FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

BACHELOR THESIS

Alexander Gažo

Pole Shifting Theorem in Control Theory

Department of Algebra

Supervisor of the bachelor thesis: doc. RNDr. Jiří Tůma, DrSc.
Study programme: Mathematics
Study branch: Mathematical Structures

Prague 2019

I declare that I carried out this bachelor thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date .... signature of the author

Dedication.

Title: Pole Shifting Theorem in Control Theory

Author: Alexander Gažo

Department: Department of Algebra

Supervisor: doc. RNDr. Jiří Tůma, DrSc., Department of Algebra

Abstract: Abstract.

Keywords: Pole-Shifting Theorem

Contents

1. Introduction

1.1 Basics

Pole shifting theorem is one of the basic results of the theory of linear dynamical systems with linear feedback. It claims that in case of controllable systems one can achieve an arbitrary asymptotic behavior by a suitably chosen feedback. To understand this crucial theorem, we must first describe few basic concepts.

1.1.1 Systems of First Order Differential Equations

Remark. Let $f(t)$ be a function of time $t \in \mathbb{R}^+$. We will denote its derivative with respect to t by

$$\dot{f}(t) = \frac{d}{dt} f(t) \ .$$

Definition. *A* **system of linear differential equations of order of one with constant coefficients** *is a system*

$$\begin{aligned} \dot{x}_1(t) &= a_{1,1}x_1(t) + \ldots + a_{1,n}x_n(t) \\ &\vdots \\ \dot{x}_n(t) &= a_{n,1}x_1(t) + \ldots + a_{n,n}x_n(t) \ . \end{aligned}$$

The system can be written in a matrix form

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

where $x(t) = (x_1(t), \ldots, x_n(t))^T \in \mathbb{C}^n$ *is a* **state vector** *(shortly* **state***) of the system. The matrix* $A \in \mathbb{C}^{n\times n}$*,* $A = (a_{i,j})$ *is a* **fundamental matrix** *of the system. The* **initial condition** *of the system is the state* $x(0)$*.*

We will use the matrix form, as it is a very compact way of describing the system.

To express solution of this system in similarly compact way we will establish the notion of a matrix exponential.

Definition. *Let* X *be a real or complex square matrix. The exponential of* X*, denoted by* e^X*, is the square matrix of the same type defined by the series*

$$e^X = \sum_{k=0}^{\infty} \frac{1}{k!} X^k \ ,$$

where X^0 *is defined to be the identity matrix* I *of the same type as* X*.*

For this definition to make sense, we need to show that the series converges for any real or complex square matrix. Firstly, we will define what does it mean for a matrix series to converge. In this text, we will be using Frobenius norm to describe the notion of the convergence of a matrix series.

Definition. Frobenius norm *is a matrix norm, denoted by* $\|\cdot\|_F$*, which for an arbitrary* $n \times m$ *matrix* A *is defined as*

$$\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}|a_{i,j}|^2} \ .$$

Remark. In what follows, $\mathbb{K}$ will denote a field of either real or complex numbers.

Lemma 1. *Then Frobenius norm satisfies following statements for any matrices* A*,* B*,* $C \in \mathbb{K}^{n\times m}$*,* $D \in \mathbb{K}^{m\times o}$ *and any scalar* $\alpha \in \mathbb{K}$*.*

1. $\|A+B\|_F \leq \|A\|_F + \|B\|_F$,
2. $\|\alpha A\|_F = |\alpha|\|A\|_F$,
3. $\|A\|_F \geq 0$ *with equality occurring if and only if* $A = O_{n\times m}$,
4. $\|CD\|_F \leq \|C\|_F\|D\|_F$.

Proof. First three points can be simply shown using the definition of the Frobenius form and properties of the absolute value.

The forth point follows from the Cauchy–Schwarz inequality

$$\|CD\|_F^2 = \sum_{i=1}^{m}\sum_{j=1}^{o}\|c_i d_j\|_2^2 \leq \sum_{i=1}^{m}\sum_{j=1}^{o}\|c_i\|_2^2\|d_j\|_2^2 = \sum_{i=1}^{m}\|c_i\|_2^2\sum_{j=1}^{o}\|d_j\|_2^2 = \|C\|_F^2\|D\|_F^2 \ ,$$

where $\|\cdot\|_2$ denotes the Euclidean norm, and c_i, d_i denote the i-th column of the matrices C and D respectively. □

Lemma 2. *Absolute value of any element of a matrix is always less than or equal to the Frobenius norm of the matrix. In particular, for a matrix* $A^k = (a_{i,j}^{(k)})_{n\times n}$*, where* $A \in \mathbb{K}^{n\times n}$*, for every position* (i,j) *it holds* $|a_{i,j}^{(k)}| \leq \|A^k\|_F \leq \|A\|_F^k$*.*

Proof. Obviously, for an arbitrary element of the matrix $X = (x_{i,j})_{n\times m}$ we can write

$$|x_{i,j}| \leq \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}|x_{i,j}|^2} = \|X\|_F \ .$$

It follows

$$|a_{i,j}^{(k)}| \leq \|A^k\|_F \leq \|A\|_F^k \ ,$$

where second inequality follows from Lemma 1.4. □

Corollary 1. Let us have $A^k = (a_{i,j}^{(k)})_{n\times n}$. Then the series $\sum_{k=0}^{\infty}\frac{b^k}{k!}a_{i,j}^{(k)}$ converges absolutely for any $b \in \mathbb{K}$.

Proof. By Lemma 2, for any $N \in \mathbb{N}$, we have

$$\sum_{k=0}^{N}\left|\frac{b^k}{k!}a_{i,j}^{(k)}\right| \leq \sum_{k=0}^{N}\frac{|b|^k}{k!}\left|a_{i,j}^{(k)}\right| \leq \sum_{k=0}^{N}\frac{|b|^k}{k!}\|A\|_F^k = \sum_{k=0}^{N}\frac{\|bA\|_F^k}{k}$$

Then

$$\sum_{k=0}^{\infty}\left|\frac{b^k}{k!}a_{i,j}^{(k)}\right| = \lim_{N\to\infty}\sum_{k=0}^{N}\left|\frac{b^k}{k!}a_{i,j}^{(k)}\right| \leq \lim_{N\to\infty}\sum_{k=0}^{N}\frac{\|bA\|_F^k}{k} = \sum_{k=0}^{\infty}\frac{\|bA\|_F^k}{k} = e^{\|bA\|_F}$$

□

Definition. *A matrix sequence $\{A_k\}_{k=0}^{\infty}$ of $n \times m$ matrices is said to* **converge** *to $n \times m$ matrix A, denoted by $A_k \longrightarrow A$, if*

$$\forall \varepsilon \in \mathbb{R}, \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \in \mathbb{N}, n \geq n_0 : ||A_n - A||_F < \varepsilon \ .$$

Lemma 3. *A matrix sequence $\{A_k = (a_{i,j}^{(k)})_{n \times m}\}_{k=0}^{\infty}$ converges to a matrix $A = (a_{i,j})_{n \times m}$ if and only if it converges elementwise, in other words*

$$\forall i \in \{1, \ldots, n\} \quad \forall j \in \{1, \ldots, m\} : a_{i,j}^{(k)} \xrightarrow{k \to \infty} a_{i,j} \ .$$

Proof. Let $A_k \to A$. Then we can for any $\varepsilon > 0$ find such n_0 that $\|A_n - A\|_F < \varepsilon$ for every $n \geq n_0$. Using Lemma 2 we can write

$$|a_{i,j}^{(n)} - a_{i,j}| \leq \|A_n - A\|_F < \varepsilon \ .$$

It follows that $\{A_k\}_{k=0}^{\infty}$ converges to A elementwise.

