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

Contents

Chapter 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. To understand this crucial theorem, we must first describe few basic concepts.

1.1.1 Systems of First Order Differential Equations

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

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

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* **starting condition** *of the system is the state $x(0)$.*

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

To express solution of this system in similarly compact matter we will establish a 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 given 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 be valid, we need to show that the series converges for any real or complex square matrix. Firstly, we will define what does it mean for matrix to converge. In this text, will be using Frobenius norm to describe the notion of matrix convergence, but it is possible to use any matrix norm which satisfies a set of properties discussed in the next lemma.

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 $\alpha \in \mathbb{K}$.*

- $\|A + B\|_F \leq \|A\|_F + \|B\|_F$

- $\|\alpha A\|_F = |\alpha| \|A\|_F$

- $\|A\|_F \geq 0$ *with equality occurring if and only if $A = O_{n \times m}$*

- $\|CD\|_F \leq \|C\|_F \|D\|_F$

Proof. Can be simply shown from the definition of Frobenius form and properties of the absolute value. $\qquad\square$

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

Proof. For non-negative real numbers a, b it holds $a \leq \sqrt{a^2 + b}$. Therefore, for arbitrary element of the matrix $X = (x_{i,j})$ we can write

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

It follows

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

$\qquad\square$

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 $\{A_k = (a_{i,j}^{(k)})\}_{k=0}^{\infty}$ of $n \times m$ matrices converges to a matrix A 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

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

$\square$

Lemma 4. *Let $\{A_k\}_{k=0}^{\infty}$ be a sequence of matrices over $\mathbb{K}$. This sequence converges if and only if it is a* **Cauchy sequence**, *that is when the following condition is satisfied*

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

Proof. By Lemma 3 the sequence converges if and only if it also converges elementwise. This happens if and only if all the sequences of elements are Cauchy what occurs if and only if the matrix sequence itself is Cauchy. The proof of the last equivalence is similar to the proof of Lemma 3. $\square$

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 the following equation holds*

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

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

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

Now we have to show that the left side converges to $B\left(\sum_{k=0}^{\infty} A_k\right) C$. Let $\varepsilon \in \mathbb{R}, \varepsilon > 0$ be fixed. We want to find such N_0 that for every $N \in \mathbb{N}, N \geq N_0$ it holds

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

Since $\sum_{k=0}^{\infty} \frac{1}{k!} A^k$ converges we can find such N_1 that $\forall N \geq N_1$ it holds

$$\left\|\sum_{k=0}^{\infty} A_k - \sum_{l=0}^{N} A_l\right\|_F < \frac{\varepsilon}{\|B\|_F \|C\|_F}$$

Now we can write

$$\left\| B\left(\sum_{k=0}^{\infty} A_k - \sum_{k=0}^{N} A_k\right)C \right\|_F \le \|B\|_F \left\| \sum_{k=0}^{\infty} A_k - \sum_{k=0}^{N} A_k \right\|_F \|C\|_F < \varepsilon$$

By putting $N_0 = N_1$ we show that the convergence holds and therefore we get the desired equation. $\qquad\square$

Remark. *We understand the differentiation of matrix elementwise, that is*

$$\frac{d}{dt}X = \left(\frac{d}{dt}x_{i,j}\right)$$

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

1. *$\sum_{k=0}^{\infty} \frac{1}{k!}X^k$ converges for any matrix X.*

2. *If $AB = BA$, then $e^A B = B e^A$*

3. *If R is invertible, then $e^{R^{-1}XR} = R^{-1}e^X R$*

4. *$\frac{d}{dt}e^{tX} = Xe^{tX}$, for $t \in \mathbb{R}$*

5. *If $AB = BA$, then $e^{A+B} = e^A e^B$*

Proof.   1. This can be shown by showing that sequence of partial sums $\{\sum_{k=0}^{N} \frac{1}{k!}X^k\}_{N=0}^{\infty}$ is Cauchy. Let $M, N \in \mathbb{N}$, then by Lemma 1 it holds

$$\left\| \sum_{k=0}^{M} \frac{1}{k!}X^k - \sum_{l=0}^{N} \frac{1}{l!}X^l \right\|_F = \left\| \sum_{k=N+1}^{M} \frac{1}{k!}X^k \right\|_F \le \sum_{k=N+1}^{M} \frac{1}{k!}\|X\|_F^k = \left\| \sum_{k=0}^{M} \frac{1}{k!}\|X\|_F^k - \sum_{l=0}^{N} \frac{1}{l!}\|X\|_F^l \right\|_F$$

