Stabilisation of the Inverted Pendulum

Dynamic Systems

MATTEK3 4.205

December 17, 2021

P3 Project

STUDENT REPORT
AALBORG UNIVERSITY

AALBORG UNIVERSITY

STUDENT REPORT

Title:
Stabilisation of the Inverted Pendulum

Theme:
Dynamic Systems

Project period:
Autumn semester 2021

Project group:
MATTEK3 4.205

Participants:
Christian Dausel Jensen
Mikkel Sebastian Lundsgaard Brøndt
Julie Timmermann Werge
Sif Bjerre Lindby
Stine Byrjalsen

Supervisors:
John-Josef Leth
Oliver Matte

Copies: 1

Number of pages: 83

Date of completion:
December 17, 2021

Abstract:

In this project, we seek to stabilise the inverted pendulum in the upright position by deploying linear control theory. The equations of motion are obtained through Lagrangian mechanics. This yields a system of non-linear equations. Afterwards we introduce controllability, which provides us the tools to control a linear system. As our equations of motion are non-linear, we linearise them in order to design a matrix K that works as a controller on the system. This allows us to control and stabilise the linearised system, and therefore, the non-linearised system. The solutions to the system are determined using numerical methods. Afterwards, the obtained solutions are simulated. From the simulation we see that we can stabilise the inverted pendulum, as we take into consideration the restraints of the system.

Preface

This project is written in the autumn semester of 2021 by the group MATTEK3 4.205, consisting of five students from Mathematics and Technology at Aalborg University.

We would like to thank our supervisors, John-Josef Leth and Oliver Matte, both of whom have contributed to our project with advice and constructive criticism throughout the entire process.

This project presupposes that the reader has a fundamental knowledge of linear algebra and real analysis.

Unless otherwise stated, all figures, graphs, etc. are made using the TikZ package in LaTeX.

Calculations and simulations are made using `Spyder (anaconda3) Python 3.8`.

To mark the end of a partial proof, we use $\diamondsuit$. To mark the end of a proof, we use $\blacksquare$.

In this project, we have utilised IEEE referencing where the sources are referenced with numbers in square brackets. The references lead to the bibliography where the full information on the source is stated next to its number.

Christian Dausel Jensen

Mikkel Sebastian Lundsgaard Brøndt

Julie Timmermann Werge

Sif Bjerre Lindby

Stine Byrjalsen

Contents

Symbols and abbreviations

Symbol	**Interpretation**	**Comment**
$\mathbf{x}$	Vector $\mathbf{x}$ or vector function $\mathbf{x}$.	Bold letter. Should be clear from context.
$\langle \bullet \vert \bullet \rangle$	Inner product between vectors.	
$\mathcal{F}$	Field.	Either real ($\mathbb{R}$) or complex ($\mathbb{C}$).
$\overline{\bullet}$	Complex conjugate.	
j	Imaginary unit.	
$\operatorname{tr} \bullet$	Trace of a square matrix.	
$\det \bullet$	Determinant of a square matrix.	
$\bullet^{\mathsf{T}}$	Matrix transpose.	
$\operatorname{end}_{\mathcal{F}}(\bullet)$	The set of all endomorphisms over $\mathcal{F}$.	
$\dot{\mathbf{x}}(t), \ddot{\mathbf{x}}(t)$	Derivatives of time.	Used in mechanics and engineering.
$\frac{\mathrm{d}}{\mathrm{d}t}$	Derivative of time.	Used in formal mathematics.
L	Lagrangian or Lipschitz constant.	Should be clear from context.
$J_{\mathbf{f}(\mathbf{x})}$	Jacobian Matrix.	Capital letter.

Table 1: List of symbols and abbreviations used in the project.

Chapter 1

Introduction

A dynamical system is a system whose state changes over time. Modelling a dynamical system is the process of describing the motion of the system using equations. For this purpose, you can utilise a wide range of methods. In this project, we utilise Lagrangian Mechanics. When such a model is established, it can be used to simulate and hereby predict the future behaviour of said dynamical system. In order to simulate a dynamical system, a set of conditions is required: a set of initial states that describe the positions or velocities of the system.

Two essential properties of dynamical systems are stability and controllability. It is crucial to examine the stability of the system when a specific behaviour is desired. For instance, a stable state is crucial in drone technology where the drone must remain in the horizontal position regardless of the surrounding conditions. For a system to remain stabilised, it is crucial to be able to control the system. Therefore, in this project we examine the stability of the inverted pendulum (see Figure 1.1) in order to control and stabilise it.

1.1 Problem Statement

How can the application of *linear* control theory on the inverted pendulum, a *non-linear* mechanical system, be utilised to stabilise the pendulum in the upright position?

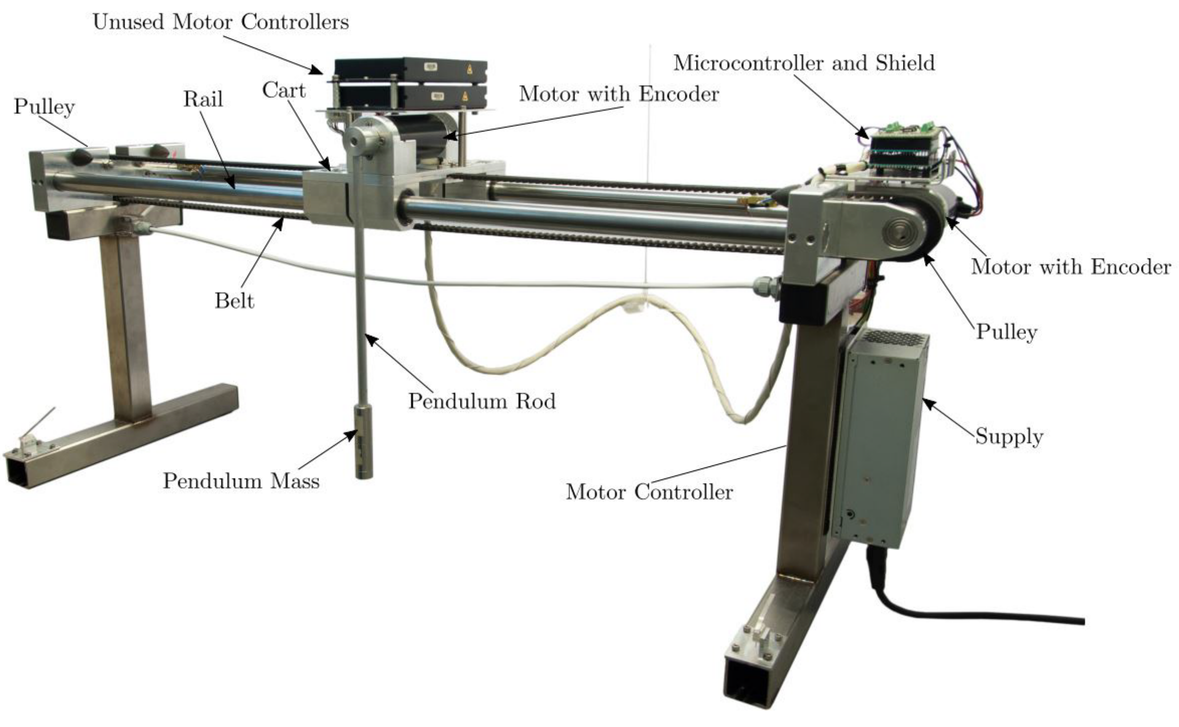


Figure 1.1: The experimental setup for the inverted pendulum. [21, p. 3]

Chapter 2

Linear Algebra

In this chapter, we will introduce a range of essential terms in linear algebra that will be used to build the foundation for the engineering later in the project.

2.1 Inner Product and Norms

In regards to proof-writing technique, it is beneficial to include inner product and norms. First, we will introduce inner products and its properties. Note that V denotes a vector space over $\mathcal{F}$ (either real $\mathbb{R}$ or complex $\mathbb{C}$).

Definition 2.1 (Inner Product and its Properties)
An inner product on V is a function that takes each ordered pair $(\mathbf{u}, \mathbf{v})$ of elements of V to a number $\langle \mathbf{u}|\mathbf{v}\rangle \in \mathcal{F}$ and has the following properties:

1. **Positivity:**
$$\langle \mathbf{v}|\mathbf{v}\rangle \geq 0 \; \forall \; \mathbf{v} \in V.$$
2. **Definiteness:**
$$\langle \mathbf{v}|\mathbf{v}\rangle = 0 \text{ if and only if } \mathbf{v} = \mathbf{0}.$$
3. **Conjugate symmetry:**
$$\langle \mathbf{u}|\mathbf{v}\rangle = \overline{\langle \mathbf{v}|\mathbf{u}\rangle} \; \forall \; \mathbf{u}, \mathbf{v} \in V.$$

[1, p. 166, Definition 6.3]

Remark for Definition 2.1: In the source, there are stated five properties of the inner product. We have chosen only to state the properties that are relevant to this project.

We utilise Definition 2.1 of inner products to compute the Euclidean norm.

Definition 2.2 (Euclidean Norm)
Let V be a vector space with an inner product. For $\mathbf{v} \in V$, the norm of $\mathbf{v}$, denoted $\|\mathbf{v}\|$, is defined as
$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}|\mathbf{v}\rangle}.$$

[1, p. 168, Definition 6.8]

The following theorem (Theorem 2.1) is essential in regards to proof-writing later in the project. The theorem and its proof include inner products and norms.

Theorem 2.1 (Cauchy-Schwarz Inequality)
Suppose $\mathbf{u}$,$\mathbf{v} \in V$. Then

$$|\langle \mathbf{u}|\mathbf{v}\rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|.$$

[1, p. 172, Theorem 6.15]

Remark for Theorem 2.1: This inequality is an equality if and only if one of $\mathbf{u}$,$\mathbf{v}$ is a scalar multiple of the other.

Proof
The proof contains two parts, of which the first is trivial.

1. If either $\mathbf{v}$ or $\mathbf{u}$ is the zero vector, then the inequality holds. ◇

2. Now assume that $\mathbf{v}, \mathbf{u} \neq \mathbf{0}$. Consider the orthogonal decomposition

$$\mathbf{u} = \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v} + \mathbf{w}$$

where $\mathbf{w}, \mathbf{v}$ are orthogonal per [1, p. 171, Theorem 6.14]. By the Pythagorean Theorem, we obtain

$$\begin{aligned}
\|\mathbf{u}\|^2 &= \left\| \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v} \right\|^2 + \|\mathbf{w}\|^2 \\
&= \frac{|\langle \mathbf{u}|\mathbf{v}\rangle|^2}{\|\mathbf{v}\|^4}\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 \\
&= \frac{|\langle \mathbf{u}|\mathbf{v}\rangle|^2}{\|\mathbf{v}\|^2} + \|\mathbf{w}\|^2.
\end{aligned}$$

Because $\|\mathbf{w}\|^2 \geq 0$, the following inequality holds:

$$\|\mathbf{u}\|^2 \geq \frac{|\langle \mathbf{u}|\mathbf{v}\rangle|^2}{\|\mathbf{v}\|^2}$$

By multiplying with a positive constant, the inequality is kept

$$\begin{gathered}
\|\mathbf{u}\|^2\|\mathbf{v}\|^2 \geq |\langle \mathbf{u}|\mathbf{v}\rangle|^2 \\
\|\mathbf{u}\| \cdot \|\mathbf{v}\| \geq |\langle \mathbf{u}|\mathbf{v}\rangle|,
\end{gathered}$$

hence proving the Cauchy-Schwarz Inequality. ■

We have defined the norm of a vector (Definition 2.2). It will be beneficial in regards to proof-writing to define the norm of matrices, the Frobenius norm.