Conversely, let ε be a positive real number. For every position (i, j) we find such $n_{i,j}$ that

$$\forall n \geq n_{i,j} : |a_{i,j}^{(n)} - a_{i,j}| < \frac{\varepsilon}{\sqrt{nm}} \ .$$

We put $N_0 = \min\{n_{i,j}\}$. Now $\forall n \in \mathbb{N}, n \geq N_0$ we have

$$||A_n - A||_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} |a_{i,j}^{(n)} - a_{i,j}|^2} < \sqrt{nm \frac{\varepsilon^2}{nm}} = \varepsilon \ .$$

□

Claim 4. *The definition of the matrix exponential makes sense, that is, the matrix series $\sum_{k=0}^{\infty} \frac{1}{k!} X^k$ converges for any matrix X.*

Proof. Let $X^k = (x_{i,j}^{(k)})_{n \times n}$. By Corollary 1 every element of matrix $\sum_{k=0}^{\infty} \frac{1}{k!} X^k = \left(\sum_{k=0}^{\infty} \frac{1}{k!} x_{i,j}^{(k)}\right)_{n \times n}$ converges absolutely. Therefore the matrix series converges elementwise to some matrix Y (we denote this matrix by e^X). □

Lemma 5. *Let $\{A_k\}_{k=0}^{\infty}$ be a matrix sequence, where $A_k \in \mathbb{K}^{n \times m}$, $B \in \mathbb{K}^{r \times n}$ and $C \in \mathbb{K}^{m \times s}$. If $\sum_{k=0}^{\infty} A_k$ converges, then also $\sum_{k=0}^{\infty} B A_k C$ converges, and the following equality holds*

$$B \left(\sum_{k=0}^{\infty} A_k \right) C = \sum_{k=0}^{\infty} B A_k C \ .$$

Proof. We know that for $N \in \mathbb{N}$ it is true

$$\sum_{k=0}^{N} B A_k C = B \left(\sum_{k=0}^{N} A_k \right) C \ .$$

We want to now show that the left hand side converges to $B \left(\sum_{k=0}^{\infty} A_k\right) C$. Let $\varepsilon_1 \in \mathbb{R}^+$ be fixed. Since the series $\sum_{k=0}^{\infty} A^k$ converges, we can find N_0 such that for every $N \in \mathbb{N}, N \geq N_0$ it holds

$$\left\| \sum_{k=0}^{\infty} A_k - \sum_{l=0}^{N} A_l \right\| < \varepsilon_1 \ .$$

Then

$$\left\| B\left(\sum_{k=0}^{\infty} A_k\right) C - \sum_{l=0}^{N} BA_lC \right\|_F = \left\| B\left(\sum_{k=0}^{\infty} A_k\right) C - B\left(\sum_{l=0}^{N} A_l\right) C \right\|_F =$$
$$= \left\| B\left(\sum_{k=0}^{\infty} A_k - \sum_{l=0}^{N} A_l\right) C \right\|_F = \|B\|_F \left\| \sum_{k=0}^{\infty} A_k - \sum_{l=0}^{N} A_l \right\|_F \|C\|_F < \|B\|_F \|C\|_F \varepsilon_1 \ .$$

That concludes the proof that $\sum_{k=0}^{\infty} BA_kC$ converges to $B\left(\sum_{k=0}^{\infty} A_k\right)C$. □

Definition. *Let us have a matrix function $X(t)\colon \mathbb{R} \to \mathbb{K}^{n\times m}$. Then the derivative of the function is*

$$\frac{d}{dt}X(t) = \left(\frac{d}{dt}x_{i,j}(t)\right)_{n\times m} = \left(\dot{x}_{i,j}(t)\right)_{n\times m} \ .$$

Lemma 6. *Let A, B and X be real or complex $n \times n$ matrices. Then*

1. *If $AB = BA$, then $e^A B = Be^A$,*
2. *If R is invertible, then $e^{R^{-1}XR} = R^{-1}e^X R$,*
3. *$\frac{d}{dt}e^{tX} = Xe^{tX}$, for $t \in \mathbb{R}$,*
4. *If $AB = BA$, then $e^{A+B} = e^A e^B$.*

Proof. 1. Because of the convergence of the matrix exponential, we can use Lemma 5 and we get

$$e^A B = \sum_{k=0}^{\infty} \frac{1}{k!}A^k B \overset{AB=BA}{=\!=\!=} \sum_{k=0}^{\infty} \frac{1}{k!}BA^k = B\sum_{k=0}^{\infty} \frac{1}{k!}A^k = Be^A \ .$$

2. Following from Lemma 5, we have

$$e^{R^{-1}XR} = \sum_{k=0}^{\infty} \frac{1}{k!}(R^{-1}XR)^k = \sum_{k=0}^{\infty} \frac{1}{k!}R^{-1}X^kR = R^{-1}\left(\sum_{k=0}^{\infty} \frac{1}{k!}X^k\right)R = R^{-1}e^X R \ .$$

3. The elements of the matrix $e^{tX} = \sum_{k=0}^{\infty} \frac{t^k}{k!}X^k = (e_{i,j}(t))_{n\times n}$ are equal to

$$e_{i,j}(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!}a_{i,j}^{(k)} \ ,$$

where $X^k = (a_{i,j}^{(k)})_{n\times n}$. By Corollary 1 the series $\sum_{k=0}^{\infty} \frac{t^k}{k!}a_{i,j}^{(k)}$ is absolutely convergent for every $t \in \mathbb{K}$. We can now differentiate the individual elements (see Pick et al. [2019], Věta 8.2.2.)

$$\frac{d}{dt}e_{i,j}(t) = \frac{d}{dt}\sum_{k=0}^{\infty} \frac{t^k}{k!}a_{i,j}^{(k)} = \sum_{k=1}^{\infty} \frac{t^{k-1}}{(k-1)!}a_{i,j}^{(k)} = \sum_{k=0}^{\infty} \frac{t^k}{k!}a_{i,j}^{(k+1)} \ .$$

Using Lemma 5 we get the desired result

$$\frac{d}{dt}e^{tX} = \left(\frac{d}{dt}e_{i,j}(t)\right)_{n\times n} = \left(\sum_{k=0}^{\infty} \frac{t^k}{k!}a_{i,j}^{(k+1)}\right)_{n\times n} = \sum_{k=0}^{\infty} \frac{t^k}{k!}X^{k+1} = X\sum_{k=0}^{\infty} \frac{t^k}{k!}X^k = Xe^{tX} \ .$$

4. Let us denote elements of matrix $A^kB^l = (\alpha_{i,j}^{(k,l)})_{n\times n}$. Then

$$\begin{aligned} e^Ae^B &= \sum_{k=0}^{\infty}\frac{1}{k!}A^k \cdot \sum_{l=0}^{\infty}\frac{1}{l!}B^l = \sum_{k=0}^{\infty}\left(\frac{1}{k!}A^k\sum_{l=0}^{\infty}\frac{1}{l!}B^l\right) = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{1}{k!l!}A^kB^l \\ &= \left(\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{1}{k!l!}\alpha_{i,j}^{(k,l)}\right)_{n\times n} = \left(\sum_{k=0}^{\infty}\sum_{l=0}^{k}\frac{1}{l!(k-l)!}\alpha^{(l,k-l)}\right)_{n\times n} \\ &= \sum_{k=0}^{\infty}\sum_{l=0}^{k}\frac{1}{l!(k-l)!}A^lB^{k-l} = \sum_{k=0}^{\infty}\sum_{l=0}^{k}\binom{k}{l}\frac{1}{k!}A^lB^{k-l} \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}(A+B)^k = e^{A+B}\ . \end{aligned}$$