Since $\{\sum_{k=0}^{N} \frac{\|X\|_F^k}{k!}\}_{N=0}^{\infty}$ is Cauchy, as it is sequence of partial sums of $e^{\|X\|_F}$, which always converges, we can for any $\varepsilon > 0$ find such N_0 that the first expression is strictly less then ε and therefore the sequence $\{\sum_{k=0}^{N} \frac{1}{k!}X^k\}_{N=0}^{\infty}$ is Cauchy.

2. Because of the convergence of the matrix exponential we can use Lemma 5 and we get

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

3. From Lemma 5 follows

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

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

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

where $X^k = (a_{i,j}^{(k)})$. To differentiate the elements of e^{tX} we need to first show that the series converges for any $t \in \mathbb{R}$. By Lemma 2 it holds $|a_{i,j}^{(k)}| \le \|X\|_F^k$. It follows

$$\left| \sum_{k=0}^{\infty} \frac{t^k}{k!} a_{i,j}^{(k)} \right| \le \sum_{k=0}^{\infty} \frac{|t|^k}{k!} |a_{i,j}^{(k)}| \le \sum_{k=0}^{\infty} \frac{\|tX\|_F^k}{k!} = e^{\|tX\|_F}$$

This shows that the series is absolutely convergent for every $t \in \mathbb{R}$. We can now differentiate the individual elements

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

Using Lemma 5 we get the desired result

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

5. Let us write

$$e^A e^B = \sum_{k=0}^{\infty} \frac{1}{k!} A^k \cdot \sum_{l=0}^{\infty} \frac{1}{l!} B^l = \sum_{k=0}^{\infty} \sum_{l=0}^{k} \frac{1}{l!(k-l)!} A^l B^{k-l} = \sum_{k=0}^{\infty} \sum_{l=0}^{k} \binom{k}{l} \frac{1}{k!} A^l B^{k-l}$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} (A + B)^k = e^{A+B}$$

The crucial point is the second equation in which we use Mertens' theorem. The theorem states that $\left(\sum_{i=0}^{\infty} a_i x^i \right) \cdot \left(\sum_{j=0}^{\infty} b_j x^j \right) = \sum_{k=0}^{\infty} a_l b_{k-l} x^k$ as long as at least one of the series on the left side is absolutely convergent. This condition is satisfied because matrix exponential converges absolutely for any matrix, as shown in the proof of the previous point. In the forth equation, we are using the assumption $AB = BA$.

□

Remark. $e^{\alpha I} = e^{\alpha} I$

Proof. Straight from the definition. □

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

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

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

is in the Jordan normal form. By substituting $y(t) = R^{-1}x(t)$, which is equivalent with changing the basis of our system, we get

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

and therefore the solution is

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

It is sufficient to show that $y(t)$ converges to $\mathbf{o}$, because since R is an invertible matrix, $y(t)$ converges to $\mathbf{o}$ if and only if $x(t)$ converges to $\mathbf{o}$.

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 $(\mathbf{o}, v_1, \ldots, v_{n-1})$.

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 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 J to the power of any natural number holds its block form, we can write

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

7

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 $\mathbf{o}$ if all the eigenvalues λ_i of matrix A have negative real parts. As the last step, we calculate $x(t) = Ry(t)$ and $x(0) = Ry(0)$.

Example 1. *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 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

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

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 $A \in \mathbb{K}^{n\times n}$ is a fundamental matrix, $B \in \mathbb{K}^{n\times m}$ is a **control matrix**, $u(t) \in \mathbb{K}^m$ is a **control vector**. Vector $x(t) \in \mathbb{K}^n$ is called a **state** of the system. The starting condition of the system is the state $x(0)$.*
We will regard this system as (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.

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.

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

$$u(t) = Fx(t)$$

*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. This means that we require that the system

converges to some fixed point. In many practical applications we require that this point is the origin of our state space, i.e., all the possible states of $x(t)$.