Definition 2.3 (Frobenius Norm)
Let $A = [a_{ij}] \in \mathcal{F}^{n\times n}$, and $i, j \in \mathbb{N}$. The Frobenius norm of A is defined as

$$\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2} = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2\right)^{\frac{1}{2}}.$$

[13, p. 333] and [7, p. 5, Definition 4.1]

The Frobenius norm can also be written as seen in Lemma 2.1.

Lemma 2.1 (Frobenius Norm for $\mathcal{F}^{n\times n}$ Matrices)
The Frobenius norm of a matrix $A \in \mathcal{F}^{n\times n}$ is

$$\|A\|_F = \left(\sum_{i=1}^{n} \langle \mathbf{s}_i|\mathbf{s}_i\rangle\right)^{\frac{1}{2}} = \left(\sum_{i=1}^{n} \|\mathbf{s}_i\|^2\right)^{\frac{1}{2}}$$

where $\mathbf{s}_i$ is the i'th column vector of A. [13, p. 333]

Proof
Let $A \in \mathcal{F}^{n\times n}$ be a matrix and $\mathbf{s}_i \in \mathcal{F}^n$ be the i'th column of A such that

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_n \end{bmatrix}.$$

By Definition 2.3, we have

$$\begin{aligned} \|A\|_F &= \left(\left(a_{11}^2 + \cdots + a_{1n}^2\right) + \cdots + \left(a_{n1}^2 + \cdots + a_{nn}^2\right)\right)^{\frac{1}{2}} \\ &= \left(\left(a_{11}^2 + \cdots + a_{n1}^2\right) + \cdots + \left(a_{1n}^2 + \cdots + a_{nn}^2\right)\right)^{\frac{1}{2}} \\ &= \left(\langle \mathbf{s}_1|\mathbf{s}_1\rangle + \cdots + \langle \mathbf{s}_n|\mathbf{s}_n\rangle\right)^{\frac{1}{2}} \\ &= \left(\sum_{i=1}^{n} \langle \mathbf{s}_i|\mathbf{s}_i\rangle\right)^{\frac{1}{2}}. \end{aligned}$$

Given $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}|\mathbf{v}\rangle}$ by Definition 2.2, we can write

$$\|A\|_F = \left(\sum_{i=1}^{n} \|\mathbf{s}_i\|^2\right)^{\frac{1}{2}},$$

hence completing the proof. ∎

The Cauchy-Schwarz Inequality (Theorem 2.1) yields the following properties of the Frobenius norm.

Theorem 2.2 (Properties of the Frobenius Norm)
Let $A \in \mathbb{R}^{n\times n}$, then the following properties hold:

1. $\|AB\|_F \leq \|A\|_F \|B\|_F$ for $B \in \mathbb{R}^{n\times n}$,
2. $\|A\mathbf{x}\| \leq \|A\|_F \|\mathbf{x}\|$ for $\mathbf{x} \in \mathbb{R}^n$.

[13, p. 333] and [7, p. 5]

Proof
The two properties are proven seperately.

1. Let $A, B \in \mathbb{R}^{n\times n}$ and look at the left-hand side of the inequality. We calculate the Frobenius norm of the matrix product AB squared:
$$\begin{aligned}\|AB\|_F^2 &= \sum_{i=1}^n \sum_{j=1}^n (AB)_{ij}^2 \\ &= \sum_{i=1}^n \sum_{j=1}^n \langle \mathbf{r}_i | \mathbf{s}_j \rangle^2,\end{aligned}$$
where $\mathbf{r}_i$ is the i'th row vector of A and $\mathbf{s}_j$ is the j'th column vector of B. The Cauchy-Schwarz Inequality (Theorem 2.1) is applied in order to obtain the following inequality:
$$\sum_{i=1}^n \sum_{j=1}^n \langle \mathbf{r}_i | \mathbf{s}_j \rangle^2 \leq \sum_{i=1}^n \sum_{j=1}^n \|\mathbf{r}_i\|^2 \|\mathbf{s}_j\|^2 = \sum_{i=1}^n \|\mathbf{r}_i\|^2 \sum_{j=1}^n \|\mathbf{s}_j\|^2.$$
Thus we have obtained the right-hand side of the inequality in Theorem 2.2 property 1. By Lemma 2.1 we get:
$$\|AB\|_F^2 \leq \|A\|_F^2 \|B\|_F^2 \quad \iff \quad \|AB\|_F \leq \|A\|_F \|B\|_F. \qquad \diamond$$

2. (2) is proven using the same method as (1) where $\mathbf{x}$ is now a $\mathbb{R}^{n\times 1}$ matrix.
Let $A \in \mathbb{R}^{n\times n}$ and $\mathbf{x} \in \mathbb{R}^n$, and look at the left-hand side of the inequality in Theorem 2.2 property 2. We calculate the norm of the matrix vector product $A\mathbf{x}$ squared:
$$\begin{aligned}\|A\mathbf{x}\|^2 &= \sum_{i=1}^n (A\mathbf{x})_i^2 \\ &= \sum_{i=1}^n \langle \mathbf{r}_i | \mathbf{x} \rangle^2\end{aligned}$$
where $\mathbf{r}_i$ is the i'th row vector of A. The Cauchy-Schwarz Inequality (Theorem 2.1) is applied in order to obtain the following inequality:
$$\sum_{i=1}^n \langle \mathbf{r}_i | \mathbf{x} \rangle^2 \leq \sum_{i=1}^n \|\mathbf{r}_i\|^2 \|\mathbf{x}\|^2.$$

Therefore we have the right-hand side of the inequality, and by Lemma 2.1 we get:

$$\|A\mathbf{x}\|^2 \leq \|A\|_F^2 \|\mathbf{x}\|^2 \quad \Longleftrightarrow \quad \|A\mathbf{x}\| \leq \|A\|_F \|\mathbf{x}\|. \qquad \blacksquare$$

2.2 Eigenvalues

Eigenvalues are essential in linear algebra as well as engineering because they are used to solve differential equations and to analyse the stability of a system (which will be explained further in Chapter 6). To find eigenvalues of a matrix, we will define the characteristic equation in the following definition.

Definition 2.4 (Characteristic Equation and Eigenvalues)
A scalar $\lambda \in \mathcal{F}$ is an eigenvalue of a matrix $A \in \mathcal{F}^{n\times n}$ if and only if the characteristic equation

$$P_A(\lambda) = \det(A - \lambda I) = 0$$

is satisfied. [9, p. 294]

In special cases, the eigenvalues can be found more conveniently than by solving the characteristic equation.

Definition 2.5 (Upper-Triangular Matrix)
Given a matrix $A \in \mathcal{F}^{n\times n}$ with zero-entries under the diagonal

$$\begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix},$$

where $* \in \mathcal{F}$. Then the eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathcal{F}$ of the matrix are the entries on the diagonal. [1, p. 185]

From the diagonal of any arbitrary square matrix, we obtain the trace.

Definition 2.6 (Trace)
The trace of a matrix $A \in \mathcal{F}^{n\times n}$ is defined to be the sum of the diagonal entries of A such that

$$\text{tr } A = \sum_{i=1}^{n} a_{ii} = a_{11} + \cdots + a_{nn}.$$

[1, p. 300, Definition 10.13]

Chapter 3

Ordinary Differential Equations

In this chapter, ordinary differential equations and systems of ordinary differential equations are introduced. Furthermore, we will assess existence and uniqueness of solutions to systems of differential equations. Lastly, Lipschitz conditions and continuity will be introduced. Going forward, an ordinary differential equation will be referred to only as an "ODE".

Definition 3.1 (Explicit ODE)
An explicit ODE of n'th order is written on the form

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}x(t) = f\left(t, x(t), \ldots, \frac{\mathrm{d}^{n-1}}{\mathrm{d}t^{n-1}}x(t)\right), \tag{3.1}$$

where f is a function describing the n'th derivative of $x(t)$.

[14, p. 2]

Remark for Definition 3.1: The ODE is called explicit since it is explicitly solved for $\frac{\mathrm{d}^n}{\mathrm{d}t^n}x(t)$.

Thus we define a first-order explicit ODE in $\mathbb{R}^n$. We define $\mathcal{D}_f \subseteq \mathbb{R} \times \mathbb{R}^n$ and $f : \mathcal{D}_f \to \mathbb{R}^n$. The equation has the form $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(t, x(t))$.

Definition 3.2 (Solution to first-order explicit ODE)
Let $I \subseteq \mathbb{R}$ and $x : I \to \mathbb{R}^n$ be a function. We say that x is a solution of the first-order explicit ODE $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(t, x(t))$ if and only if:

- The set I is an interval.
- The function x is differentiable in I.
- For all $t \in I$, we have $(t, x(t)) \in \mathcal{D}_f$.
- For all $t \in I$, we have $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(t, x(t))$.

[13, p. 3]

There exists different types of ODE's, including linear and non-linear. Linear ODE's can be solved analytically whereas non-linear ODE's must be solved numerically (see Section 9.1.1).

Definition 3.3 (Linear ODE)
An ODE is linear if it can be written as a linear combination:

$$a_0(t)x(t) + a_1(t)\frac{\mathrm{d}}{\mathrm{d}t}x(t) + \cdots + a_{n-1}(t)\frac{\mathrm{d}^{n-1}}{\mathrm{d}t^{n-1}}x(t) + a_n(t)\frac{\mathrm{d}^n}{\mathrm{d}t^n}x(t) = b(t).$$

If $b(t) = 0$, the ODE is called homogeneous. Otherwise it is called inhomogeneous.

[14, p. 3]

In Definition 3.1 we looked at the general case of an n'th order ODE. Now, we will consider the n'th order ODE written as a system expressed on the following form

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}x_1(t) &= f_1(t, x_1(t), x_2(t), \ldots, x_n(t)) \\
\frac{\mathrm{d}}{\mathrm{d}t}x_2(t) &= f_2(t, x_1(t), x_2(t), \ldots, x_n(t)) \\
&\vdots \\
\frac{\mathrm{d}}{\mathrm{d}t}x_n(t) &= f_n(t, x_1(t), x_2(t), \ldots, x_n(t)),
\end{aligned}$$

where $f_1, f_2, \ldots, f_n$ are functions of $n+1$ variables $(t, x_1, x_2, \ldots, x_n)$. This system can be written using vector notation

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} \in \mathbb{R}^n, \quad \mathbf{f}(t, \mathbf{x}(t)) = \begin{bmatrix} f_1(t, x_1(t), x_2(t), \ldots, x_n(t)) \\ f_2(t, x_1(t), x_2(t), \ldots, x_n(t)) \\ \vdots \\ f_n(t, x_1(t), x_2(t), \ldots, x_n(t)) \end{bmatrix} \in \mathbb{R}^n.$$

Then it is possible to write the system as the n-dimensional first-order differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \mathbf{f}(t, \mathbf{x}(t)).$$

The initial value problem (IVP) of the system is

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \mathbf{f}(t, \mathbf{x}(t)), \qquad \mathbf{x}(t_0) = \mathbf{x}_0.$$

The ODE itself has multiple solutions whereas the initial condition ensures the uniqueness of the solution.

3.1 Existence and Uniqueness

In this section, the Lipschitz condition will be introduced. The Lipschitz condition is sufficient to guarantee existence and uniqueness of a solution to a differential equation. We will first introduce Definition 3.4 to further understand theorems later in this section.

Definition 3.4 (Open Subset)
A subset $U \subseteq \mathbb{R}^n$ is called open if all points in U are inner points, such that

$$\forall\, x \in U\, \exists\, r > 0 : B_r(x) \subseteq U,$$

where $B_r(x)$ is a ball with radius r and centre point x.

[15, p. 91, Definition 6.11]

Therefore the subset Ω in Theorem 3.1 is required to have every point be an inner point.