The second and the third equalities hold by Lemma 5, and the penultimate equality holds by the assumption $AB = BA$. The crucial point is the fifth equality in which we reorder the elements of the series as depicted in Figure 1.1. The equality holds as long as the original series is absolutely

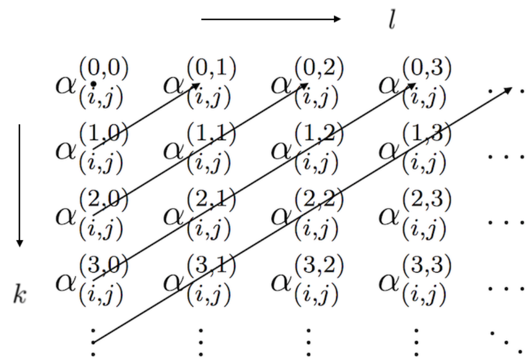

Figure 1.1

convergent. This follows from the inequality

$$\sum_{k=0}^{N}\sum_{l=0}^{M}\left|\frac{1}{k!l!}\alpha_{i,j}^{(k,l)}\right| \le \sum_{k=0}^{N}\frac{\|A\|_F^k}{k!}\sum_{l=0}^{M}\frac{\|B\|_F^l}{l!} = \left(\sum_{k=0}^{N}\frac{\|A\|_F^k}{k!}\right)\cdot\left(\sum_{l=0}^{M}\frac{\|B\|_F^l}{l!}\right)\ ,$$

and the fact that the expression on the rightmost side is bounded by $e^{\|A\|_F}e^{\|B\|_F}$ for $M, N \to \infty$.

□

Lemma 7. *For any $\alpha \in \mathbb{K}$ we have $e^{\alpha I} = e^{\alpha}I$.*

Proof. Follows straight from the definition of the matrix exponential. □

Now, using properties in Lemma 6, we can see that $\dot{x}(t) = Ax(t)$ is solved by $x(t) = e^{tA}x(0)$. The solution is unique which follows from advanced calculus. Let us now discuss under what circumstances does the state $x(t)$ converge to o for $t \to \infty$.

Let A be a real or complex matrix. Then there is a regular matrix $R \in \mathbb{K}^{n\times n}$ such that the matrix

$$J = R^{-1}AR$$

is in the Jordan normal form. By substituting $x(t) = Ry(t)$, which is equivalent with changing the basis of our system, we get

$$\begin{aligned} R\dot{y}(t) &= ARy(t) \\ \dot{y}(t) &= R^{-1}ARy(t) \\ \dot{y}(t) &= Jy(t) \end{aligned}$$

and therefore the unique solution is

$$y(t) = e^{tJ}y(0) \ ,$$

where $y(0) = R^{-1}x(0)$. It is sufficient to show that $y(t)$ converges to o, because since R is an invertible matrix, $x(t)$ converges to o if and only if $y(t)$ converges to o.

We know that every Jordan block $J_{\lambda,n}$ in the matrix J can be decomposed as $J_{\lambda,n} = \lambda I_n + N_n$, $n \in \mathbb{N}$ where N_n is $n \times n$ nilpotent matrix satisfying $n_{i,j} = \delta_{i,j-1}$. For example, in case of $n = 4$ we have

$$N_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, (N_4)^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, (N_4)^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

It is also true that $(N_n)^k_{i,j} = \delta_{i,j-k}$ and $(N_n)^n = O_{n\times n}$, since every right multiplication by matrix N shifts the multiplied matrix's columns to the right by one column, that is, it maps matrix $(v_1, \ldots, v_n)$ onto $(o, v_1, \ldots, v_{n-1})$.

By using Lemma 6, we now for each Jordan block $J_{\lambda,n}$ have

$$e^{tJ_{\lambda,n}} = e^{t(\lambda I + N)} = e^{t\lambda I}e^{tN} = e^{\lambda t}e^{tN} \ .$$

Let $\lambda = a + ib$ where $a,b \in \mathbb{R}$. Then we have

$$e^{tJ_{\lambda,n}} = e^{at}e^{ibt}e^{tN} \ .$$

We know that $|e^{ibt}| = 1$ and that

$$e^{tN} = \sum_{k=0}^{\infty} \frac{t^k}{k!}N^k = \sum_{k=0}^{n-1} \frac{t^k}{k!}N^k$$

since $(N_n)^n = O_{n\times n}$. Therefore, we can see that every element of matrix e^{tN} is a polynomial in t of degree less than n. It follows that $e^{tJ_{\lambda,n}}$ approaches $O_{n\times n}$ in infinity if

$$\lim_{t\to\infty} e^{at}t^{n-1} = 0 \ .$$

This holds for any $n \in \mathbb{N}$ as long as $a < 0$.

Because any block diagonal matrix to the power of any natural number preserves its block form, we can write

$$J = \begin{pmatrix} J_{\lambda_1,n_1} & 0 & \cdots & 0 \\ 0 & J_{\lambda_2,n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{\lambda_r,n_r} \end{pmatrix}, \quad e^J = \begin{pmatrix} e^{J_{\lambda_1,n_1}} & 0 & \cdots & 0 \\ 0 & e^{J_{\lambda_2,n_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{J_{\lambda_r,n_r}} \end{pmatrix},$$

where zeroes in the matrices represent zero matrices of appropriate sizes. Therefore, since $y(0)$ is a constant vector, we see that $y(t) = e^{tJ}y(0)$ converges to o if (and only if, because of the uniqueness of the solution) all the eigenvalues λ_i of the matrix A have negative real parts. As the last step, we calculate $x(t) = Ry(t)$ and $x(0) = Ry(0)$. We formulate this result into a theorem.

Theorem 8. *The system $\dot{x} = Ax(t)$ is stable if and only if all eigenvalues of the matrix A have negative real parts.*

Example. Consider higher order differential equation

$$x^{(n)}(t) + a_1 x^{(n-1)}(t) + \ldots + a_{n-1}x'(t) + a_n x(t) = 0 \ ,$$

where $x(t)\colon \mathbb{C} \to \mathbb{C}$. This equation can be solved as a system of linear differential equations of first order $\dot{z}(t) = Az(t)$ by choosing fundamental matrix A and state vector $z(t)$ as follows

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix}, z(t) = \begin{pmatrix} x(t) \\ x'(t) \\ \vdots \\ x^{(n-1)}(t) \end{pmatrix} .$$

1.1.2 Linear System With Control

Definition. *A* **continuous dynamical linear system with control u** *is a system of linear differential equations of first order with constant coefficients in the form*

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

where $x(t) \in \mathbb{K}^n$ is a **state vector** *(shortly* **state***) of the system, $A \in \mathbb{K}^{n\times n}$ is a* **fundamental matrix** *of the system, $B \in \mathbb{K}^{n\times m}$ is a* **control matrix** *of the system and $u(t) \in \mathbb{K}^m$ is a* **control vector** *of the system. The* **initial condition** *of the system is the state $x(0)$.*

We will call this system shortly (A, B) system.

In a general case, this is called an **open-loop control** system because the control is not dependent on previous state of the system.

