalues $1,\ 2,\ 2$ and thus is unstable. Moreover we consider an LQ problem with the cost function matrices

$$R = (1) \text{ and } Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

The pair (A, Q) is observable, hence by Theorem 1.6 there exists a feedback control such that the resulting closed loop system is stable. Indeed, the eigenvalues of the closed loop system are $-2.8326,\ -1.3409,\ -1.0856$.

We perturb the cost function matrix Q by adding a positive value ε to the element in the upper left corner. Then the eigenvalues of the closed loop system resulting from solving the perturbed LQ problem move towards the zeros of the polynomial $q(\lambda) = \lambda^6 - (p + \varepsilon)$ for $\varepsilon \to \infty$. Figure 1 visualizes the paths of the system eigenvalues in black and the paths of the zeros of $q(\lambda)$ in grey.

The lower bound for $(p + \varepsilon)$ is 2530944. The black crosses and grey circles mark the locations of the system eigenvalues and the zeros of $q(\lambda)$ for this value. The circles show the areas, where exactly one eigenvalue of the system must lie. Their centres are the zeros of $q(\lambda)$ and their radius is 5.8370.

5.2 Pole Placement for One Dimensional Input

The next example deals with pole placement for one dimensional input systems. We consider the system Σ of the last section and the cost function

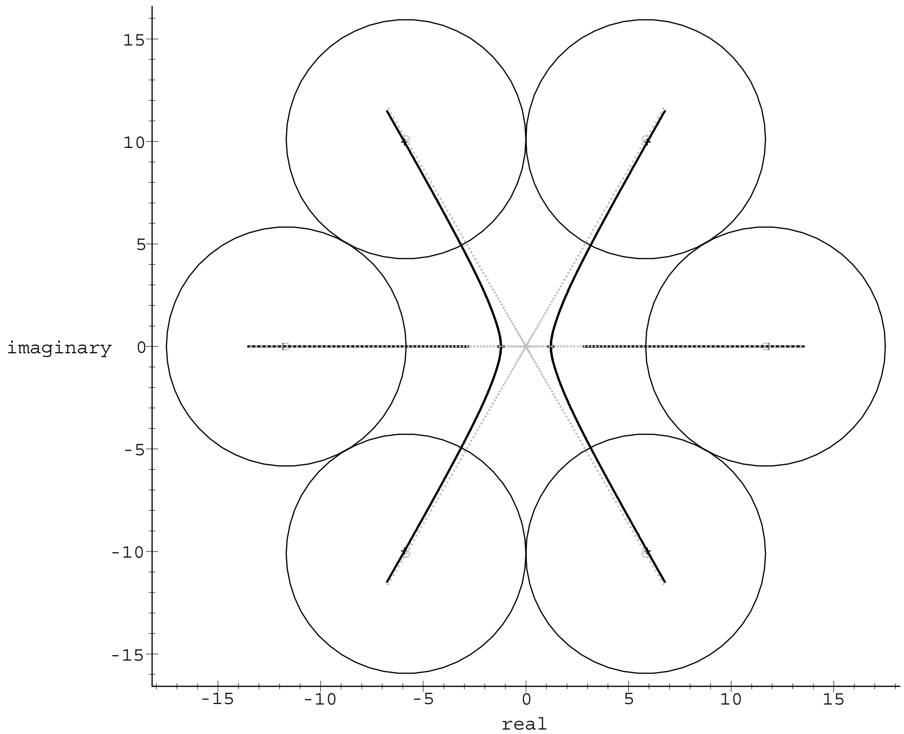

Figure 1: Eigenvalues under perturbation

matrix $R = (1)$. We compute a cost function matrix Q such that (A, Q) is observable and the closed loop system resulting from solving the LQ problem has the eigenvalues $-3,\ -2-i,\ -2+i$.

The characteristic polynomial of A is

$$p_A(\lambda) = \lambda^3 - 5\lambda^2 + 8\lambda - 4.$$

By Theorem 3.1 there exists such a matrix Q, if the term

$$r(\lambda) := (-1)^n \prod_{i=1}^{n} (\lambda - \lambda_i)(\lambda + \overline{\lambda_i}) - (p_A(\lambda)\overline{p}_A(-\overline{\lambda})$$

is nonnegative for all $\lambda \in i\mathbb{R}$. The term $r(\lambda)$ computes to