Theorem 3.1 (Local Existence and Uniqueness)
Let $\Omega \subseteq \mathbb{R} \times \mathbb{R}^n$ be an open non-empty subset and $\mathbf{f} : \Omega \to \mathbb{R}^n$ be a continuous function. Assume there exists an $L > 0$ such that:

$$\|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \tilde{\mathbf{x}})\| \leq L\|\mathbf{x} - \tilde{\mathbf{x}}\| \quad \forall (t, \mathbf{x}), (t, \tilde{\mathbf{x}}) \in \Omega, \tag{3.2}$$

where $\mathbf{x}, \tilde{\mathbf{x}}$ are inner points in Ω.
Let $(t_0, \mathbf{x}_0) \in \Omega$. Then there exists a $\delta > 0$ and a unique, continuously differentiable function $\mathbf{g} : (t_0 - \delta, t_0 + \delta) \to \mathbb{R}^n$, such that $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{g}(t) = \mathbf{f}(t, \mathbf{g}(t)), t \in (t_0 - \delta, t_0 + \delta)$ and $\mathbf{g}(t_0) = \mathbf{x}_0$.

[13, p. 234, Theorem 9.8] and [7, p. 13, Theorem A.1]

Theorem 3.1 ensures the local existence and uniqueness of solutions. We can rewrite (3.2) by isolating L, which yields the Lipschitz condition

$$\frac{\|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \tilde{\mathbf{x}})\|}{\|\mathbf{x} - \tilde{\mathbf{x}}\|} \leq L, \quad \mathbf{x} \neq \tilde{\mathbf{x}}.$$

The Lipschitz condition restricts the slope of the function $\mathbf{f}$ in the x-direction on the subset Ω. Rewriting provides us with a formula for the Lipschitz constant.

A function $\mathbf{f}$ can be locally Lipschitz continuous in $\mathbf{x}$. The concept "Lipschitz continuity" is introduced in Definition 3.5.

Definition 3.5 (Lipschitz Continuity)

1. Let $V \subset \mathbb{R} \times \mathbb{R}^n$ be open and $\mathbf{f} : V \to \mathbb{R}^n$ be a function. Then the function $\mathbf{f}$ fulfills the Lipschitz condition on V if and only if there exists a Lipschitz constant $L \in [0, \infty)$ such that:

$$\|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \tilde{\mathbf{x}})\| \leq L\|\mathbf{x} - \tilde{\mathbf{x}}\| \quad \forall\, (t, \mathbf{x}), (t, \tilde{\mathbf{x}}) \in V.$$

2. We call $\mathbf{f}$ locally Lipschitz continuous in $\mathbf{x}$ or say it fulfills a Lipschitz condition locally if and only if for every $(t, \mathbf{x}) \in V$ we can find some open subset $U \subset V$ with $(t, \mathbf{x}) \in U$ such that the restriction of $\mathbf{f}$ to the subset U fulfills a Lipschitz condition on U. We set:

$$C^{0,1-}(V, \mathbb{R}^n) := \{\mathbf{f} \in C(V, \mathbb{R}^n) |\ \mathbf{f} \text{ is locally Lipschitz continuous in } x\}.$$

[12, p. 3, Definition 2.1]

Remark (1) for Definition 3.5: The definition begins by introducing V and afterwards it describes a function in the exact same way as Theorem 3.1. The reason we introduce this definition is that we wish to describe $C^{0,1-}(V, \mathbb{R}^n)$ since it is necessary for understanding global Lipschitz later in the chapter.

Example 3.1

Let $f(t, x) = t^2 + 2x$ for each $(t, x), (t, \tilde{x}) \in \mathbb{R}^2$, as shown on Figure 3.1.

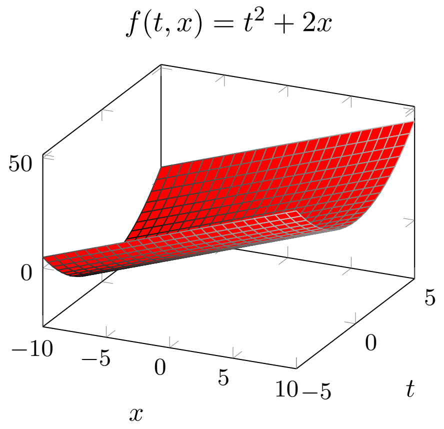


Figure 3.1: The function f as a surface plot.

We use Definition 3.5 to calculate the Lipschitz constant:

$$|f(t, x) - f(t, \tilde{x})| = |(t^2 + 2x) - (t^2 + 2\tilde{x})| = 2|(x - \tilde{x})| \quad \Rightarrow \quad L = 2.$$

We say f satisfies a Lipschitz condition on $\mathbb{R}^2$ because $L \geq 0$. Hence

$$\frac{|f(t, x) - f(t, \tilde{x})|}{|x - \tilde{x}|} \leq 2, \qquad x \neq \tilde{x}.$$

Since x is one-dimensional, the calculated Lipschitz constant will not exceed a slope ± 2. This means that any two points, $(x, \tilde{x})$, we will be able to plot:

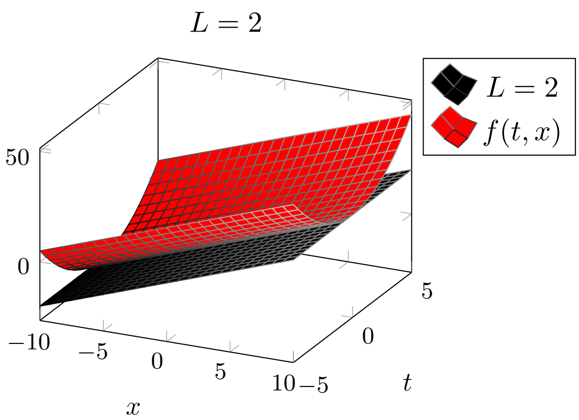


Figure 3.2: Lipschitz constant, $-2 \leq L \leq 2$, creating a slope plotted as the plane underneath our surface plot of Figure 3.1

.

Remark: Figure 3.2 does not tell us anything about the slope, when comparing two points in the t-direction.

We have defined the Local Existence and Uniqueness conditions. Furthermore, we expand Theorem 3.1 to describe the Global Existence and Uniqueness of a solution. It can ensure that if $\mathbf{f}$ is Lipschitz continuous, then the solution exists and is unique globally.

Theorem 3.2 (The Global Existence and Uniqueness Theorem)
Let $V \subset \mathbb{R} \times \mathbb{R}^n$ be open and $\mathbf{f} \in C^{0,1-}(V, \mathbb{R}^n)$ and $(\tau, \xi) \in V$. Then the IVP

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \mathbf{f}(t, \mathbf{x}(t)), \quad \mathbf{x}(\tau) = \xi$$

is uniquely solvable.

[12, p. 3, Definition 2.3]

If the function within coordinates in V, (τ, ξ), is Lipshitz on V, provided that Theorem 3.2 is satisfied $\forall \tau \in V$ with the same Lipschitz constant, then $\mathbf{f}$ must be globally Lipschitz.

3.2 Autonomous Linear Systems

A system of first-order ODE's can be autonomous and/or linear. An autonomous system is a system where the independent variable $t \in \mathbb{R}$ does not explicitly appear in the equations describing the system. A such system can be expressed as

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{f}(\mathbf{x}),$$

where $\mathbf{x} \in \mathbb{R}^n$. If $f_1, f_2, \ldots, f_n$ are linear and depend only on the variables $x_1, x_2, \ldots, x_n$, then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}x_1(t) &= a_{11}x_1(t) + a_{12}x_2(t) + \cdots + a_{1n}x_n(t) \\
\frac{\mathrm{d}}{\mathrm{d}t}x_2(t) &= a_{21}x_1(t) + a_{22}x_2(t) + \cdots + a_{2n}x_n(t) \\
&\vdots \\
\frac{\mathrm{d}}{\mathrm{d}t}x_n(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + \cdots + a_{nn}x_n(t)
\end{aligned}$$

is a linear and autonomous system. A such system can be written as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = A\mathbf{x}(t),$$

where A is an $n \times n$ matrix containing all the coefficients $[a_{ij}]$, and $\mathbf{x}$ is a vector containing the coefficients $\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^\intercal$. The IVP for this system is

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = A\mathbf{x}(t), \qquad \mathbf{x}(t_0) = \mathbf{x}_0, \tag{3.3}$$

where $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$ and $\mathbf{x}_0 \in \mathbb{R}^n$. This can be generalised to a system, where the solution is a matrix, rather than a vector.

Theorem 3.3 (Existence of Unique Solution to Autonomous Linear System)
Let $A \in \mathbb{R}^{n\times n}$, $\Phi : \mathbb{R} \to \mathbb{R}^{n\times n}$ and I the identity matrix. Then the autonomous linear homogeneous system

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t) = A\Phi(t), \quad \Phi(0) = I \tag{3.4}$$

has a unique matrix value solution. The solution is defined on the real number axis.

[13, p. 218] and [7, p. 5]

The matrix Φ consists of n column vectors $\mathbf{s}_i$, where $i = 1, 2, \ldots, n$, and can be expressed as

$$\Phi(t) = \begin{bmatrix} s_1(t) & s_2(t) & \cdots & s_n(t) \end{bmatrix}.$$

Then it follows that we can write (3.4) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} s_1(t) & s_2(t) & \cdots & s_n(t) \end{bmatrix} = A\begin{bmatrix} s_1(t) & s_2(t) & \cdots & s_n(t) \end{bmatrix}.$$

As differentiating a matrix is done element-wise, we obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} s_1(t) &= A s_1(t), & s_1(0) &= \mathbf{e}_1, \\
\frac{\mathrm{d}}{\mathrm{d}t} s_2(t) &= A s_2(t), & s_2(0) &= \mathbf{e}_2, \\
&\vdots & &\vdots \\
\frac{\mathrm{d}}{\mathrm{d}t} s_n(t) &= A s_n(t), & s_n(0) &= \mathbf{e}_n,
\end{aligned}$$

where $I = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{bmatrix}$, and $\mathbf{e}_i$ is the standard basis vector of $\mathbb{R}^n$ for $i = 1, 2, \ldots, n$. Each of these differential equations is on the form (3.3).

Chapter 4

The Matrix Exponential

In this chapter, the matrix exponential function and its properties are introduced, since it is essential for autonomous linear systems. The solution for the linear systems can be obtained by using the matrix exponential function.

4.1 Introduction to the Matrix Exponential

This section provides a method for computing the matrix exponential function. This method is called Putzer's Algorithm. In order to derive Putzer's Algorithm, the following prerequisites are needed.

4.1.1 Prerequisites

Theorem 4.1 (Cayley-Hamilton)
Suppose V is a complex vector space and $T \in \mathrm{end}_{\mathbb{C}}(V)$. Let P denote the characteristic polynomial of T. Then $P(T) = 0$.