TODO

We can imagine this system as follows. The first part of the right side, $Ax(t)$, of the equation $\dot{x}(t) = Ax(t) + Bu(t)$ can be thought of as the model of machine or event that we want to control and the second part, $Bu(t)$, as our control mechanism. The B matrix is our "control board" and the control vector $u(t)$ is us deciding, which levers and buttons we want to push.

Of course, if we want this system to be self-regulating, we cannot input our own values into $u(t)$, and therefore it has to be calculated from the current state of our system.

END TODO

Definition. *Let us have linear differential system with control $u(t)$ defined as*

$$u(t) = Fx(t) \ ,$$

where $F \in \mathbb{C}^{m\times n}$ *is a* **feedback matrix**. *This system is then called a* **closed-loop control** *system or a* **linear feedback control** *system.*

Typically, we require a feedback control system to stabilize itself back into its stable state after some disturbances.

Definition. *We say that the system* (A, B) *is* **stable**, *if it converges to the null vector.*

The feedback control system can be expressed as a first order linear differential system

$$\dot{x}(t) = Ax(t) + BFx(t) = (A + BF)x(t) \ .$$

As discussed in the first section, we now know that the system converges to o if all of the eigenvalues of matrix $A + BF$ have negative real parts.

Therefore, the system can stabilize itself if we find such matrix $F \in \mathbb{C}^{n\times n}$ that all the eigenvalues of the matrix $A + BF$ have negative real parts. This requirement can be expressed through characteristic polynomial of the matrix $A+BF$, since roots of the characteristic polynomial of a matrix are precisely the eigenvalues of the matrix.

Definition. *Let* A *be a* $n \times n$ *matrix. Then the* **characteristic polynomial** *of* A*, denoted by* χ_A*, is defined as*

$$\chi_A(s) = \det(sI_n - A) \ .$$

Through our observations we got to a conclusion, that we need to find a feedback matrix F such that the characteristic polynomial of the matrix $A+BF$ is

$$\chi_{A+BF} = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n) \ ,$$

where all its roots $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ have negative real parts. This leads to an important definition.

Definition. *Let* $\mathbb{K}$ *be a field and let* $A \in \mathbb{K}^{n\times n}$*,* $B \in \mathbb{K}^{n\times m}$*,* $n, m \in \mathbb{N}$*. We say that polynomial* χ *is* **assignable** *for the pair* (A, B) *if there exists such matrix* $F \in \mathbb{K}^{m\times n}$ *that*

$$\chi_{A+BF} = \chi \ .$$

The pole shifting theorem states, that if A and B are "sensible" in a sense that we will discuss in the next section, then an arbitrary monic polynomial χ of degree n can be assigned to the pair (A, B). It also claims that it is immaterial over what field A and B are.

1.2 Controllable pairs

In this section we will establish the notion of controllability. We will first explain this concept for *discrete-time systems* and then we will show that the requirement for controllability of *continuous-time systems* is the same as the one for *discrete-time systems*.

1.2.1 Discrete-time systems

Let us have a continuous dynamical system $\dot{x}(t) = A_1 x(t)$, where A_1 is a real or complex square matrix. We *discretize* the time, that is, instead of using continuous real time values of $x(t)$ and $\dot{x}(t)$, we are interested in these values only at discrete *sampling times* $0, \delta, 2\delta, \ldots, k\delta, \ldots$ where $\delta \in \mathbb{R}^+$. We will denote the states at each sampling time as follows

$$x_k = x(k\delta)\ , k \in \mathbb{N}_0\ .$$

The solution of this system is precisely $x(t) = e^{tA_1}x(0)$, as discussed in previous sections. For some $k \in \mathbb{N}$ we get $x_k = x(k\delta) = e^{k\delta A_1}x(0)$ and $x_{k+1} = e^{(k+1)\delta A_1}x(0) = e^{\delta A_1 + k\delta A_1}x(0)$ which by Lemma 6.5 equals $x_{k+1} = e^{\delta A_1 + k\delta A_1}x(0) = e^{\delta A_1}e^{k\delta A_1}x(0) = e^{\delta A_1}x_k$.

Using Lemma 6 we obtain

$$\begin{aligned} x_{k+1} &= e^{(k+1)\delta A_1}x(0) \\ &= e^{\delta A_1 + k\delta A_1}x(0) \\ &= e^{\delta A_1}e^{k\delta A_1}x(0) \\ &= e^{\delta A_1}x_k \\ &= Ax_k \end{aligned}$$

by choosing $A = e^{\delta A_1}$. We see that we can calculate the next value of x from its previous value. We will now define this system. The definition holds for any field $\mathbb{K}$.

Definition. *A* **discrete dynamical linear system** *is a system of equations*

$$x_{k+1} = Ax_k,\ \ k \in \mathbb{N}_0\ ,$$

where $x_k \in \mathbb{K}^n$ is a **state vector** *(shortly* **state***) of the system, the matrix $A \in \mathbb{K}^{n\times n}$ is a* **fundamental matrix** *of the system. The* **initial condition** *of the system is the state $x(0)$.*

Definition. *A* **discrete dynamical linear system with control u** *is a system of equations*

$$x_{k+1} = Ax_k + Bu_k,\ \ k \in \mathbb{N}_0\ ,$$

where $x_k \in \mathbb{K}^n$ is a **state vector** *(shortly* **state***) of the system, $A \in \mathbb{K}^{n\times n}$ is a fundamental matrix, $B \in \mathbb{K}^{n\times m}$ is a control matrix, $u_k \in \mathbb{K}^m$ is a control vector. The* **initial condition** *of the system is the state x_0.*

We will call this system **discrete** (A, B) *system.*

Definition. *We say that a state x can be* **reached** *in time $k \in \mathbb{N}_0$ if there exists such a sequence of control vectors $u_0, u_1, \ldots, u_k$ that for starting condition $x_0 = o$ we get $x = x_k$, after k iterations of $x_{l+1} = Ax_l + Bx_l$, where $l \in \{0, 1, \ldots, k-1\}$.*

States that we can reach in set number of iterations in a open-loop control discrete-time systems can be derived as follows. From state x_k and control vector u_k is the next state x_{k+1} computed by equation

$$x_{k+1} = Ax_k + Bu_k$$

The starting condition is $x_0 = \mathbf{o}$ and we can choose an arbitrary u_k. Then, for $k = 0$ we have

$$x_1 = Ax_0 + Bu_0 = Bu_0 \in \mathrm{Im}B.$$

For $k = 1$ we get

$$x_2 = Ax_1 + Bu_1 = ABu_0 + Bu_1 \in (AB|B)$$

It is clear, that

$$x_k \in \mathrm{Im}(A^{k-1}B|\cdots|AB|B)$$

We can observe that $\mathrm{Im}(B|AB|\cdots|A^kB) \subseteq \mathrm{Im}(B|AB|\cdots|A^{k+1}B)$. By Cayley-Hamilton theorem we know that $\chi_A(A) = O_{n\times n}$. That means that A^n can be expressed as linear combination of matrices $\{I, A, \ldots, A^{n-1}\}$ or equivalently that A^nB can be expressed as linear combination of matrices $\{B, AB, \ldots, A^{n-1}B\}$. We now see that $\mathrm{Im}(B|AB|\cdots|A^kB) \supseteq \mathrm{Im}(B|AB|\cdots|A^{k+1}B)$ holds. It follows

$$\mathrm{Im}(B|AB|\cdots|A^{n-1}B) = \mathrm{Im}(B|AB|\cdots|A^{n-1}B|A^nB) \tag{1.1}$$

Therefore, all the states we could ever reach are already in space

$$\mathrm{Im}(B|AB|\cdots|A^{n-1}B)$$

Definition. *Let $\mathbb{K}$ be a field and let $A \in \mathbb{K}^{n\times n}$, $B \in \mathbb{K}^{n\times m}$, $n, m \in \mathbb{N}$. We define* **reachable space** *$\mathcal{R}(A, B)$ of the pair (A, B) as Im$(B|AB|\cdots|A^{n-1}B)$.*

We have seen that by left multiplying $\mathcal{R}(A, B)$ by A, we get the same subspace. This leads to an important property of some subspaces.