$$r(\lambda) = 6\lambda^4 - 55\lambda^2 + 209.$$

The zeros of $r(\lambda)$ are $2.2897 \pm 0.8120i,\ -2.2897 \pm 0.8120i$ and therefore it is nonnegative on $\imath\mathbb{R}$. Thus the condition of Theorem 3.1 is fulfilled and we can find a factorization into $r(\lambda) = v(\lambda)\overline{v(-\overline{\lambda})}$. A possible factorization is

$$v(\lambda) = \sqrt{6}(\lambda^2 + 4.5794\lambda + 5.9010).$$

The next step is to compute a realization of the real rational function

$$\frac{v(\lambda)}{p_A(\lambda)} = \frac{\sqrt{6}(\lambda^2 + 4.5794\lambda + 5.9010)}{\lambda^3 - 5\lambda^2 + 8\lambda - 4}.$$

The following matrices yield a realization of this function. The algorithm to compute such a realization is described in detail in [4].

$$\widetilde{A} = \begin{pmatrix} 2.7150 & -5.0504 & 7.8946 \\ 0.1107 & 1.8552 & -2.8502 \\ 0.0113 & 0.1855 & 0.4298 \end{pmatrix}, \quad \widetilde{B} = \begin{pmatrix} -113.5424 \\ 15.7412 \\ -1.4533 \end{pmatrix},$$
$$\widetilde{S} = \begin{pmatrix} -0.0907 & -0.5739 & -0.8139 \end{pmatrix}$$

To compute the state space transformation back to the original system matrix, we follow the steps of Theorem 3.1. The last column of the transformation matrix T is equal to $\widetilde{B}$. The second column computes to $T_2 = (\widetilde{A} - 5I)\widetilde{B}$ and the first to $T_1 = \widetilde{A}T_2 + 8\widetilde{B}$. These computations yield the transform matrix

$$T = \begin{pmatrix} -92.9839 & 168.4727 & -113.5424 \\ 13.4986 & -57.9274 & 15.7412 \\ -16.9160 & 8.2836 & -1.4533 \end{pmatrix}.$$

With this T we compute S, and furthermore the cost function matrix $Q = S^*S$:

$$Q = \begin{pmatrix} 209.0000 & 162.1640 & 35.4119 \\ 162.1640 & 125.8237 & 27.4762 \\ 35.4119 & 27.4762 & 6.0000 \end{pmatrix}.$$

5.2.1 Finding a Definite Q

Furthermore we want to compute a cost function matrix Q_1 which is positive definite and still produces a solution to the LQ problem with the desired eigenvalues. Assume that we have found such a matrix Q_1. Then we know from Theorem 3.2 that the difference $Q_\Delta = Q_1 - Q$ meets the equation (3.2). Theorem 3.3 then gives us a characterization of all matrices solving equation (3.2).

Conversely, if we find a hermitian matrix Q_Δ that solves equation (3.2) and such that $Q_1 = Q + Q_\Delta$ is positive definite then Q_1 yields the same solution to the pole placement problem.

We can parameterize all Q_Δ meeting the condition of Theorem 3.3 in the following way

$$Q_\Delta(a,b,c) = \begin{pmatrix} 0 & a & b \\ a & 2b & c \\ b & c & 0 \end{pmatrix}.$$

Adding this to Q we have to find parameters $a,\ b,\ c$ such that the sum is positive definite. With some experimenting, one can find that $a = -2,\ b = -1,\ c = -1$ yield a matrix with the desired properties. We get the cost function matrix

$$Q_1 = \begin{pmatrix} 209.0000 & 160.1640 & 34.4119 \\ 160.1640 & 125.8237 & 26.4762 \\ 34.4119 & 26.4762 & 6.0000 \end{pmatrix}$$

which is positive definite and for which the closed loop system resulting from solving the LQ problem has the desired eigenvalues $-3,\ -2-i,\ -2+i$.

5.3 Pole Placement for Multi Dimensional Input

The next example corresponds to the Algorithm 4.5. It uses all three cases for Step 4 and will also demonstrate that a different choice of the order of blocks can stop the algorithm.