[1, p. 261, Theorem 8.37]

Definition 4.1 (The Matrix Exponential Function)
The solution to the IVP

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t) = A\Phi(t), \quad \Phi(0) = I \tag{3.4}$$

is given by the matrix exponential function e^{tA}

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{tA} = A\mathrm{e}^{tA}, \quad \mathrm{e}^{tA}\big|_{t=0} = I.$$

[13, p. 153] and [7, pp. 5-6]

Theorem 4.2 (Properties of the Matrix Exponential Function)
Let $A, B \in \mathbb{R}^{n\times n}$ and assume $AB = BA$. The matrix exponential function $\mathrm{e}^{tA} : \mathbb{R} \to \mathbb{R}^{n\times n}$ then has the properties:

1. $B\mathrm{e}^{tA} = \mathrm{e}^{tA}B, \quad \forall t \in \mathbb{R}.$
2. $\mathrm{e}^{(s+t)A} = \mathrm{e}^{sA}\mathrm{e}^{tA} \quad \forall s, t \in \mathbb{R}.$

[13, p. 142] and [7, p. 7]

Proof

1. To prove that e^{tA} commutes, we apply the following: B is an arbitrary matrix, $B \in \mathbb{R}^{n\times n}$. We define

$$\Psi_1(t) = B\mathrm{e}^{tA},$$
$$\Psi_2(t) = \mathrm{e}^{tA}B.$$

To show they can commute they must both be a solution to the IVP in Definition 4.1. Let $AB = BA$, hence the commutative rewriting:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi_1(t) = BA\mathrm{e}^{tA} = AB\mathrm{e}^{tA} = A\Psi_1(t).$$

Furthermore:

$$\Psi_1(0) = B.$$

Now we use the same method for $\Psi_2(t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi_2(t) = A\mathrm{e}^{tA}B = A\Psi_2(t).$$

Furthermore:

$$\Psi_2(0) = B.$$

Theorem 3.3 states that there is a unique solution, therefore we have $\Psi_1(t) = \Psi_2(t)$, which means the equation works for an arbitrary matrix B. This proof also accounts for the special case where $A = B$. $\diamond$

2. We define:

$$\Psi_1(t) = \mathrm{e}^{(s+t)A}. \tag{4.1}$$

The goal is to differentiate $\Psi_1(t)$. We show that $\Psi_1(t)$ is a solution to the IVP

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t) = A\Phi(t), \quad \Phi(0) = \mathcal{E}(s).$$

The chain rule also applies to matrices. However, this is not proven. Therefore, we apply the chain rule to $\Psi_1(t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{(s+t)A}\right) = A\mathrm{e}^{(s+t)A} \cdot \frac{\mathrm{d}}{\mathrm{d}t}(s+t) = A\mathrm{e}^{(s+t)A}.$$

Choose $s = 0$, to fulfill the IVP

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi_1(t) = \mathrm{e}^{tA}A = A\mathrm{e}^{tA}.$$

We know from Theorem 4.2 property 1 that the exponential matrix function is commutative, hence $\Psi_1(t)$ is a solution to the IVP in Definition 4.1. Now we define:

$$\Psi_2(t) = \mathrm{e}^{sA}\mathrm{e}^{tA}.$$

The product rule also applies to matrices. However, this is not proven. As before, we want to see if $\Psi_2(t)$ is a solution to Theorem 3.3, and we differentiate the matrix:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{sA}\mathrm{e}^{tA}\right) = \frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{sA}) \cdot \mathrm{e}^{tA} + \mathrm{e}^{sA} \cdot \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{tA} = 0 + \mathrm{e}^{sA}A\mathrm{e}^{tA} = A\mathrm{e}^{sA}\mathrm{e}^{tA}.$$

When $s = 0$, e^{tA} satisfies the Theorem 3.3, and then we have:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi_2(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{0A}\mathrm{e}^{tA}\right) = A\mathrm{e}^{tA}.$$

This shows $\Psi_1(t)$ and $\Psi_2(t)$ are solutions to Theorem 3.3, and by uniqueness we obtain

$$\Psi_1(t) = \Psi_2(t).$$ ■

In the derivation for Putzer's Algorithm, we must solve inhomogeneous linear first-order ODE's. Therefore, we introduce the general solution for said ODE's.

Theorem 4.3 (Solution for Inhomogeneous Linear First-Order ODE's)
Let $J \subset \mathbb{R}$ be an interval and $a, b : J \to \mathbb{R}$ be continuous. Then the IVP

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = a(t)x(t) + b(t), \qquad x(t_0) = x_0,$$

is uniquely solvable on the the whole interval J. Its unique solution is given by:

$$x(t) = x_0\mathrm{e}^{\int_{t_0}^{t} a(r)\mathrm{d}r} + \int_{t_0}^{t} \mathrm{e}^{\int_{s}^{t} a(r)\mathrm{d}r} b(s)\mathrm{d}s.$$

[12, p. 5, Theorem 3.2]

4.1.2 Putzer's Algorithm

Putzer's Algorithm computes e^{tA}:

$$\mathrm{e}^{tA} = \sum_{k=0}^{n-1} r_{k+1}(t) P_k. \tag{4.2}$$

From [13, p. 151] and [7, p. 9], we define

$$\begin{aligned} P_0 &= I, \\ P_k &= (A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_k I), \quad k = 1, 2, \ldots, n \end{aligned} \tag{4.3}$$

Note that, per Theorem 4.1, $P_n = 0$, and $\lambda_1, \ldots, \lambda_k$ are eigenvalues.

We now look at e^{tA}, where $t = 0$:

$$\mathrm{e}^{0 \cdot A} = I. \tag{4.4}$$

We now see that (4.4) gives us the requirement:

$$I = \sum_{k=0}^{n-1} r_{k+1}(0) P_k \tag{4.5}$$

From (4.3) we see that (4.5) is true if:

$$r_1(0) = 1 \text{ and } r_k(0) = 0, \quad k = 2, 3, \ldots, n. \tag{4.6}$$

We now differentiate both sides of (4.2) with respect to t and we get:

$$A\mathrm{e}^{tA} = \sum_{k=0}^{n-1} \frac{\mathrm{d}}{\mathrm{d}t}(r_{k+1}(t)) P_k. \tag{4.7}$$

The left-hand side comes from Definition 4.1, and we see that P_k is not a function of t, therefore, only making it necessary to differentiate $r_{k+1}(t)$.

Looking further into the left-hand side we see that

$$A\mathrm{e}^{tA},$$

can also be written as

$$A \sum_{k=0}^{n-1} r_{k+1}(t) P_k = \sum_{k=0}^{n-1} r_{k+1}(t) A P_k, \tag{4.8}$$

per (4.2).

An important "rewriting" to note is:

$$AP_k = AP_k - \lambda_{k+1} P_k + \lambda_{k+1} P_k = (A - \lambda_{k+1} I) P_k + \lambda_{k+1} P_k = P_{k+1} + \lambda_{k+1} P_k. \tag{4.9}$$

Equation (4.9) is now inserted into (4.8). Hereby we obtain:

$$\sum_{k=0}^{n-1} r_{k+1}(t)(P_{k+1} + \lambda_{k+1} P_k) = \sum_{k=0}^{n-1} r_{k+1}(t) \lambda_{k+1} P_k + \sum_{k=1}^{n-1} r_k(t) P_k. \tag{4.10}$$

If we look at $\sum_{k=0}^{n-1} r_{k+1}(t)P_{k+1}$, we see that the summation index can be changed from $k = 0$ and $n-1$ to $k = 1$ and n. We then note that the n'th index of the sum is zero per Theorem 4.1. Therefore, the summation index is $k = 1$ and $n-1$.
We now see that the right-hand side in (4.10) is equal to the right-hand side in (4.7):

$$\sum_{k=0}^{n-1} r_{k+1}(t)\lambda_{k+1}P_k + \sum_{k=1}^{n-1} r_k(t)P_k = \sum_{k=0}^{n-1} \frac{\mathrm{d}}{\mathrm{d}t}(r_{k+1}(t))P_k. \tag{4.11}$$

From (4.11) we get the following differential equations with the starting conditions mentioned in (4.6):

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}r_1(t) &= \lambda_1 r_1(t), & r_1(0) &= 1,\\
\frac{\mathrm{d}}{\mathrm{d}t}r_2(t) &= \lambda_2 r_2(t) + r_1(t), & r_2(0) &= 0,\\
&\vdots & &\vdots\\
\frac{\mathrm{d}}{\mathrm{d}t}r_n(t) &= \lambda_n r_n(t) + r_{n-1}(t), & r_n(0) &= 0.
\end{aligned}$$

The general solution for these differential equations is given by Theorem 4.3 as:

$$\begin{aligned}
r_1(t) &= \mathrm{e}^{t\lambda_1},\\
r_k(t) &= \mathrm{e}^{t\lambda_k}\int_0^t \mathrm{e}^{-s\lambda_k} r_{k-1}(s)\mathrm{d}s, \qquad k = 2, 3, \dots, n.
\end{aligned}$$

If you want further information about utilisation and generalisations regarding the matrix exponential, see Appendices A and B, as they include generalisations for $\mathbb{R}^{2\times 2}$ and $\mathbb{R}^{3\times 3}$, respectively.

Chapter 5

Lagrangian Mechanics

In this chapter, the essential concepts of Lagrangian mechanics are introduced. These concepts are beneficial in order to model and simulate the stabilisation of the pendulum later in the project. The material in this chapter is based on [11, Chapter 1].

5.1 Prerequisites for Lagrangian Mechanics

In order to understand the arguments in this chapter, we will first need to recall some basic terms from physics.

Definition 5.1 (Kinetic Energy of a Particle)
The kinetic energy of a particle with mass m and velocity v is defined as

$$T = \frac{1}{2}mv^2.$$

[5, p. 15]

Remark for Definition 5.1: Note that the velocity of the particle may be a vector. This yields $T = \frac{1}{2}m\langle \mathbf{v}|\mathbf{v}\rangle$.

Potential energy yields conservative forces on a system, and conservative forces yield potential energy. Conservative forces do not alter the mechanical state of the system, meaning that no energy is converted to other types of energy other than mechanical, i.e. thermal energy. In this project, we utilise a special case of potential energy, which is the potential energy caused by the interaction between a particle and the gravitational field of the Earth. This yields

$$U = mgh,$$

where m is the mass of the particle, g is the gravity acceleration, and h is the particle's perpendicular height to the surface of the Earth. [5, p. 16] Note that this is an approximation and only accounts for systems near the surface of the Earth where the gravitational field is strong.

The kinetic and potential energy are important as they both appear in the Lagrangian of a closed system of n particles.

Definition 5.2 (The Lagrangian)
The Lagrangian $L : \mathbb{R}^{2s+1} \to \mathbb{R}$ of a closed system with n particles,

$$L = \sum_{i=1}^{n} \frac{1}{2} m_i v_i^2 - U(r_1, r_2, \ldots, r_n),$$

where $T = \sum_{i=1}^{n} \frac{1}{2} m_i v_i^2$ is the total kinetic energy of the system, and $U = U(r_1, r_2, \ldots, r_n)$ is the total potential energy of the system.

[11, p. 8]

The position of a system can be described by a set of coordinates. Therefore, it is necessary to define generalised coordinates, which describes the position of said system.