Definition. *Let V be a vector space, W be its subspace and f be a mapping from V to V. We call W an* **invariant subspace** *of f if $f(W) \subseteq W$.*

We also say that W is **f-invariant**. *When $f = f_A$ for some matrix A, we also shortly say that W is* **A-invariant**.

Remark. $\mathcal{R}(A, B)$ is a A-invariant subspace.

The maximum dimension of $\mathcal{R}(A, B)$ is, of course, n. This leads us to important property of pair (A, B), where we want to be able to get the system into any state in state space by controlling it with our control vector $u(t)$, i.e., choosing appropriate $u(t)$. Therefore, we desire that $\mathcal{R}(A, B) = \mathbb{K}^n$. The equivalent condition is $\dim\mathcal{R}(A, B) = n$.

Definition. *Let $\mathbb{K}$ be a field and let $A \in \mathbb{K}^{n\times n}$, $B \in \mathbb{K}^{n\times m}$, $n, m \in \mathbb{N}$. The pair (A, B) is* **controllable** *if $\dim\mathcal{R}(A, B) = n$.*

1.2.2 Continuous-time systems

Remark. In this section we assume, that $\mathbb{K}$ is a field of either real or complex numbers and that $A \in \mathbb{K}^{n\times n}$, $B \in \mathbb{K}^{n\times m}$.

We will now show that the condition for controllability of *discrete-time systems* also characterizes controllable *continuous-time systems*. For this we have to express solution of such system using matrices A^iB for $i \in \mathbb{N}_0$.

We utilize matrix exponential in solving system of inhomogeneous linear system $\dot{x}(t) = Ax(t) + Bu(t)$. By left multiplying it by e^{-tA} we get

$$e^{-tA}\dot{x}(t) - e^{-tA}Ax(t) = e^{-tA}Bu(t)$$
$$\frac{d}{dt}(e^{-tA}x(t)) = e^{-tA}Bu(t)$$

Note, we used $-AA = A(-A) \Rightarrow e^{-tA}A = Ae^{-tA}$ from Lemma 6. After integrating both sides with respect to t on interval (t_0, t_1) we have

$$[e^{-tA}x(t)]_{t_0}^{t_1} = \int_{t_0}^{t_1} e^{-tA}Bu(t)dt$$
$$e^{-t_1A}x(t_1) - e^{-t_0A}x(t_0) = \int_{t_0}^{t_1} e^{-tA}Bu(t)dt$$
$$x(t_1) = e^{(t_1-t_0)A}x(t_0) + \int_{t_0}^{t_1} e^{(t_1-t)A}Bu(t)dt$$

Now it is clear that in system where $x(0) = o$ can every state in time $t \in \mathbb{R}^+$ be expressed as

$$x(t) = \int_0^t e^{(t-s)A}Bu(s)ds$$

Definition. *We say that a state $x \in \mathbb{K}^n$ can be* **reached in the time** t*, if there exists a control $u(x)\colon [0,t] \to \mathbb{K}^m$ such that*

$$x = \int_0^t e^{(t-s)A}Bu(s)ds\ .$$

The set of all states that can be reached in the time t is denoted by $\mathcal{R}^t$. The set $\mathcal{R} = \cup_{t\in\mathbb{R}^+}\mathcal{R}^t$ of all states that can be reached, is called a **reachable space**.

Definition. *A n-dimensional continuous-time linear system is* **controllable**, *if $\mathcal{R} = \mathbb{K}^n$.*

Theorem 9. *The n-dimensional continuous-time linear system is controllable if and only if $dim\mathcal{R}(A,B) = n$.*

Proof. From discussion above we have

$$x(t) = \int_0^t e^{(t-s)A}Bu(s)ds = \int_0^t \sum_{k=0}^{\infty} \frac{(t-s)^k}{k!}A^kBu(s)ds$$

The n-dimensional vector $w^{(k)}(s) = A^kBu(s)$ has elements

$$w_i^{(k)}(s) = \sum_{j=1}^{m} \alpha_{i,j}^{(k)}u_j(s)$$

where $\alpha_{i,j}^{(k)}$ is the element of the matrix A^kB on the position (i,j). Therefore, the elements of $x(t)$ are

$$x_i(t) = \int_0^t \sum_{k=0}^{\infty} \frac{(t-s)^k}{k!}w_i^{(k)}ds = \int_0^t \sum_{k=0}^{\infty}\sum_{j=1}^{m} \frac{(t-s)^k}{k!}\alpha_{i,j}^{(k)}u_j(s)ds$$

Now, in order to be able to modify this expression, we will prove that the series $\sum_{k=0}^{\infty} \frac{(t-s)^k}{k!} \alpha_{i,j}^{(k)} u_j(s)$ is convergent for every position (i,j). This follows from

$$\left| \sum_{k=0}^{\infty} \frac{(t-s)^k}{k!} \alpha_{i,j}^{(k)} u_j(s) \right| \leq \sum_{k=0}^{\infty} \frac{|t-s|^k}{k!} |\alpha_{i,j}^{(k)}| |u_j(s)| \leq \sum_{k=0}^{\infty} \frac{\|(t-s)A\|_F^k}{k!} \|B\|_F |u_j(s)| \leq |u_j(s)| \|B\|_F e^{\|(t-s)A\|} ,$$

where the second inequality holds by Corollary 1, and the third inequality by the fact that we can factor constants $\|B\|_F$ and $u_j(s)$ out of convergent number series $e^{\|(t-s)A\|}$. Because of the absolute convergence, we can now swap the integral and the series:

$$\begin{aligned}
x_i(t) &= \int_0^t \sum_{k=0}^{\infty} \sum_{j=1}^{m} \frac{(t-s)^k}{k!} \alpha_{i,j}^{(k)} u_j(s) ds = \int_0^t \sum_{j=1}^{m} \sum_{k=0}^{\infty} \frac{(t-s)^k}{k!} \alpha_{i,j}^{(k)} u_j(s) ds \\
&= \sum_{j=1}^{m} \int_0^t \sum_{k=0}^{\infty} \frac{(t-s)^k}{k!} \alpha_{i,j}^{(k)} u_j(s) ds = \sum_{j=1}^{m} \sum_{k=0}^{\infty} \int_0^t \frac{(t-s)^k}{k!} \alpha_{i,j}^{(k)} u_j(s) ds \\
&= \sum_{j=1}^{m} \sum_{k=0}^{\infty} \alpha_{i,j}^{(k)} \int_0^t \frac{(t-s)^k}{k!} u_j(s) ds = \sum_{k=0}^{\infty} \sum_{j=1}^{m} \alpha_{i,j}^{(k)} \int_0^t \frac{(t-s)^k}{k!} u_j(s) ds \\
&= \sum_{k=0}^{\infty} \sum_{j=1}^{m} \alpha_{i,j}^{(k)} v_i^{(k)}(t) ,
\end{aligned}$$

where $v^{(k)}(t) = \int_0^t \frac{(t-s)^k}{k!} u(s) ds$ is a vector of length m. Therefore, we have

$$x(t) = \sum_{k=0}^{\infty} A^k B v_k(t) = \sum_{k=0}^{\infty} A^k B \int_0^t \frac{(t-s)^k}{k!} u(s) ds$$

Now it its clear that

$$x(t) \in \mathrm{Im}(B|AB|\dots|A^k B|\dots) = \mathrm{Im}(B|AB|\dots|A^{n-1}B) = \mathcal{R}(A,B)$$

The first equality follows from the equality (1.1).