The system $\Sigma = (A, B)$ is given by the following controllable pair of matrices:

$$A = \begin{pmatrix} -4.3173 & -6.5467 & -4.1661 & -7.4271 & 5.5611 \\ 6.7380 & 7.7913 & 1.9065 & 5.6858 & -6.6460 \\ 2.7412 & 0.3797 & 1.6670 & -0.3939 & -3.2227 \\ -2.2311 & -2.5102 & 0.0908 & 0.2577 & 2.4016 \\ -4.2429 & -2.7632 & 0.4777 & -2.3180 & 6.6014 \end{pmatrix},$$

$$B = \begin{pmatrix} -0.1983 & 0.0118 & -0.5116 \\ -0.0197 & 0.0160 & 0.5793 \\ 0.0261 & 0.0509 & 0.2670 \\ 0.0378 & 0.0159 & -0.4467 \\ -0.1783 & -0.0359 & -0.2497 \end{pmatrix}.$$

The system matrix A has the eigenvalues $2,\ 2\pm 4i,\ 3\pm i$. The set of complex numbers is $\lambda_j = \{-4+i, -4-i, -5+2i, -5-2i, -3\}$. The notation used will be the same as in the definition of the algorithm.

Step 1. Transform A to the Real Schur Form with an orthogonal transformation K, which is computed such that the blocks on the diagonal of A correspond to the eigenvalues or pairs of conjugate complex eigenvalue in the order $2,\ 2\pm 4i,\ 3\pm i$. The system matrices become

$$A = \begin{pmatrix} 2 & 2.2419 & -9.8087 & -3.0411 & 7.7630 \\ 0 & 0.1134 & 11.4682 & 4.1734 & -8.6016 \\ 0 & -1.7055 & 3.8866 & -2.2149 & -4.4243 \\ 0 & 0 & 0 & 2.6979 & 2.3482 \\ 0 & 0 & 0 & -0.4647 & 3.3021 \end{pmatrix},$$

$$B = \begin{pmatrix} 0.2713 & 0.0209 & 0.4595 \\ 0 & 0.0097 & -0.7543 \\ 0 & -0.0632 & -0.2587 \\ 0 & 0 & 0.2790 \\ 0 & 0 & -0.0744 \end{pmatrix}.$$

Step 2. Set $S = 0$ and $Q = 0$, where $S,\ Q \in \mathbb{R}^{5\times 5}$, and set $k = 1$.

Step 3. We pick the last block of A and the corresponding rows of B.

$$\overline{A} = \begin{pmatrix} 2.6979 & 2.3482 \\ -0.4647 & 3.3021 \end{pmatrix}, \quad \overline{B} = \begin{pmatrix} 0 & 0 & 0.2790 \\ 0 & 0 & -0.0744 \end{pmatrix}$$

$\overline{A}$ has the eigenvalues $3\pm i$ and we will assign $-4\pm i$ to this block in the next step.

Step 4. Now we have to decide how to compute a stabilizing feedback matrix $\overline{S}$. Obviously $\overline{B}$ has rank 1 and so we use the not invertible case described in Section 4.3.2. We set $\overline{B}_1$ to

$$\overline{B}_1 = \begin{pmatrix} 0.2790 \\ -0.0744 \end{pmatrix}.$$

Observe that $\overline{B}\overline{B}^* = \overline{B}_1\overline{B}_1^*$ and therefore any solution for $\overline{B}_1$ is also a solution for $\overline{B}$. Then we transform the system $(\overline{A}, \overline{B}_1)$ to control canonical form and denote the new system matrices as $\widetilde{A}$ and $\widetilde{B}$.

Finally we use Theorem 3.1 to compute a semidefinite $\widetilde{Q}$ that shifts the eigenvalues of $\widetilde{A}$ to $\lambda_1 = -4+i,\ \lambda_2 = -4-i$. Corollary 3.6 states conditions