5.2 Generalised Coordinates

The smallest number of independent quantities to uniquely describe the position of a system is called the degree of freedom. The independent quantities $\mathbf{q} = [q_1, q_2, \ldots, q_s]^\intercal$ where s is the degree of freedom, are called generalised coordinates of the system. The derivatives of the generalised coordinates $\dot{\mathbf{q}} = [\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_s]^\intercal$ are called its generalised velocities. Furthermore the derivatives of the generalised velocities give us the generalised accelerations $\ddot{\mathbf{q}} = [\ddot{q}_1, \ddot{q}_2, \ldots, \ddot{q}_s]^\intercal$. The relation between the generalised coordinates, velocities and accelerations are called the equations of motion. The equations of motion are second order differential equations, and the number of equations equal the number of degrees of freedom. [11, pp. 1-2]

5.3 The Euler-Lagrange Equations

The motion of any mechanical system can be described either by Hamilton's Principle, also known as The Principle of Least Action, or Newton's Laws of Motion. Newton's Laws are based on forces affecting the system whereas Hamilton's Principle is based on its energies: potential and kinetic. The Principle of Least Action reads that any mechanical system can be characterised by the Lagrangian (Definition 5.2)

$$L(q_1(t), q_2(t), \ldots, q_s(t), \dot{q}_1(t), \dot{q}_2(t), \ldots, \dot{q}_s(t), t),$$

or simply

$$L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t), \tag{5.1}$$

where $\mathbf{q}(t) = [q_1(t), q_2(t), \cdots, q_s(t)]^\intercal$ and $\dot{\mathbf{q}}(t) = [\dot{q}_1(t), \dot{q}_2(t), \cdots, \dot{q}_s(t)]^\intercal$ denote the positions and the velocities of the system, respectively, at the instants $1, 2, \ldots, s$. To ease the reading, the Lagrangian, L, will from now on only be denoted $L(\mathbf{q}, \dot{\mathbf{q}}, t)$, as well as $\mathbf{q}(t)$ and $\dot{\mathbf{q}}(t)$ being denoted only as $\mathbf{q}$

and $\dot{\mathbf{q}}$, respectively. Note that all further information regarding the system can be derived from the positions $\mathbf{q}$ and the velocities $\dot{\mathbf{q}}$ of the system alone, why the Lagrangian only depends on these two conditions. Furthermore, the system must move between positions $\mathbf{q}(t_1)$ and $\mathbf{q}(t_2)$, where $t_2 > t_1$, in such a way that the definite integral

$$S = \int_{t_1}^{t_2} L(\mathbf{q}, \dot{\mathbf{q}}, t)\mathrm{d}t, \tag{5.2}$$

is minimised. The integral (5.2) is called the "action".

The question now is, "how do we minimise S?" For simplicity, let us first assume that the system has only one degree of freedom. This assumption yields that we only determine one function, $\mathbf{q}$. Now, let $\mathbf{q}$ be the function for which S is minimal. This means that S increases when $\mathbf{q}$ is replaced by any other function of the form:

$$\mathbf{q} + \delta\mathbf{q}, \tag{5.3}$$

where $\delta\mathbf{q}$ is a function providing small function values everywhere in the interval (t_1, t_2). For $t = t_1, t_2$, all functions on the form (5.3) must take values $\mathbf{q}(t_1) + \delta\mathbf{q}(t_1) = \mathbf{q}(t_1)$ and $\mathbf{q}(t_2) + \delta\mathbf{q}(t_2) = \mathbf{q}(t_2)$, respectively. This yields the boundary condition:

$$\delta\mathbf{q}(t_1) = \delta\mathbf{q}(t_2) = 0, \tag{5.4}$$

as illustrated by Figure 5.1.

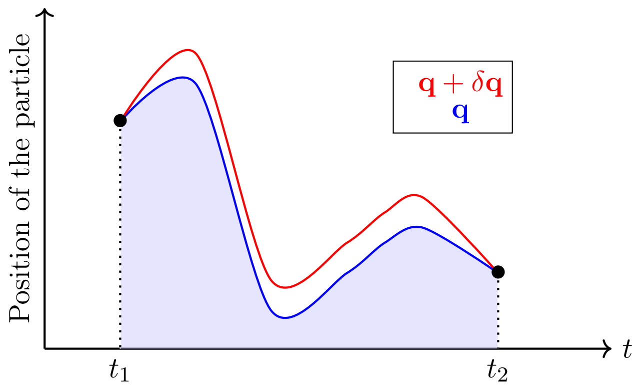


Figure 5.1: Visual representation of the action, (5.2), where the blue graph is $\mathbf{q}$, and the red graph is $\mathbf{q} + \delta\mathbf{q}$. The action, S, is the blue area under the blue graph between limits t_1 and t_2.

We can write the change in the action [18]:

$$\delta S = S(\mathbf{q} + \delta\mathbf{q}) - S(\mathbf{q}) = \int_{t_1}^{t_2} L(\mathbf{q} + \delta\mathbf{q}, \dot{\mathbf{q}} + \delta\dot{\mathbf{q}}, t) - L(\mathbf{q}, \dot{\mathbf{q}}, t)\mathrm{d}t = 0. \tag{5.5}$$

We recognize that $S(\mathbf{q}+\delta\mathbf{q})$ can be rewritten by using a Second-degree Taylor polynomial of a function of two variables [18]:

$$f(x,y) \approx f(a,b) + \frac{\partial}{\partial x} f(a,b)(x-a) + \frac{\partial}{\partial y} f(a,b)(y-a),$$

such that

$$\begin{aligned}
L(\mathbf{q}+\delta\mathbf{q}, \dot{\mathbf{q}}+\delta\dot{\mathbf{q}}, t) &\approx L(\mathbf{q},\dot{\mathbf{q}},t) + \left(\frac{\partial}{\partial \mathbf{q}} L(\mathbf{q},\dot{\mathbf{q}},t)\right)\left(\begin{bmatrix}\mathbf{q}\\ \dot{\mathbf{q}}\end{bmatrix} - \begin{bmatrix}\delta\mathbf{q}\\ \delta\dot{\mathbf{q}}\end{bmatrix}\right) + \left(\frac{\partial}{\partial \dot{\mathbf{q}}} L(\mathbf{q},\dot{\mathbf{q}},t)\right)\left(\begin{bmatrix}\mathbf{q}\\ \dot{\mathbf{q}}\end{bmatrix} - \begin{bmatrix}\delta\mathbf{q}\\ \delta\dot{\mathbf{q}}\end{bmatrix}\right) \\
&\approx L(\mathbf{q},\dot{\mathbf{q}},t) + \left(\left(\frac{\partial}{\partial \mathbf{q}} L(\mathbf{q},\dot{\mathbf{q}},t)\right) + \left(\frac{\partial}{\partial \dot{\mathbf{q}}} L(\mathbf{q},\dot{\mathbf{q}},t)\right)\right)\begin{bmatrix}\mathbf{q}-\delta\mathbf{q}\\ \dot{\mathbf{q}}-\delta\dot{\mathbf{q}}\end{bmatrix} \\
&\approx L(\mathbf{q},\dot{\mathbf{q}},t) + \left(\frac{\partial}{\partial \mathbf{q}} L(\mathbf{q},\dot{\mathbf{q}},t)\right)\left(\mathbf{q}-\delta\mathbf{q}\right) + \left(\frac{\partial}{\partial \dot{\mathbf{q}}} L(\mathbf{q},\dot{\mathbf{q}},t)\right)\left(\dot{\mathbf{q}}-\delta\dot{\mathbf{q}}\right) \\
&\approx L(\mathbf{q},\dot{\mathbf{q}},t) + \left(\frac{\partial}{\partial \mathbf{q}} L\right)\delta\mathbf{q} + \left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right)\delta\dot{\mathbf{q}} \qquad (5.6)
\end{aligned}$$

Replace the new term in (5.6) with the existing in (5.5), and $L(\mathbf{q},\dot{\mathbf{q}},t)$ will be cancelled out.

$$\delta S = \int_{t_1}^{t_2} \underbrace{L(\mathbf{q},\dot{\mathbf{q}},t) - L(\mathbf{q},\dot{\mathbf{q}},t)}_{0} + \frac{\partial}{\partial \mathbf{q}} L\delta\mathbf{q} + \frac{\partial}{\partial \dot{\mathbf{q}}} L\delta\dot{\mathbf{q}}\mathrm{d}t = 0$$

The change in action, denoted δS, can now be seen as the change of the Lagrangian in the $\mathbf{q}$ direction multiplied by the small variation $\delta\mathbf{q}$ as well as varying in $\dot{\mathbf{q}}$ and multiplying by $\delta\dot{\mathbf{q}}$:

$$\delta S = \int_{t_1}^{t_2} \left(\frac{\partial}{\partial \mathbf{q}} L\right)\delta\mathbf{q} + \left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right)\delta\dot{\mathbf{q}}\mathrm{d}t = 0,$$

As seen on Figure 5.1, δS is the area between the red and the blue graph. To obtain the minimal value of S, the variation of S has to be zero. A change in the time derivative of $\mathbf{q}$ is the same as the time derivative of the change:

$$\delta\dot{\mathbf{q}} = \delta\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\delta\mathbf{q}.$$

Hence we can rewrite the second term in δS and obtain:

$$\delta S = \int_{t_1}^{t_2} \left(\frac{\partial}{\partial \mathbf{q}} L\right)\delta\mathbf{q} + \left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right)\frac{\mathrm{d}}{\mathrm{d}t}\delta\mathbf{q}\mathrm{d}t = 0, \qquad (5.7)$$

We use the product rule on the second term in (5.7):

$$f \cdot \frac{\mathrm{d}}{\mathrm{d}t} g = \frac{\mathrm{d}}{\mathrm{d}t}(f\cdot g) - \frac{\mathrm{d}}{\mathrm{d}t}(f)\cdot g$$

with the work shown:

$$f = \left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right), \quad g = \delta\mathbf{q}, \quad \frac{\mathrm{d}}{\mathrm{d}t}(f) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial \dot{\mathbf{q}}} L,\right) \quad \frac{\mathrm{d}}{\mathrm{d}t}(g) = \frac{\mathrm{d}}{\mathrm{d}t}\delta\mathbf{q}.$$

The product rule yields:

$$\left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right)\frac{\mathrm{d}}{\mathrm{d}t}\delta\mathbf{q} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right)\delta\mathbf{q}\right) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right)\delta\mathbf{q}. \qquad (5.8)$$

The rewritten second term, equation 5.8, will replace the second term in (5.7):

$$\delta S = \int_{t_1}^{t_2} \left(\frac{\partial}{\partial \mathbf{q}} L\right) \delta \mathbf{q} + \frac{\mathrm{d}}{\mathrm{d}t} \left(\left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right) \delta \mathbf{q} \right) - \frac{\partial}{\partial t} \left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right) \delta \mathbf{q} \mathrm{d}t = 0,$$

since the middle term is both differentiated and integrated, we can evaluate it at t_1 and t_2 easily:

$$\int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t} \left(\left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right) \delta \mathbf{q} \right) \mathrm{d}t = \left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right) \delta \mathbf{q} \Big|_{t1}^{t2} = 0.$$

We know from the boundary conditions equation (5.4) on $\delta \mathbf{q}$ that on both t_1 and t_2, it equals zero. Therefore the middle term equals zero, thus we have the remaining:

$$\delta S = \int_{t_1}^{t_2} \left(\frac{\partial}{\partial \mathbf{q}} L\right) \delta \mathbf{q} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right) \delta \mathbf{q} \mathrm{d}t = 0.$$

We can take the common $\delta \mathbf{q}$ to obtain:

$$\delta S = \int_{t_1}^{t_2} \left(\frac{\partial}{\partial \mathbf{q}} L\right) - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial}{\partial \dot{\mathbf{q}}} L\right) \delta \mathbf{q} \mathrm{d}t = 0.$$

Since $\delta \mathbf{q}$ is an arbitrary function, the only way the change in the action is guaranteed to equal zero, is if and only if:

$$\frac{\partial}{\partial \mathbf{q}} L - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial \dot{\mathbf{q}}} L = 0.$$

Remember that we assumed the system only had one degree of freedom, such that $s = 1$. The s different functions $q_1, \ldots, q_s$ must be varied independently, hence we can rewrite for a general differential equations for the s unknown functions q_i:

$$\frac{\partial}{\partial q_i} L - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial \dot{\mathbf{q}}_i} L = 0, \quad i = 1, 2, \ldots, s.$$

[11, p. 2]

The Euler-Lagrange Equations evaluate only the conservative forces. They do not account for non-conservative forces like friction and air resistance. If these need to be included, then d'Alembert's principle can be taken into consideration. By d'Alembert's Principle, the Euler-Lagrange Equation can be written as

$$\frac{\partial}{\partial q_i} L - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial \dot{q}_i} L = Q_i, \quad i = 1, 2, \ldots, n,$$

where Q_i is the sum of non-conservative forces. [17, pp. 98-99]

5.4 Equations of Motion for the Pendulum on a Cart

Having introduced the Lagrangian and the Euler-Lagrange Equations, we can now derive the equations of motion for the inverted pendulum on a cart. For simplicity of the derivations, assume that the pendulum has a point of mass m at the end of a massless, rigid rod of length l, and a cart of mass M (a point mass in the contact point between the pendulum rod and the cart). Now assume friction in the contact point between the pendulum rod and the cart, as well as between the cart and the rail (the potential zero line). Hence, we illustrate the system in Figure 5.2. Note that the placement of the cart is not restricted to the origin of the coordinate system, but can move freely along the x-axis.