If the system is controllable then $x(t)$ can be equal to any of the vectors of an arbitrary basis of $\mathbb{K}^n$. Therefore, we know that n linearly independent vectors belong into $\mathcal{R}(A,B)$, and naturally it follows $\dim\mathcal{R}(A,B) = n$.

Conversely, if dimension of reachable space is equal to n we then have $x(t) \in \mathcal{R}(A,B) = \mathbb{C}^n$, therefore the system is controllable. □

1.2.3 Decomposition theorem

Lemma 10. *Let W be an invariant subspace of linear mapping $f\colon V \to V$. Then there exists a basis C of V such that*

$$[f]_C^C = \begin{pmatrix} F_1 & F_2 \\ 0 & F_3 \end{pmatrix}$$

where F_1 is $r \times r$, $r = dimW$.

Proof. Let $(w_1, \dots, w_r)$ be an arbitrary basis of subspace W. We complete this sequence into basis C of V with vectors $v_1, \dots, v_{n-r}$ where $n = \dim V$, thus $C = (w_1, \dots, w_r, v_1, \dots, v_{n-r})$. We know that

$$[f]_C^C = ([f(w_1)]_C, \dots, [f(w_r)]_C, [f(v_1)]_C, \dots, [f(v_{n-r})]_C)$$

Since W is an A-invariant subspace, it holds that $f(w_i) \in W$ and therefore, because of our choice of the basis C, the matrix $[f]_C^C$ is of the desired form. □

If (A, B) is not controllable, then there exists subspace of our state space that is not affected by our input. This can be shown using following theorem.

Theorem 11. *Let (A, B) represent a dynamical system and let $dim\mathcal{R}(A, B) = r \leq n$. Then there exists invertible $n \times n$ matrix T over $\mathbb{K}$ such that the matrices $\widetilde{A} := T^{-1}AT$ and $\widetilde{B} := T^{-1}B$ have the block structure*

$$\widetilde{A} = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix} \qquad \widetilde{B} = \begin{pmatrix} B_1 \\ 0 \end{pmatrix} \tag{1.2}$$

where A_1 is $r \times r$ and B_1 is $r \times m$.

Proof. We know that $\mathcal{R}(A, B)$ is an A-invariant subspace (Remark 1.2.1). Using Lemma 10 on the matrix mapping f_A we get a basis C for which it holds that

$$[f_A]_C^C = [\text{id}]_C^K [f_A]_K^K [\text{id}]_K^C = [\text{id}]_C^K A [\text{id}]_K^C$$

is in a block triangular form. By putting $T = [\text{id}]_K^C = C$ we get that $\widetilde{A} = [f_A]_C^C$ is now in the desired form.

Consider now matrix mapping f_B. We have

$$\widetilde{B} = TB = [\text{id}]_C^{K_n} [f_B]_{K_n}^{K_m} = [f_B]_C^{K_m} = ([f_B(e_1)]_C, \ldots, [f_B(e_m)]_C)$$

Since $f_B(e_i)$ is i-th column of matrix B and trivially by definition of a reachable space it holds $\text{Im}(B) \subseteq \mathcal{R}(A, B)$, we see that $\widetilde{B}$ is in the requested form. $\square$

We achieved the new form of matrices A and B by changing the basis of our state space. We now define the relation between (A, B) and $(\widetilde{A}, \widetilde{B})$.

Definition. *Let (A, B) and $(\widetilde{A}, \widetilde{B})$ be pairs as in Theorem 11 above. Then (A, B)* **is similar to** *$(\widetilde{A}, \widetilde{B})$, denoted $(A, B) \sim (\widetilde{A}, \widetilde{B})$, if there exists an invertible matrix T for which it holds that*

$$\widetilde{A} = T^{-1}AT \quad \text{and} \quad \widetilde{B} = T^{-1}B$$

Lemma 12. *Let A and B be similar matrices, that is, there exists an invertible matrix R such that $A = R^{-1}BR$. Then $\chi_A = \chi_B$.*

Proof. We will use the properties of the matrix determinant.

$$\begin{aligned}
\chi_A &= \det(sI - A) = \det(sI - R^{-1}BR) \\
&= \det(sR^{-1}IR - R^{-1}BR) = \det(R^{-1}(sI - B)R) \\
&= (\det R)^{-1}\det(sI - B)\det R = \det(sI - B) \\
&= \chi_B
\end{aligned}$$

$\square$

Lemma 13. *If $(A, B) \sim (\widetilde{A}, \widetilde{B})$ the they can be assigned the same polynomials.*

Proof. Since a matrix similar to $A + BF$ is in the form $T^{-1}(A + BF)T = T^{-1}ATT^{-1}BFT = \widetilde{A} + \widetilde{B}\widetilde{F}$, where $\widetilde{F} = FT$ ans T is some invertible matrix, using Lemma 12 we can write

$$\chi_{A+BF} = \chi_{\widetilde{A}+\widetilde{B}\widetilde{F}}$$

Therefore, the same polynomial can be assigned to pair $(\widetilde{A}, \widetilde{B})$ using matrix $\widetilde{F}$. $\square$

We can interpret the decomposition as follows. Consider our system $\dot{x}(t) = Ax(t) + Bu(t)$. By changing the basis by putting $x(t) = Ty(t)$ we get

$$T\dot{y}(t) = ATy(t) + Bu(t)$$

which we can write as

$$\dot{y}(t) = T^{-1}ATy(t) + T^{-1}Bu(t) = \widetilde{A}y(t) + \widetilde{B}u(t)$$

Which gives us

$$\begin{aligned} \dot{y}_1(t) &= A_1y_1(t) + A_2y_2(t) + B_1u_1(t) \\ \dot{y}_2(t) &= \qquad\qquad\; A_3y_2(t) \end{aligned}$$

where the vector $y_1(t)$ is composed of the first r elements of the vector $y(t)$, $y_2(t)$ is composed of the last $n-r$ elements of $y(t)$ and the vector $u_1(t)$ is composed of the first r elements of the vector $u(t)$. Since $\dot{y}_2(t)$ does not depend on the control vector $u(t)$, we see that we cannot change the last $n-r$ coordinates of $y(t)$ by changing the vector $u(t)$.

It is also true that (A_1, B_1) from Theorem 11 is a controllable pair which we will state in a lemma.