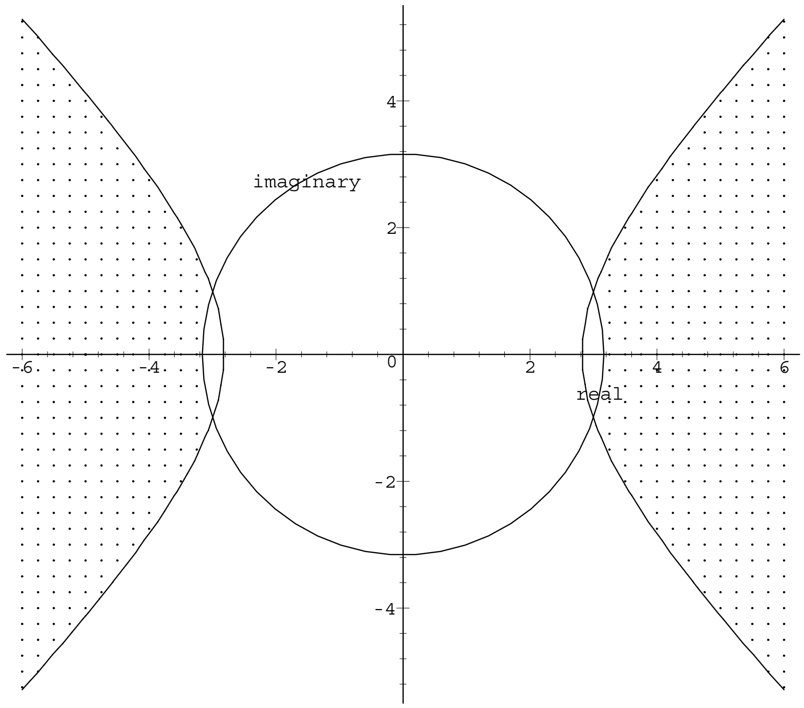

Figure 2: Valid eigenvalues for one dimensional input

on the desired eigenvalues. In Figure 2, the dotted area visualizes the set of possible eigenvalues. Clearly, the pair of eigenvalues $-4 \pm i$ meets the conditions.

Following the procedure described in Section 5.2, we find the cost function matrix

$$\widetilde{Q} = \begin{pmatrix} 189 & -51.4393 \\ -51.4393 & 14 \end{pmatrix}.$$

We will also look for a positive definite $\widetilde{Q}$ and use Theorem 3.2 and 3.3 to find one. A suitable solution is

$$\widetilde{Q} \leftarrow \widetilde{Q} + \begin{pmatrix} 0 & 10 \\ 10 & 0 \end{pmatrix} = \begin{pmatrix} 189 & -41.4393 \\ -41.4393 & 14 \end{pmatrix}.$$

Finally we compute the stabilizing feedback matrix $\widetilde{S}$:

$$\widetilde{S} = \begin{pmatrix} 237.4393 & 7 \\ 7 & 14 \end{pmatrix}.$$

Transforming back to $(\overline{A}, \overline{B})$ and applying the fifth step of the algorithm yields

$$S = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 383.2696 & 1766.7970 \\ 0 & 0 & 0 & 1766.7970 & 10386.9196 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 178.1902 & 213.4402 \\ 0 & 0 & 0 & 213.4402 & 1623.4117 \end{pmatrix}$$

and

$$A = \begin{pmatrix} 2 & 2.2419 & -9.8087 & 8.2539 & 136.5326 \\ 0 & 0.1134 & 11.4682 & -14.3674 & -219.9774 \\ 0 & -1.7055 & 3.8866 & -8.5742 & -76.9244 \\ 0 & 0 & 0 & 9.5554 & 80.5274 \\ 0 & 0 & 0 & -2.2942 & -17.5554 \end{pmatrix}.$$

Step 5. Here we deviate from the algorithm a little. Instead of moving the last block of A to position $(1,1)$, we exchange the two 2×2 blocks. This may yield a different outcome, but will also give a solution.

We transform A and B in such a way that the two 2×2 blocks on the diagonal of A exchange their places and correspond to the following order of the eigenvalues : $2, -4 \pm i, 2 \pm 4i$. This is accomplished with an orthogonal transformation K_1. Furthermore $K \leftarrow K_1 K$, $B \leftarrow K_1 B$, $S \leftarrow K_1 S K_1^*$ and $Q \leftarrow K_1 Q K_1^*$ are computed. This yields the following state space matrices :

$$A = \begin{pmatrix} 2 & -7.2719 & -44.3844 & -128.2351 & 18.5315 \\ 0 & 13.5898 & 73.0398 & 220.8940 & -39.9979 \\ 0 & -4.2497 & -21.5898 & -71.4695 & 7.7012 \\ 0 & 0 & 0 & 0.5684 & 2.8012 \\ 0 & 0 & 0 & -6.4435 & 3.4316 \end{pmatrix},$$

$$B = \begin{pmatrix} -0.2713 & -0.0209 & -0.4595 \\ 0 & 0.0258 & 0.8061 \\ 0 & 0.0409 & -0.2621 \\ 0 & -0.0046 & -0.0135 \\ 0 & 0.0416 & -0.0263 \end{pmatrix}.$$

Step 6. We add the size of the last block to the index variable k and get $k = 3$. As we have not yet finished we continue with the third step of the algorithm. Again we pick the last block of A and the corresponding rows of B.

$$\overline{A} = \begin{pmatrix} 0.5684 & 2.8012 \\ -6.4435 & 3.4316 \end{pmatrix}, \quad \overline{B} = \begin{pmatrix} 0 & -0.0046 & -0.0135 \\ 0 & 0.0416 & -0.0263 \end{pmatrix}$$

$\overline{A}$ has the eigenvalues $2 \pm 4i$ and we will assign $-5 \pm 2i$ to this block in the next step.