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

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

TODO 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 characteristic polynomial of a matrix are precisely the eigenvalues of said 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 through our choice of the feedback matrix F we need to satisfy the condition

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

where $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ 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 arbitrary polynomial χ of degree that depends on how "sensible" A and B are, 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 for *discrete-time systems* also holds for *continuos-time systems*.

1.2.1   Discrete-time systems

Let $\mathbb{K}$ be a field of real or complex numbers and let (A_1, B_1) be a continuos dynamical system $\dot{x}(t) = A_1 x(t) + B_1 u(t)$. We discretize the time, that is, instead of using continuous real time values of $x(t)$ and $\dot{x}(t)$, we will measure these values only at discrete sampling times $0, \delta, 2\delta, \ldots, k\delta, \ldots$ where $\delta \in \mathbb{R}^+$. The control

vector u is updated at each time $t = k\delta, k \in \mathbb{N}_0$ and is held constant until next update at $t = k\delta + \delta$. We will denote the states at each point in time as follows.

$$u_k = u(t) \quad x_k = x(t), t \in [k\delta, k\delta + \delta)$$

The derivative is approximated by the function

$$\frac{d}{dt}x_k \approx \frac{x_{k+1} - x_k}{\delta}$$

and therefore, we obtain

$$\frac{x_{k+1} - x_k}{\delta} = A_1 x_k + B_1 u_k$$
$$x_{k+1} = \delta A_1 x_k + x_k + \delta B_1 u_k$$
$$x_{k+1} = (\delta A_1 + I)x_k + \delta B_1 u_k$$
$$x_{k+1} = A x_k + B u_k$$

by choosing $A = \delta A_1 + I$ and $B = \delta B_1$. We see that we can calculate the next value of x from its previous value and a control vector. We will now define this system. The definition holds for any field $\mathbb{K}$.

Definition. *A **discrete dynamical linear system with control u** is a system of equations of form*
$$x_{k+1} = A x_k + B u_k, k \in \mathbb{N}_0$$
where $A \in \mathbb{K}^{n \times n}$ 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. Vector $x_k \in \mathbb{K}^n$ is called a state of the system. The starting condition of the system is the state x_0.
We will regard this system as (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 $x_0 = \boldsymbol{o}$ we get $x = x_k$, after k iterations of $x_{l+1} = A x_l + B x_l$, 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

$$x_{k+1} = A x_k + B u_k$$

The starting condition is $x_0 = \boldsymbol{o}$ and we can choose an arbitrary u_k. Then, for $k = 0$ we have
$$x_1 = A x_0 + B u_0 = B u_0 \in \text{Im}B.$$
For $k = 1$ we get

$$x_2 = A x_1 + B u_1 = ABu_0 + Bu_1 \in (AB|B)$$

It is clear, that
$$x_k \in \text{Im}(A^{k-1}B|\ldots|AB|B)$$
We can observe that $\text{Im}(B|AB|\ldots|A^k B) \subseteq \text{Im}(B|AB|\ldots|A^{k+1}B)$. From 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 $A^n B$ can be expressed as linear combination of matrices $\{B, AB, \ldots, A^{n-1}B\}$. We now see that $\text{Im}(B|AB|\ldots|A^k B) \supseteq \text{Im}(B|AB|\ldots|A^{k+1}B)$ holds. I t follows

$$\text{Im}(B|AB|\ldots|A^{n-1}B) = \text{Im}(B|AB|\ldots|A^{n-1}B|A^n B)$$

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

$$\text{Im}(B|AB|\ldots|A^{n-1}B)$$

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 $\text{Im}(B|AB|\ldots|A^{n-1}B)$.*

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 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 *discrete-time systems* also characterizes *continuous-time systems*. For this we have to express solution of such system using matrices $A^i B$ 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

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

$$e^{-t_1 A}x(t_1) - e^{-t_0 A}x(t_0) = \int_{t_0}^{t_1} e^{-tA}Bu(t)dt$$

$$x(t_1) = e^{(t_1-t_0)A}x(t_0) + \int_{t_0}^{t_1} e^{(t_1-t)A}Bu(t)dt$$

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

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

Theorem 1. *The n-dimensional continuos-time linear system is controllable, meaning that $x(t)$ can be equal to any vector in $\mathbb{K}^n$, if and only if $dim\mathcal{R}(A, B) = n$.*