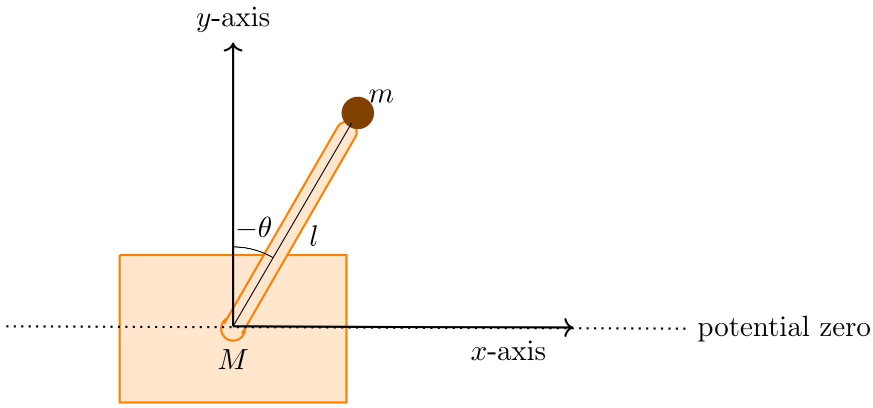


Figure 5.2: Schematic of the inverted pendulum on a cart.

Having introduced our coordinate system, we can now determine the position vectors of the cart, $\mathbf{r}_c$, and the pendulum, $\mathbf{r}_p$:

$$\mathbf{r}_c = \begin{bmatrix} x(t) & 0 \end{bmatrix}^\intercal,$$
$$\mathbf{r}_p = \begin{bmatrix} x(t) + l\sin(-\theta(t)) & l\cos(-\theta(t)) \end{bmatrix}^\intercal = \begin{bmatrix} x(t) - l\sin(\theta(t)) & l\cos(\theta(t)) \end{bmatrix}^\intercal.$$

The coordinates of the vectors are made obvious by Figure 5.3.

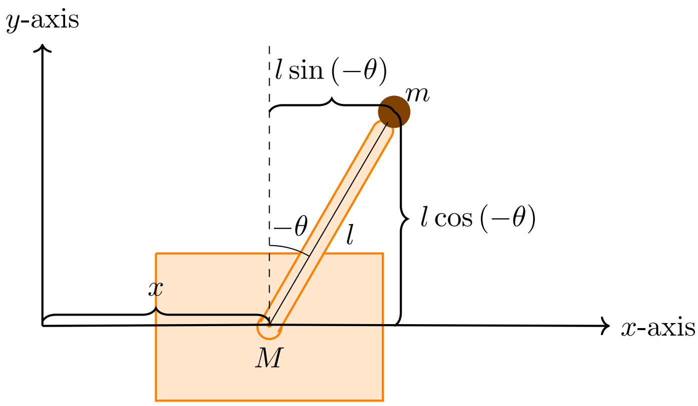


Figure 5.3: Visual representation of the coordinates of the pendulum and the cart.

To ease further reading, $x(t)$ and $\theta(t)$ will be referred to only as x and θ, respectively.

We first look at the Lagrangian

$$L = \sum_{i=1}^{n} \frac{1}{2} m_i v_i^2 - U(r_1, r_2, \dots, r_n).$$

We determine the kinetic energy T:

$$T = \frac{1}{2}M\langle \mathbf{v}_c|\mathbf{v}_c\rangle + \frac{1}{2}m\langle \mathbf{v}_p|\mathbf{v}_p\rangle, \tag{5.9}$$

per the remark for Definition 5.1 where $\mathbf{v}_c$ is the velocity vector of the cart and $\mathbf{v}_p$ is the velocity vector of the pendulum. As the movement of the cart is restricted to the x-axis, we obtain

$$\mathbf{v}_c = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{r}_c = \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x & 0 \end{bmatrix}^\intercal = \begin{bmatrix} \dot{x} & 0 \end{bmatrix}^\intercal .$$

The velocity vector of the pendulum is the time derivative of the position vector of the pendulum and is dependent on the angle θ and the displacement of the cart:

$$\mathbf{v}_p = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{r}_p = \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x - l\sin(\theta) & l\cos(\theta) \end{bmatrix}^\intercal = \begin{bmatrix} \dot{x} - l\cos(\theta)\dot{\theta} & -l\sin(\theta)\dot{\theta} \end{bmatrix}^\intercal .$$

We now calculate $\langle \mathbf{v}_p|\mathbf{v}_p\rangle$:

$$\begin{aligned}
\langle \mathbf{v}_p|\mathbf{v}_p\rangle &= (\dot{x} - l\cos(\theta)\dot{\theta})^2 + (-l\sin(\theta)\dot{\theta})^2 \\
&= \dot{x}^2 + l^2\cos^2(\theta)\dot{\theta}^2 - 2l\cos(\theta)\dot{\theta}\dot{x} + l^2\sin^2(\theta)\dot{\theta}^2 \\
&= \dot{x}^2 + l^2\dot{\theta}^2\underbrace{(\cos^2(\theta) + \sin^2(\theta))}_{1} - 2l\cos(\theta)\dot{\theta}\dot{x} \\
&= \dot{x}^2 + l^2\dot{\theta}^2 - 2l\cos(\theta)\dot{\theta}\dot{x}.
\end{aligned}$$

We can now insert the calculated expressions for the velocities $\mathbf{v}_c^2$ and $\mathbf{v}_p^2$ in (5.9) and we obtain:

$$\begin{aligned}
T &= \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m(\dot{x}^2 + l^2\dot{\theta}^2 - 2l\cos(\theta)\dot{\theta}\dot{x}) \\
&= \frac{1}{2}(M+m)\dot{x}^2 + \frac{1}{2}ml^2\dot{\theta}^2 - ml\cos(\theta)\dot{\theta}\dot{x}. \qquad (5.10)
\end{aligned}$$

We now look at the potential energy U of the system. It is only the gravitational potential energy which will affect the pendulum since the zero potential of the system is located along the cart's rail (see Figure 5.2), hence the potential energy of the cart equaling zero. Therefore we get:

$$\begin{aligned}
U &= mgh, \qquad \text{where } h = l\cos(\theta) \\
&= mlg\cos(\theta). \qquad (5.11)
\end{aligned}$$

By insertion of (5.10) and (5.11), the Lagrangian of the system is as follows,

$$L = \frac{1}{2}(M+m)\dot{x}^2 + \frac{1}{2}ml^2\dot{\theta}^2 - ml\cos(\theta)\dot{\theta}\dot{x} - mlg\cos(\theta).$$

5.4.1 Equations of Motion

The equations of motion can now be found, by using the Euler-Lagrange Equations as well as d'Alemberts Principle. As the system has two degrees of freedom, we get:

$$\frac{\partial L}{\partial q_1} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q_1}} = Q_1,$$
$$\frac{\partial L}{\partial q_2} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q_2}} = Q_2,$$

where Q_1 and Q_2 are the sums of the generalised forces acting on the system. As mentioned earlier, the coordinates necessary for describing the position of the pendulum are the displacement of the cart x and the angle relative to the y-axis θ. Q_1 and Q_2 are seen as the external, non-conservative forces, i.e. friction, for the cart and the pendulum. We define the friction as viscosity velocity and we get the following expressions

$$\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} = \mathbf{u}(t) - \mu_1\dot{x}, \qquad \text{where } \mathbf{u}(t) \text{ is the input,}$$
$$\frac{\partial L}{\partial \theta} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\theta}} = -\mu_2\dot{\theta}.$$

Here, $\mathbf{u}(t)$ acts as an external force we apply in order to force the system into equilibrium (which will be explored further in Chapter 6). To ease the reading, $\mathbf{u}(t)$ will be denoted as $\mathbf{u}$.
Remark (1): In our calculations, $\mathbf{u}$ is a one-dimensional vector which is equivalent to a scalar. This is shown in Chapter 8.
Remark (2): $-\mu_1\dot{x}$ describes the friction between the rail and the cart, and$-\mu_2\dot{\theta}$ is the friction between the pendulum rod and the cart.

First we look at:

$$\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} = \mathbf{u} - \mu_1\dot{x}.$$

On the left-hand side, we will look at the individual parts of the Euler-Lagrange equation, and from the first term get:

$$\frac{\partial L}{\partial x} = 0,$$

as x is not a part of L.

Next we look at the second term $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}}$, where the parts that do not involve $\dot{x}$ are zero and therefore, are not included. This continues throughout this section of deriving the equations of motion.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{x}}\left(\frac{1}{2}(M+m)\dot{x}^2 - ml\cos(\theta)\dot{x}\dot{\theta}\right)\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left((M+m)\dot{x} - ml\cos(\theta)\dot{\theta}\right).\end{aligned}$$

When differentiating $\cos(\theta)\dot{\theta}$ we apply the product rule and the chain rule respectively:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} = (M+m)\ddot{x} - ml\frac{\mathrm{d}}{\mathrm{d}t}\cos(\theta)\dot{\theta}).$$

First the product rule:

$$w = \dot{\theta}, \quad \dot{w} = \ddot{\theta}, \quad v = \cos(\theta).$$

To obtain $\frac{\mathrm{d}}{\mathrm{d}t}v$ we need to apply the chain rule:

$$f = \cos(\theta), \quad \frac{\mathrm{d}}{\mathrm{d}t}f = -\sin(\theta), \quad g = \theta, \quad \frac{\mathrm{d}}{\mathrm{d}t}g = \dot{\theta}, \quad \text{hence, } \frac{\mathrm{d}}{\mathrm{d}t}v = (-\sin(\theta)\dot{\theta}).$$

Thus by completing the product rule, differentiating $\cos(\theta)\dot{\theta}$ gives us:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\cos(\theta)\dot{\theta}) = \ddot{\theta}\cos(\theta) + \dot{\theta}(-\sin(\theta)\dot{\theta}).$$

The second term is:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} = (M+m)\ddot{x} - ml\cos(\theta)\ddot{\theta} - \sin(\theta)\dot{\theta}^2.$$

We can now write the full Euler-Lagrange equation for the displacement

$$\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} = -(M+m)\ddot{x} + ml\cos(\theta)\ddot{\theta} - ml\sin(\theta)\dot{\theta}^2 = \mathbf{u} - \mu_1\dot{x}. \tag{5.12}$$

Then we can isolate $\ddot{x}$ in (5.12)

$$\begin{aligned}
-(M+m)\ddot{x} + ml\cos(\theta)\ddot{\theta} - ml\sin(\theta)\dot{\theta}^2 &= \mathbf{u} - \mu_1\dot{x} \\
-(M+m)\ddot{x} &= -ml\cos(\theta)\ddot{\theta} + ml\sin(\theta)\dot{\theta}^2 - \mu_1\dot{x} + \mathbf{u} \\
\ddot{x} &= \frac{ml\cos(\theta)}{M+m}\ddot{\theta} - \frac{ml\sin(\theta)}{M+m}\dot{\theta}^2 + \frac{\mu_1}{M+m}\dot{x} - \frac{\mathbf{u}}{M+m}.
\end{aligned} \tag{5.13}$$