Step 7. We set $\widetilde{B}$ to be the invertible 2×2 block of $\overline{B}$ and $\widetilde{A} = \overline{A}$. If we find a feedback matrix for the pair $(\widetilde{A}, \widetilde{B})$, the same matrix is also a solution for the pair $(\overline{A}, \overline{B})$.

Now $\widetilde{B}$ is invertible. We try to find a matrix C such that the eigenvalues of C are $-5 \pm 2i$ and it meets the two conditions of the invertible case. A suitable C is

$$C = \begin{pmatrix} -1.5 & 1.9922 \\ -8.1570 & -8.5 \end{pmatrix}.$$

Furthermore the feedback matrix $\overline{S}$ and the cost function matrix $\overline{Q}$ compute to

$$\overline{S} = \begin{pmatrix} 10142.7787 & 24.9029 \\ 24.9029 & 4932.42035 \end{pmatrix}, \quad \overline{Q} = \begin{pmatrix} 9812.7833 & 11773.6575 \\ 11773.6575 & 24880.0255 \end{pmatrix}.$$

The fifth step of the algorithm yields

$$S = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 81.2187 & 279.2251 & 740.9266 & -243.0759 \\ 0 & 279.2251 & 1098.4442 & 3081.5113 & -912.6813 \\ 0 & 740.9266 & 3081.5113 & 18962.9998 & -2489.6391 \\ 0 & -243.0759 & -912.6813 & -2489.6391 & 5702.7257 \end{pmatrix},$$

$$Q = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 37.7603 & 60.2644 & 76.1492 & -74.3378 \\ 0 & 60.2644 & 180.6516 & 447.4062 & -165.6092 \\ 0 & 76.1492 & 447.4062 & 11223.5109 & 11442.5504 \\ 0 & -74.3378 & -165.6092 & 11442.5504 & 25052.4880 \end{pmatrix}$$

and

$$A = \begin{pmatrix} 2 & -7.2719 & -44.3844 & -192.4148 & -36.9090 \\ 0 & 13.5898 & 73.0398 & 332.9767 & 59.4823 \\ 0 & -4.2497 & -21.5898 & -105.6490 & -34.7416 \\ 0 & 0 & 0 & -1.5 & 1.9922 \\ 0 & 0 & 0 & -8.1570 & -8.5 \end{pmatrix}.$$

Step 8. To move the last remaining unstable eigenvalue, we transform A and B such that the 1×1 block is in the lower right corner of A. The blocks now correspond to the following ordering of the eigenvalues : $-5 \pm 2i, -4 \pm i, 2$. Again an orthogonal transformation K_1 is used. Furthermore $K \leftarrow K_1 K$, $B \leftarrow K_1 B$, $S \leftarrow K_1 S K_1^*$ and $Q \leftarrow K_1 Q K_1^*$ are computed. This yields the following state space matrices :

$$A = \begin{pmatrix} -4.5097 & -69.0646 & 61.4010 & -341.5588 & -217.2932 \\ 0.0614 & -5.4903 & -11.1823 & -17.3236 & -18.3695 \\ 0 & 0 & -2.9491 & -3.4261 & -5.5612 \\ 0 & 0 & 0.6142 & -5.0509 & -5.3381 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix},$$

$$B = \begin{pmatrix} 0.1328 & 0.0232 & 0.9626 \\ 0.0436 & -0.0514 & 0.0529 \\ -0.0937 & -0.0348 & 0.0158 \\ -0.1336 & -0.0118 & 0.0296 \\ 0.1657 & 0 & 0 \end{pmatrix}.$$

Step 9. Again we add the size of the last block to the index set k and get $k = 5$. Again we continue with the third step of the algorithm. We pick the last block of A and the corresponding rows of B.

$$\overline{A} = \begin{pmatrix} 2 \end{pmatrix}, \quad \overline{B} = \begin{pmatrix} 0.1657 & 0 & 0 \end{pmatrix}$$

$\overline{A}$ has the eigenvalue 2 and we will assign -3 to this block in the next step.