Lemma 14. *The pair (A_1, B_1) is controllable.*

Proof. We know that $\dim\mathcal{R}(A, B) = r$. We desire $\dim\mathcal{R}(A_1, B_1) = r$. We will show that $\mathcal{R}(\widetilde{A}, \widetilde{B}) = \mathcal{R}(A, B)$ and that each vector in $\mathcal{R}(\widetilde{A}, \widetilde{B})$ has its last $n-r$ elements equal to 0 and that $\mathcal{R}(\widetilde{A}, \widetilde{B})$ restricted on its first r coordinates is equal to $\mathcal{R}(A_1, B_1)$.

$$\begin{aligned} \mathcal{R}(\widetilde{A}, \widetilde{B}) &= \mathrm{Im}(\widetilde{A}^{n-1}\widetilde{B}|\cdots|\widetilde{A}\widetilde{B}|\widetilde{B}) \\ &= \mathrm{Im}((T^{-1}AT)^{n-1}T^{-1}B|\cdots|T^{-1}ATT^{-1}B|T^{-1}B) \\ &= \mathrm{Im}(T^{-1}A^{n-1}B|\cdots|T^{-1}AB|T^{-1}B) \\ &= \{(T^{-1}A^{n-1}B|\cdots|T^{-1}AB|T^{-1}B)\cdot v|v \in \mathbb{K}^{n\cdot m}\} \\ &= \{T^{-1}(A^{n-1}B|\cdots|AB|B)\cdot v|v \in \mathbb{K}^{n\cdot m}\} \\ &= T^{-1}\cdot\{(A^{n-1}B|\cdots|AB|B)\cdot v|v \in \mathbb{K}^{n\cdot m}\} \\ &= T^{-1}\cdot(\mathrm{Im}(A^{n-1}B|\cdots|AB|B)) \\ &= T^{-1}\cdot(\mathcal{R}(A, B)) \end{aligned}$$

where by $\cdot : \mathbb{K}^{n\times m} \times V \to W$ where V, W are vector spaces is defined as $A\cdot V = \{Av|v \in V\}$.

Since T is an invertible matrix, which preserves linear independence, we have

$$\dim\mathcal{R}(\widetilde{A}, \widetilde{B}) = \dim(T^{-1}\mathcal{R}(A, B)) = \dim(\mathcal{R}(A, B)) = r$$

Now let us focus on the structure of $\mathcal{R}(\widetilde{A}, \widetilde{B})$: We know that last $n-r$ rows of $\widetilde{B}$ are o. Also because of structure of $\widetilde{A}$ we have for an arbitrary matrix $X \in \mathbb{K}^{r\times m}$ that

$$\widetilde{A}\begin{pmatrix} X \\ 0 \end{pmatrix} = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}\begin{pmatrix} X \\ 0 \end{pmatrix} = \begin{pmatrix} A_1X \\ 0 \end{pmatrix}$$

where again are the last $n-r$ rows equal to o. Therefore we see that for any positive integer k we have

$$\widetilde{A}^k\widetilde{B} = \begin{pmatrix} A_1^k B_1 \\ 0 \end{pmatrix}, A_1^k B_1 \in \mathbb{K}^{r\times r}$$

It follows

$$\mathcal{R}(\widetilde{A}, \widetilde{B}) = \left(\begin{pmatrix} A_1^{n-1} B_1 \\ 0 \end{pmatrix} \middle| \cdots \middle| \begin{pmatrix} A_1 B_1 \\ 0 \end{pmatrix} \middle| \begin{pmatrix} B_1 \\ 0 \end{pmatrix} \right)$$

By Cayle-Hemilton theorem we therefore again have that the restriction to first r coordinates (those which are not 0) of $\mathcal{R}(\widetilde{A}, \widetilde{B})$ are equal to $\mathcal{R}(A_1, B_1)$. Finally, it follows that

$$\dim\mathcal{R}(A_1, B_1) = \dim\mathcal{R}(\widetilde{A}, \widetilde{B}) = \dim\mathcal{R}(A, B) = r$$

$\square$

Now we can see that the decomposition from Theorem 11 decomposes the matrix A into "controllable" and "uncontrollable" parts A_1 and A_3 respectively.

Corollary 2. For matrix A and its similar matrix $\widetilde{A}$ as in Theorem 11 it holds

$$\chi_A = \chi_{\widetilde{A}} = \chi_{A_1}\chi_{A_3}$$

Definition. *The characteristic polynomial of matrix A splits into* **controllable** *and* **uncontrollable parts** *with respect to pair (A, B) which we denote by χ_c and χ_u respectively. We define these polynomials as*

$$\chi_c = \chi_{A_1} \qquad \chi_u = \chi_{A_3}$$

In case $r = 0$ we put $\chi_c = 1$ and in case $r = n$ we put $\chi_u = 1$.

For this definition to be correct, we need to show that polynomials χ_{A_1} and χ_{A_3} are not dependent on the choice of a basis on $\mathcal{R}(A, B)$. Since $\chi_{A_3} = \chi_A/\chi_{A_1}$, it is enough to show that χ_{A_1} is independent of the choice.

Lemma 15. *χ_c is independent of the choice of basis on $\mathcal{R}(A, B)$.*

Proof. By definition we have $\chi_c = \chi_{A_1}$ where A_1 is some matrix for a specific decomposition of matrix A thanks to basis C used in a proof of Theorem 11. Consider different basis D which suffices the conditions from the said proof. Then we obtain a similar matrix $\widetilde{B} = [id]^K_D A [id]^D_K$. We denote the $r \times r$ matrix in the top left corner of $\widetilde{B}$ by B_1. We want to show $\chi_{A_1} = \chi_{B_1}$.

Let us have matrices

$$A' = \begin{pmatrix} A_1 & 0 \\ 0 & I_{n-r} \end{pmatrix} \qquad A'' = \begin{pmatrix} B_1 & 0 \\ 0 & I_{n-r} \end{pmatrix}$$

It holds

$$\chi_{A'}(s) = \chi_{A_1} \cdot s^n \quad \chi_{A''}(s) = \chi_{B_1} \cdot s^n$$

Therefore, it is sufficient to prove $\chi_{A'} = \chi_{A''}$ which according to Lemma 12 holds if A' and A'' are similar. Since $A' =$ $\square$

2. The Pole Shifting Theorem

Definition. *The* **controller form** *associated to the pair* (A, b) *is the pair*

$$A^\flat = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \end{pmatrix}, \quad b^\flat = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

where $s^n - \alpha_n s^{n-1} - \ldots - \alpha_2 s - \alpha_1$ *is the characteristic polynomial of* A.

Lemma 16. *The characteristic polynomial of* $A^\flat$ *is* $s^n - \alpha_n s^{n-1} - \ldots - \alpha_2 s - \alpha_1$.

Proof. Can be shown using simple properties of the matrix determinant. □

Lemma 17. *The pair* $(A^\flat, b^\flat)$ *is controllable.*

Proof. Because of the form of the matrix $A^\flat$ and $b^\flat$ is $(A^\flat)^k b^\flat$ equal to the last column of $(A^\flat)^k$ which creates vectors of form

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & \beta_{k-1} & \cdots & \beta_1 \end{pmatrix}^T$$

for some $\beta_1, \ldots, \beta_{k-1} \in \mathbb{K}$. Therefore $\mathcal{R}(A^\flat, b^\flat) = n$. □

Lemma 18. *Let* $\mathbb{K}$ *be a field and let* $A_1, A_2 \in \mathbb{K}^{n\times n}, b_1, b_2 \in \mathbb{K}^n$, *such that the pairs* $(A_1, b_1), (A_2, b_2)$ *are controllable. The pairs* $(A_1, b_1), (A_2, b_2)$ *are similar if and only if characteristic polynomials of* A_1 *and* A_2 *are the same.*

Proof. TODO □

Corollary 3. If the **single-input** $(m = 1)$ pair (A, b) is controllable, then it is similar to its controller form.