Now we look at $\frac{\partial L}{\partial \theta} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\theta}} = -\mu_2\dot{\theta}$ to find an expression for $\ddot{\theta}$, with the same method. We focus on the left-hand side, and complete our first term:

$$\begin{aligned}
\frac{\partial L}{\partial \theta} &= \frac{\partial}{\partial \theta}(-ml\cos(\theta)\dot{x}\dot{\theta} - mlg\cos(\theta)) \\
&= ml\sin(\theta)\dot{x}\dot{\theta} + mlg\sin(\theta).
\end{aligned} \tag{5.14}$$

Then the second term:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\theta}} &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{\theta}}\left(\frac{1}{2}ml^2\dot{\theta}^2 - ml\cos(\theta)\dot{x}\dot{\theta}\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(ml^2\dot{\theta} - ml\cos(\theta)\dot{x}\right) \\
&= ml^2\ddot{\theta} - ml\cos(\theta)\ddot{x} + ml\sin(\theta)\dot{x}\dot{\theta}.
\end{aligned} \tag{5.15}$$

We can now insert (5.14) and (5.15) in $\frac{\partial L}{\partial \theta} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\theta}} = -\mu_2\dot{\theta}$ and isolate $\ddot{\theta}$:

$$\begin{aligned}
ml\sin(\theta)\dot{x}\dot{\theta} + mlg\sin(\theta) - ml^2\ddot{\theta} + ml\cos(\theta)\ddot{x} - ml\sin(\theta)\dot{x}\dot{\theta} &= -\mu_2\dot{\theta} \\
\underbrace{ml\sin(\theta)\dot{x}\dot{\theta} - ml\sin(\theta)\dot{x}\dot{\theta}}_{0} + mlg\sin(\theta) + ml\cos(\theta)\ddot{x} + \mu_2\dot{\theta} &= ml^2\ddot{\theta} \\
\frac{g}{l}\sin(\theta) + \frac{\cos(\theta)}{l}\ddot{x} + \frac{\mu_2}{ml^2}\dot{\theta} &= \ddot{\theta}.
\end{aligned} \tag{5.16}$$

We can then see that $\ddot{x}$ and $\ddot{\theta}$ are interdependent, and we wish to make them independent. Therefore, we insert (5.16) into (5.13). For simplicity we define $\alpha = \frac{ml}{M+m}$. Thus we get

$$\begin{aligned}
\ddot{x} &= \alpha\cos(\theta)\left(\frac{\cos(\theta)}{l}\ddot{x} + \frac{g}{l}\sin(\theta) + \frac{\mu_2}{ml^2}\dot{\theta}\right) - \alpha\sin(\theta)\dot{\theta}^2 + \frac{\mu_1}{M+m}\dot{x} - \frac{\mathbf{u}}{M+m} \\
&= \frac{\alpha}{l}\cos^2(\theta)\ddot{x} + \frac{\alpha g}{l}\cos(\theta)\sin(\theta) + \alpha\frac{\mu_2}{ml^2}\cos(\theta)\dot{\theta} - \alpha\sin(\theta)\dot{\theta}^2 + \frac{\mu_1}{M+m}\dot{x} - \frac{\mathbf{u}}{M+m} \\
\ddot{x} - \frac{\alpha}{l}\cos^2(\theta)\ddot{x} &= \frac{\alpha g}{l}\cos(\theta)\sin(\theta) + \alpha\frac{\mu_2}{ml^2}\cos(\theta)\dot{\theta} - \alpha\sin(\theta)\dot{\theta}^2 + \frac{\mu_1}{M+m}\dot{x} - \frac{\mathbf{u}}{M+m} \\
\ddot{x}\left(1 - \frac{\alpha}{l}\cos^2(\theta)\right) &= \frac{\alpha g}{l}\cos(\theta)\sin(\theta) + \alpha\frac{\mu_2}{ml^2}\cos(\theta)\dot{\theta} - \alpha\sin(\theta)\dot{\theta}^2 + \frac{\mu_1}{M+m}\dot{x} - \frac{\mathbf{u}}{M+m} \\
\ddot{x} &= \frac{\frac{\alpha g}{l}\cos(\theta)\sin(\theta) + \alpha\frac{\mu_2}{ml^2}\cos(\theta)\dot{\theta} - \alpha\sin(\theta)\dot{\theta}^2 + \frac{\mu_1}{M+m}\dot{x} - \frac{\mathbf{u}}{M+m}}{\left(1 - \frac{\alpha}{l}\cos^2(\theta)\right)}.
\end{aligned}$$

We now re-insert $\alpha = \frac{ml}{M+m}$ and get

$$\begin{aligned}
\ddot{x} &= \frac{\frac{mlg}{(M+m)l}\cos(\theta)\sin(\theta) + \frac{ml\mu_2}{(M+m)ml^2}\cos(\theta)\dot{\theta} - \frac{ml}{M+m}\sin(\theta)\dot{\theta}^2 + \frac{\mu_1}{M+m}\dot{x} - \frac{\mathbf{u}}{M+m}}{\left(1 - \frac{ml}{(M+m)l}\cos^2(\theta)\right)} \\
&= \frac{mg\cos(\theta)\sin(\theta) + \frac{\mu_2}{ml}\cos(\theta)\dot{\theta} - ml\sin(\theta)\dot{\theta}^2 + \mu_1\dot{x} - \mathbf{u}}{M+m-m\cos^2(\theta)}.
\end{aligned} \tag{5.17}$$

Then we can insert (5.17) into (5.16) and we get

$$\begin{aligned}
\ddot{\theta} &= \frac{\cos(\theta)}{l}\left(\frac{mg\cos(\theta)\sin(\theta) + \frac{\mu_2}{ml}\cos(\theta)\dot{\theta} - ml\sin(\theta)\dot{\theta}^2 + \mu_1\dot{x} - \mathbf{u}}{M+m-m\cos^2(\theta)}\right) + \frac{g}{l}\sin(\theta) + \frac{\mu_2}{ml^2}\dot{\theta} \\
&= \frac{mg\cos^2(\theta)\sin(\theta) + \frac{\mu_2}{ml}\cos^2(\theta)\dot{\theta} - ml\cos(\theta)\sin(\theta)\dot{\theta} + \cos(\theta)\mu_1\dot{x} - \mathbf{u}\cos(\theta)}{(M+m)l - ml\cos^2(\theta)} + \frac{g}{l}\sin(\theta) + \frac{\mu_2}{ml^2}\dot{\theta}.
\end{aligned} \tag{5.18}$$

Thus we have obtained two independent equations of motion, (5.17), and (5.18), one for the cart and one for the pendulum, both with viscous friction.

Chapter 6

Stabilisation

This chapter assesses a range of prerequisites of stabilisation, including vector fields and phase portraits with respect to the solutions to linear autonomous systems. The chapter is based on [2] and [8].

6.1 Linear Autonomous Systems

When considering an autonomous system of ODE's, there is a unique solution to each different initial condition (see Theorem 3.1). To get a complete view of the solutions to these systems, we have to look at vector fields as they will be used to make a graphical representation of the behaviour of the solutions.

Definition 6.1 (Vector Fields)
Let $\mathcal{O} \subseteq \mathbb{R}^n$ be an open set. A vector field on $\mathcal{O}$ is a function $\mathbf{f} : \mathcal{O} \to \mathbb{R}^n$ given as

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{bmatrix},$$

where $\mathbf{x} \in \mathcal{O}$ and $f_i : \mathcal{O} \to \mathbb{R}$ for $i = 1, 2, \ldots, n$ are called the components or element functions of $\mathbf{f}$. [2, p. 17, Definition 1.1]

A vector field consists of vectors with varying directions and lengths, which correspond to the magnitude of the vector. Vector fields can be used in addition to solution curves, which is a convenient way of viewing solutions of systems of differential equations. Such a curve $\boldsymbol{\alpha}$ is also called an integral curve of the vector field $\mathbf{f}$.

Definition 6.2 (Solutions of Autonomous Systems)
A curve in $\mathbb{R}^n$ is a function $\boldsymbol{\alpha} : I \to \mathbb{R}^n$ where $I \subset \mathbb{R}^n$. If $\boldsymbol{\alpha}$ is differentiable, then it is called a differentiable curve. If $\mathbf{f}$ is a vector field on $\mathcal{O}$, then a solution of the system $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \mathbf{f}(\mathbf{x})$ is a differentiable curve $\boldsymbol{\alpha}$ with the properties:

1. $\boldsymbol{\alpha}(t) \in \mathcal{O} \forall t \in I$,
2. $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\alpha}(t) = \mathbf{f}(\boldsymbol{\alpha}(t)) \forall t \in I$.

[2, p. 19, Definition 1.2]

Remark for Definition 6.2 Property 1: The curve $\boldsymbol{\alpha}$ lies in the open set $\mathcal{O}$ for all $t \in I$, which is necessary for $\mathbf{f}(\boldsymbol{\alpha}(t))$ in property 2 as $\mathbf{f}(\mathbf{x})$ is only defined at points $\mathbf{x}$ in $\mathcal{O}$.
Remark for Definition 6.2 Property 2: $\boldsymbol{\alpha}$ satisfies the system of ODE's, and therefore we can write it as a system of first-order ODE's:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\alpha_1(t) &= f_1(\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t)) \\
\frac{\mathrm{d}}{\mathrm{d}t}\alpha_2(t) &= f_2(\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t)) \\
&\vdots \\
\frac{\mathrm{d}}{\mathrm{d}t}\alpha_n(t) &= f_n(\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t)),
\end{aligned}$$

for all $t \in I$.

There is one (or more) specific point(s) of the vector field and solution curves which is(are) of significant importance. Such a point is called an equilibrium point.

Definition 6.3 (Equilibrium Points)
Consider an autonomous system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{6.1}$$

where $\mathbf{f} : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a locally Lipschitz function. Then $\overline{\mathbf{x}}$ is called an equilibrium point of $\mathbf{f}$ if

$$\mathbf{f}(\overline{\mathbf{x}}) = \mathbf{0}.$$

[8, p. 112]

When positioned in an equilibrium point, the system is at rest. If it is positioned near an equilibrium point, we look at whether it moves towards the equilibrium point or away from it. Thus, we define stability of an equilibrium point.

Definition 6.4 (Stability of an Equilibrium Point)
Let $\overline{\mathbf{x}} = \mathbf{0}$ be an equilibrium point of an autonomous system. Then the following applies for equilibrium points:

- It is stable if

$$\forall \epsilon > 0 \exists \delta > 0 : \|\overline{\mathbf{x}}(0)\| < \delta \Rightarrow \|\overline{\mathbf{x}}(t)\| < \epsilon, \quad \forall t \geq 0.$$

- It is unstable if it is not stable.

- It is asymptotically stable if it is stable and

$$\exists \delta > 0 : \|\overline{\mathbf{x}}(0)\| < \delta \Rightarrow \lim_{t \to \infty} \overline{\mathbf{x}}(t) = 0.$$

[8, p. 112]

A visual representation of Definition 6.4 can be seen on Figure 6.1, in which we have a system with an equilibrium point $\overline{\mathbf{x}} = 0$. If we choose a point $\overline{\mathbf{x}}(0)$ within δ, we see, whether it stays within ϵ or not. If it stays within ϵ, it is stable, and if it converges towards $\overline{\mathbf{x}}$, it is asymptotically stable. Otherwise, if it does not stay within ϵ, it is unstable.