Proof. From discussion above we have

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

The n-dimensional vector $w^{(k)}(s) = A^k Bu(s)$ 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 $A^k B$ 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 $\sum_{k=0}^{\infty} \frac{(t-s)^k}{k!} \alpha_{i,j}^{(k)} u_j(s)$ are convergent for every position (i, j). To show this, we will use the estimate from Lemma 2, that is, $|\alpha_{i,j}^{(k)}| \leq \left\|A^k B\right\|_F \leq \|A\|_F^k \|B\|_F$. Then

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

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

Because of the the convergence, we can now swap the integral and the series.

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

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

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

$$= \sum_{k=0}^{\infty} \sum_{j=1}^{m} \alpha_{i,j}^{(k)} v_i^{(k)}(t)$$

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

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

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 $\text{dim}\mathcal{R}(A, B) = n$.

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. $\qquad\square$

1.2.3 Decomposition theorem

Lemma 7. *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 $(w_1, \ldots, w_r)$ be an arbitrary basis of subspace W. We complete this sequence into basis C of V with vectors $v_1, \ldots, v_{n-r}$ where $n = \text{dim}V$, thus $C = (w_1, \ldots, w_r, v_1, \ldots, v_{n-r})$. We know that

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

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. $\qquad\square$

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 2. *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*

$$\widetilde{A} = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix} \qquad \widetilde{B} = \begin{pmatrix} B_1 \\ 0 \end{pmatrix} \tag{1.1}$$

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 7 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 from definition of reachable space it holds $\text{Im}(B) \subseteq \mathcal{R}(A, B)$, we see that $\widetilde{B}$ is in the requested form. $\qquad\square$

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 2 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*

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

Lemma 8. *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.

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

$\qquad\square$

Lemma 9. *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 8 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 $\widetilde{F}$.

$\qquad\square$

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

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

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 2 is a controllable pair which we will state in a lemma.

Lemma 10. *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 $\mathcal{R}(A_1, B_1)$.

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

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 $\mathbf{o}$. Also because of structure of $\widetilde{A}$ we have for an arbitrary matrix $X \in \mathbb{K}^{r\times m}$ that

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

where again are the last $n-r$ rows equal to $\mathbf{o}$. 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

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

From 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 2 decomposes the matrix A into "controllable" and "uncontrollable" parts A_1 and A_3 respectively.

Corollary 1. *For matrix A and its similar matrix $\widetilde{A}$ as in Theorem 2 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 11. *χ_c is independent of the choice of basis on $\mathcal{R}(A, B)$.*

Proof. From 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 2. Consider different basis D which suffices the conditions from the said proof. Then we obtain a similar matrix $\widetilde{B} = [id]_D^K A [id]_K^D$. 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 8 holds if A' and A'' are similar. Since $A' =$ □

Chapter 2

The Pole Shifting Theorem

Theorem 3. *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 2 and Lemma 9 we can assume that the pair (A, B) is in the same form as $(\widetilde{A}, \widetilde{B})$ in (1.1). 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

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

It follows

$$\chi_{A+BF} = \chi_{A_1+B_1 F_1} \chi_{A_3} = \chi_{A_1+B_1 F_1} \chi_u$$

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 F. Assume that we are given a polynomial χ. If we find such a matrix F_1 that

$$\chi_{A_1+B_1 F_1} = \chi$$

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 10, it is sufficient only to prove that the controllable systems can be assigned an arbitrary polynomial χ. Therefore, from this point on, we assume that the pair (A, B) is controllable.

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

So let $m = 1$. From Lemmas ????????? we can consider the pair (A, b) to be in the controller form. For a vector

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

we have

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

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

Now we can see that for a given polynomial

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

we can choose

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

and we then obtain $\chi_{A+bf} = \chi$.

For a general case, where m is arbitrary, we choose any vector $u \in \mathbb{K}^m$ such that $Bv \neq \mathbf{o}$, and let $b = Bv$. If we show that the pair $(A + BF_1, b)$ is controllable, we can then use the result for the case $m = 1$, and for any desired polynomial χ find such f that the characteristic polynomial of matrix

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

is χ and the Theorem will be proved using $F = (F_1 + vf)$. $\qquad \square$