Proof. Follows from Lemmas 16, 17 and 18. □

Theorem 19. *Let* $\mathbb{K}$ *be a field. Let* $A \in \mathbb{K}^{n\times n}$, $B \in \mathbb{K}^{n\times m}$. *The assignable polynomials for pair* (A, B) *are precisely of the form*

$$\chi_{AB+F} = \chi\chi_u$$

where χ *is an arbitrary monic polynomial of degree* $r = dim\mathcal{R}(A, B)$ *and* χ_u *is the uncontrollable part of the assignable polynomial.*

In particular, the pair (A, B) *is controllable if and only if every nth degree monic polynomial can be assigned to it.*

Proof. By Theorem 11 and Lemma 13 we can assume that the pair (A, B) is in the same form as $(\widetilde{A}, \widetilde{B})$ in (1.2). Let us write $F = (F_1, F_2) \in \mathbb{K}^{m\times n}$, where $F_1 \in \mathbb{K}^{m\times r}, F_2 \in \mathbb{K}^{m\times(n-r)}$. Then

$$\begin{aligned} A + BF &= \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix} + \begin{pmatrix} B_1 \\ 0 \end{pmatrix} \begin{pmatrix} F_1 & F_2 \end{pmatrix} = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix} + \begin{pmatrix} B_1F_1 & B_1F_2 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} A_1 + B_1F_1 & A_2 + B_1F_2 \\ 0 & A_3 \end{pmatrix} \end{aligned}$$

It follows

$$\chi_{A+BF} = \chi_{A_1+B_1F_1}\chi_{A_3} = \chi_{A_1+B_1F_1}\chi_u$$

We see that any assignable polynomial has the desired factored form.

Conversely, we want to show that the first factor can be made arbitrary by a suitable choice of F_1. This does make sense only for $r > 0$, otherwise the assignable polynomial is equal to χ_u, which cannot be changed by modifying the matrix F. Assume that we are given a monic polynomial χ. If we find such a matrix F_1 that

$$\chi_{A_1+B_1F_1} = \chi$$

then by putting $F = (F_1, 0)$ we get the desired characteristic polynomial, that is, $\chi_{A+BF} = \chi\chi_u$. Since the pair (A_1, B_1) is controllable as shown in Lemma 14, it is sufficient only to prove that controllable systems can be assigned an arbitrary monic polynomial χ or respective degree. Therefore, from this point on, we assume that the pair (A, B) is controllable.

We will first prove the theorem for the case $m = 1$ and then we will express a general case as the case $m = 1$. That will conclude the proof.

Let $m = 1$. By Lemmas 13 and ???? we can consider the pair (A, b) to be in the controller form. For a vector

$$f = \begin{pmatrix} f_1 & f_2 & \dots & f_n \end{pmatrix}$$

we have

$$A + bf = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} f_1 & f_2 & \dots & f_n \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ \alpha_1 + f_1 & \alpha_2 + f_2 & \alpha_3 + f_3 & \dots & \alpha_n + f_n \end{pmatrix}.$$

One can see that for a given monic polynomial

$$\chi = s^n - \beta_n s^{n-1} - \dots - \beta_2 s - \beta_1$$

we can choose

$$f = \begin{pmatrix} \beta_1 - \alpha_1 & \beta_2 - \alpha_2 & \dots & \beta_n - \alpha_n \end{pmatrix},$$

and then it holds $\chi_{A+bf} = \chi$. We have shown that for the case $m = 1$, a controllable pair (A, b) can be assigned an arbitrary monic polynomial of degree n.

For a general case, where m is arbitrary, we choose any vector $v \in \mathbb{K}^m$ such that $Bv \neq o$, and let $b = Bv$. We will use the fact that a pair $(A + BF_1, b)$ can be assigned the same polynomials as a pair (A, B), because for any $f \in \mathbb{K}^{1\times n}$ it holds

$$A + BF_1 + bf = A + BF_1 + Bvf = A + B(F_1 + vf)$$

and therefore, for $F = F_1 + vf$ we have

$$\chi_{A+BF_1+bf} = \chi_{A+BF} \ .$$

Using the result for $m = 1$, the proof will be concluded by showing that the pair $(A + BF_1, b)$ is controllable.

Let us have an arbitrary sequence of linearly independent vectors $\{Bv = x_1, \ldots, x_k\}$, $x_i \in \mathbb{K}^n$, where

$$x_i = Ax_{i-1} + Bu_{i-1}, \ i \in \{1, \ldots, k\} \tag{2.1}$$

for some $u_i \in \mathbb{K}^m$, and $x_0 = o$. Consider k to be as large as possible. We denote $\mathcal{V} = \mathrm{Im}\{x_1, \ldots, x_k\}$. By maximality of k we have

$$x_{k+1} = Ax_k + Bu \in \mathcal{V} \tag{2.2}$$

for any $u \in \mathbb{K}^m$. Therefore, in particular for $u = o$, we get

$$Ax_k \in \mathcal{V} \ .$$

It follows by (2.2), that for any $u \in \mathbb{K}^m$ it holds

$$Bu \in \mathcal{V} - \{Ax_k\} \subseteq \mathcal{V} \ .$$

Thus, for the column space of the matrix B (denoted by $\mathcal{B}$) we have $\mathcal{B} \subseteq \mathcal{V}$. By equality (2.1) it is true

$$x_i - Ax_{i-1} \in \mathcal{B} \subseteq \mathcal{V} \ ,$$

for $i \in \{1, \ldots, k\}$. We can now write

$$-Ax_{i-1} \in \mathcal{V} - \{x_i\} \subseteq \mathcal{V} \ .$$

Since $\mathcal{V}$ is an vector space, also $Ax_{i-1} \in \mathcal{V}$. This means, that $\mathcal{V}$ is an A-invariant subspace containing $\mathcal{B}$. One can see that

$$\mathcal{V} \supseteq \mathrm{Im}(B|AB|A^2B|\ldots|A^kB|\ldots) = \mathrm{Im}(B|AB|A^2B|\ldots|A^kB) = \mathcal{R}(A, B) \ ,$$

where the second equality holds by discussion around the equality (1.1). Because of this relation, we observe that

$$\dim\mathcal{R}(A, B) \leq \dim\mathcal{V} = k \leq \dim\mathbb{K}^n = n \ .$$

Since the pair (A, B) is by the assumption controllable, we know that $\dim\mathcal{R}(A, B) = n$. Therefore, $n = k$.

Finally, we need to show that

$$\dim\mathcal{R}(A + BF_1, x_1) = n \ .$$

By letting

$$F_1x_i = u_i \ ,$$

which is equivalent with

$$F_1 \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & \dots & u_n \end{pmatrix} \ , \tag{2.3}$$

we achieve

$$(A + BF_1)x_i = Ax_i + BF_1x_i = Ax_i + Bu_i = x_{i+1} \ ,$$

and therefore

$$\mathcal{R}(A + BF_1, x_1) = \mathrm{Im}\{x_1, x_2, \dots, x_n\} \ . \tag{2.4}$$

We can define the matrix F_1 in this way, because the matrix $\begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}$ is a square matrix with linearly independent columns and therefore, its rows are also linearly independent. Using this fact, and the fact that i-th row of the matrix $\begin{pmatrix} u_1 & u_2 & \dots & u_n \end{pmatrix}$ is the result of linear combination of the rows of the matrix $\begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}$ with coefficients in i-th row of the matrix F_1, one can see that the definition of F_1 is valid.

Finally, the equality (2.4) implies, by linear independence of the vectors $x_1, \dots, x_n$, that $\dim \mathcal{R}(A + BF_1, x_1) = n$. We have shown that the pair $\mathcal{R}(A + BF_1, Bv)$ is controllable, and thus the proof is concluded. $\square$

Bibliography

L. Pick, S. Hencl, J. Spurný, and M. Zelený. Matematická analýza 1 (velmi předběžná verze). Preprint available at `http://www.karlin.mff.cuni.cz/~pick/`, April 2019.