Step 10. This is the 1×1 case and very easy to compute. As the conditions for the invertible case are met, we can directly compute the feedback matrix $\overline{S}$ and cost function matrix $\overline{Q}$.

$$\begin{aligned} \overline{S} &= (\overline{BB}^*)^{-1}(\overline{A} - C) = \frac{2 - (-3)}{0.02745} = 182.1750 \\ \overline{Q} &= S\overline{BB}^* S - SA - A^* S = 182.1750 \end{aligned}$$

Applying finally the fifth step of the algorithm yields

$$S = \begin{pmatrix} 7.2206 & -3.7345 & 8.5309 & 11.5994 & 9.3735 \\ -3.7345 & 668.3657 & 90.4523 & 903.2771 & 606.5477 \\ 8.5309 & 90.4523 & 7562.1195 & -6607.4338 & -1082.2162 \\ 11.5994 & 903.2771 & -6607.4338 & 13088.7778 & 6570.5137 \\ 9.3735 & 606.5477 & -1082.2162 & 6570.5137 & 4701.0798 \end{pmatrix},$$

$$Q = \begin{pmatrix} 13.2750 & -16.7669 & 8.1460 & 13.3596 & 9.1537 \\ -16.7669 & 901.2277 & 3143.5409 & 1815.2697 & 3017.5441 \\ 8.1460 & 3143.5409 & 13849.2349 & 5939.3491 & 11786.6218 \\ 13.3596 & 1815.2697 & 5939.3491 & 8213.0072 & 9492.5195 \\ 9.1537 & 3017.5441 & 11786.6218 & 9492.5195 & 13699.8411 \end{pmatrix},$$

and

$$A = \begin{pmatrix} -4.5097 & -69.0646 & 61.4010 & -341.5588 & -221.3003 \\ 0.0614 & -5.4903 & -11.1823 & -17.3236 & -19.6857 \\ 0 & 0 & -2.9491 & -3.4261 & -2.7338 \\ 0 & 0 & 0.6142 & -5.0509 & -1.3064 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix}.$$

Step 11. Adding the block size to the index variable yields $k = 6$ and the algorithm stops. The matrix A now has the desired eigenvalues. To compute the feedback matrix and the cost function matrix in the original state space, we use the accumulated transformation K and follow the eighth step of the algorithm.

5.3.1 Influence of the Order

If we would have computed the Real Schur Form in *Step 1.* such that the blocks on the diagonal of A would have corresponded to the ordering $2,\ 3 \pm i,\ 2 \pm 4i$ of the eigenvalues, then the algorithm would have stopped. A transformation producing this ordering yields the system matrices

$$A = \begin{pmatrix} 2 & 0.0844 & 3.7120 & 8.4587 & 9.2420 \\ 0 & 2.4018 & -4.7777 & -10.0550 & -9.0138 \\ 0 & 0.2842 & 3.5982 & 1.3934 & 4.5547 \\ 0 & 0 & 0 & 1.7124 & -4.8502 \\ 0 & 0 & 0 & 3.3159 & 2.2876 \end{pmatrix},$$

$$B = \begin{pmatrix} 0.2713 & 0.0209 & 0.4595 \\ 0 & -0.0124 & -0.8362 \\ 0 & -0.0211 & 0.0194 \\ 0 & 0.0289 & -0.1292 \\ 0 & 0.0516 & 0.0553 \end{pmatrix}.$$

We set

$$\widetilde{A} = \begin{pmatrix} 1.7124 & -4.8502 \\ 3.3159 & 2.2876 \end{pmatrix} \text{ and } \widetilde{B} = \begin{pmatrix} 0.0289 & -0.1292 \\ 0.0516 & 0.0553 \end{pmatrix}.$$

These matrices describe the system, whose eigenvalues we want to shift to $-4 \pm i$. But there is no matrix C meeting the two conditions for the invertible case. Thus the algorithm stops, because there is no feedback matrix that shifts the eigenvalues to the desired locations.

References

[1] B.D.O. Anderson and J.B. Moore: Optimal Contro, Linear Quadratic Methods, Prentice Hall, Inc., 1990

[2] R.A. Horn and C.R. Johnson: Matrix Analysis, Cambridge University Press, Cambridge etc., 1985.

[3] P. Lancaster and M. Tismenetsky: The Theory of Matrices, Second Edition with Applications, Academic Press, San Diego etc., 1985.

[4] E.D. Sontag: Mathematical Control Theory, Springer Verlag, New York etc., 1990.

[5] A. Varga: A Schur method for pole assignment, IEEE Trans. Autom. Control, AC-26, (1981), 517-519.

[6] Mathematical System Theory text book, used at the Vrije Universiteit of Amsterdam.