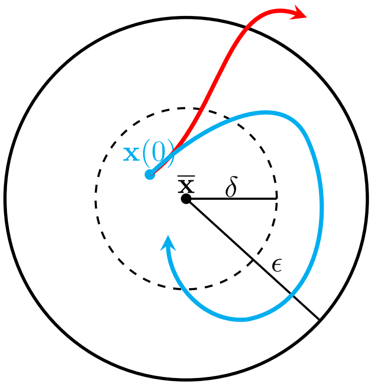


Figure 6.1: Stability of an equilibrium point. The red curve is an unstable point, and the cyan curve is a stable point.

Now consider the linear autonomous system

$$\dot{\mathbf{x}} = A\mathbf{x}. \tag{6.2}$$

To determine if this system has multiple equilibria, we have to look at $\det A$. If $\det A \neq 0$ then $\overline{\mathbf{x}} = \mathbf{0}$ is the only equilibrium of (6.2), represented by the origin, $(0, 0)$. If $\det A = 0$, then the null space of A is an equilibrium subspace for the system. We can determine when a system on the form (6.2) is stable.

Theorem 6.1 (Stability of a Linear System)
Let $\overline{\mathbf{x}} \in \mathbb{R}^n$ be an equilibrium point of an autonomous linear system as

$$\dot{\mathbf{x}} = A\mathbf{x}.$$

The system is stable if and only if the corresponding eigenvalues λ_i of A follow the requirements:

- $\text{Re}(\lambda_i) \leq 0$ for $i = 1, 2, \ldots, n$.
- If there exists two or more eigenvalues which satisfy $\text{Re}(\lambda_i) = 0$, then $\text{rank}(A - \lambda_i I) = n - q_i$, where $q_i \geq 2$ is the algebraic multiplicity and n is the dimension of $\overline{\mathbf{x}}$.

[8, p. 134]

6.2 Phase Portraits

This section is based on [2, pp. 212-214] and [8, pp. 23-30].
To determine if an autonomous system is stable or unstable, we can also look at phase portraits, which gives a visual understanding. Phase portraits connect the theory we have about vector fields, integral curves and stability of an equilibrium point $\overline{\mathbf{x}} = \mathbf{0}$. We want to look at the autonomous system

$$\dot{\mathbf{x}} = A\mathbf{x},$$

where $A \in \mathbb{R}^{2\times 2}$. The matrix A will have two eigenvalues. These eigenvalues can either be real or complex, where the real eigenvalues can be divided into two different groups: distinct or repeated eigenvalues.

6.2.1 Real Distinct Eigenvalues

The real eigenvalues $\lambda_1 \neq \lambda_2$ can be divided into three cases:

1. $\lambda_1 < 0 < \lambda_2$: In this case, λ_1 is called the stable eigenvalue and λ_2 is called the unstable eigenvalue. The equilibrium point is called a saddle point, which is always unstable. See Figure 6.2b.

2. $\lambda_1 < \lambda_2 < 0$: In this case, λ_1 is called the fast eigenvalue and λ_2 is called the slow eigenvalue. The phase portrait is called a stable proper node, and since the solution curves tend toward the equilibrium point, the system is asymptotically stable.

3. $0 < \lambda_1 < \lambda_2$: In this case, all solutions tend away from the equilibrium point, and goes to ∞ when $t \to \infty$. The phase portrait is then called an unstable proper node.

6.2.2 Real Repeated Eigenvalues

The phase portrait of a system with repeated eigenvalues depends on the sign of the repeated real eigenvalue.

1. $\lambda_1 = \lambda_2 < 0$: In this case, the solutions move toward the equilibrium point, which is asymptotically stable. This phase portrait is called an stable improper node, see Figure 6.2d.

2. $\lambda_1 = \lambda_2 > 0$: In this case, the solutions move away from the equilibrium point, which is unstable. This phase portrait is called an unstable improper node.

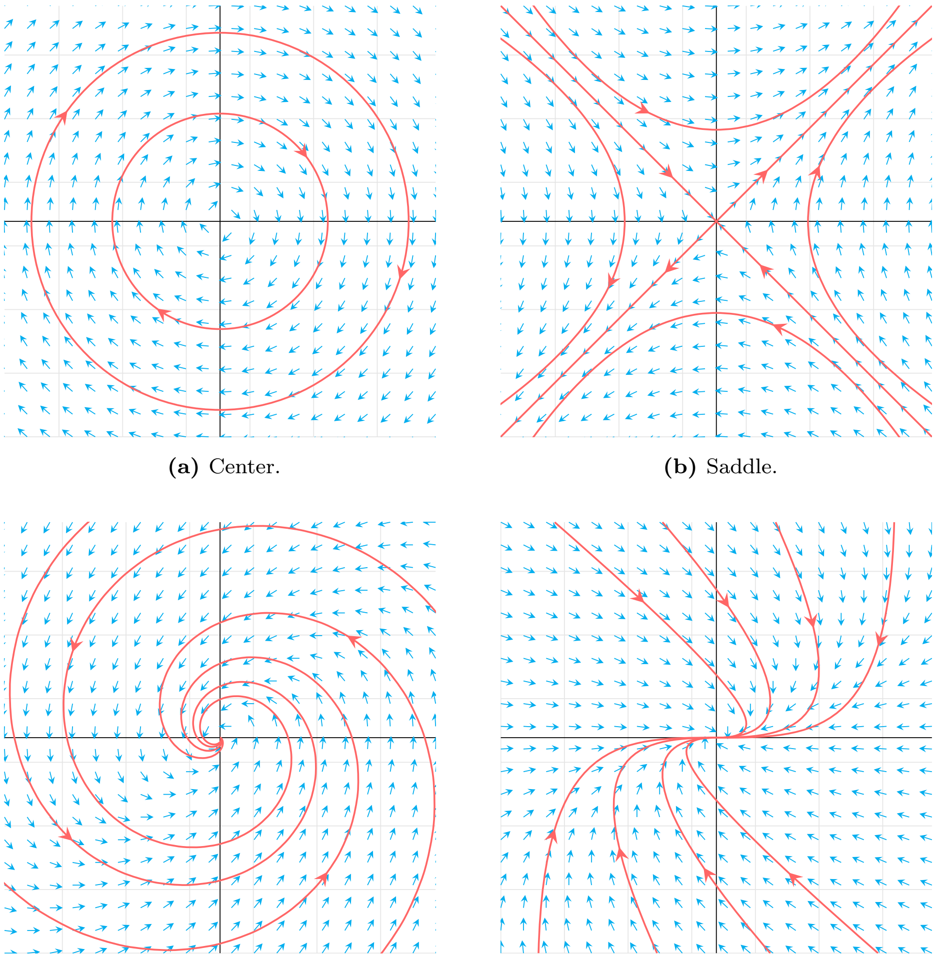
6.2.3 Complex Eigenvalues

The last case is complex eigenvalues on the form $\lambda_{1,2} = \alpha \pm j\beta$. There are three different cases of complex eigenvalues, where the phase portrait of a system depends on the value of the real part of the eigenvalues:

1. $\mathrm{Re}(\lambda_{1,2}) = 0$: In this case, the eigenvalues are completely imaginary. The solutions then form ellipses or circles around the equilibrium point. If $\beta < 0$ the circles move in the counterclockwise direction, and if $\beta > 0$ the circles move in the clockwise direction, see Figure 6.2a. This phase portrait is called a center, and it is stable.

2. $\mathrm{Re}(\lambda_{1,2}) < 0$: The solutions are spirals and these move toward the equilibrium point. The phase portrait is called a stable spiral, see Figure 6.2c.

3. $\mathrm{Re}(\lambda_{1,2}) > 0$: This case is the opposite of case (2). The solutions are also spirals, but instead of moving toward the equilibrium point, they move away from it. The phase portrait is then called an unstable spiral.



(a) Center. **(b)** Saddle. **(c)** Spiral. **(d)** Improper node.

Figure 6.2: Four different phase portraits that each describe stability and eigenvalues for a system.

If we look at a phase portrait for a pendulum, as per Figure 6.3, we can tell, there are multiple equilibria. The stable equilibrium point is a center point, where the eigenvalues for that point would be completely imaginary and $\beta > 0$. The unstable equilibrium is a saddle point, where the eigenvalues for that point would be real distinct eigenvalues of opposite signs.

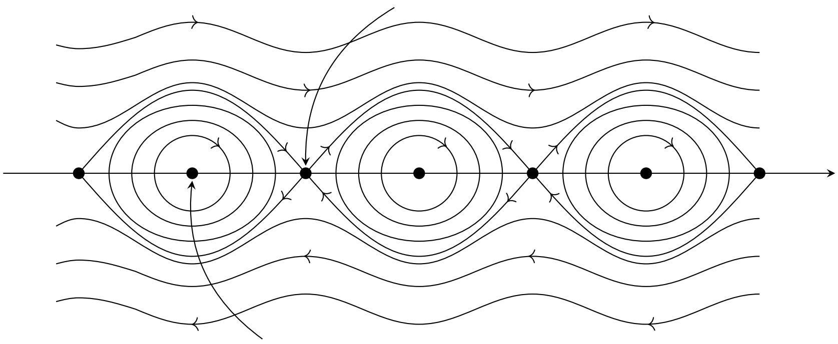


Figure 6.3: A phase portrait of a system with stable and unstable equilibrium points.

6.3 Linearisation of Non-linear Autonomous Systems

In this project, the equations of motion are non-linear differential equations. Therefore we will be using linearisation to approximate non-linear autonomous systems using linear systems, as we already know how to analyse linear systems. The linearised model of the non-linear system can then be used to determine the stability of the non-linear system in the equilibrium point $\overline{\mathbf{x}}$. First we will define the Jacobian matrix.

Definition 6.5 (Jacobian Matrix)
Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous, differentiable function. Let $\overline{\mathbf{x}} \in \mathbb{R}^n$ be an equilibrium point. Then, the Jacobian Matrix $J_{\mathbf{f}(\overline{\mathbf{x}})} \in \mathbb{R}^{n\times n}$ of $\mathbf{f}(\mathbf{x})$ evaluated at $\overline{\mathbf{x}}$ is

$$J_{\mathbf{f}(\overline{\mathbf{x}})} = \frac{\partial}{\partial \mathbf{x}}\mathbf{f}(\overline{\mathbf{x}}) = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \frac{\partial \mathbf{f}}{\partial x_2} & \cdots & \frac{\partial \mathbf{f}}{\partial x_n} \end{bmatrix}\Bigg|_{\mathbf{x}=\overline{\mathbf{x}}} = \left.\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}\right|_{\mathbf{x}=\overline{\mathbf{x}}}$$

[8, p. 52]

The eigenvalues of the Jacobian matrix are used to determine if an equilibrium point is stable or unstable for a non-linear system.

Theorem 6.2 (Lyapunov's Indirect Theorem)
Let $\overline{\mathbf{x}}$ be an equilibrium point for the non-linear autonomous system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, where $\mathbf{f} : \mathcal{O} \to \mathbb{R}^n$ is continuous differentiable and $\mathcal{O} \subset \mathbb{R}^n$ is a neighbourhood of $\overline{\mathbf{x}}$. Let $J_{\mathbf{f}(\overline{\mathbf{x}})}$ be the Jacobian matrix of the linearised system evaluated at $\overline{\mathbf{x}} = \mathbf{0}$. Then:

- The equilibrium point is asymptotically stable if $\mathrm{Re}(\lambda_i) < 0$ for all eigenvalues of $J_{\mathbf{f}(\overline{\mathbf{x}})}$.
- The equilibrium point is unstable if $\mathrm{Re}(\lambda_i) > 0$ for one or more eigenvalues of $J_{\mathbf{f}(\overline{\mathbf{x}})}$.

[8, p. 127]

Now consider the non-linear system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \tag{6.3}$$

where $\mathbf{f} : \mathcal{O} \to \mathbb{R}^n$ is a continuously differentiable function and $\mathcal{O} \subset \mathbb{R}^n$. Let $\overline{\mathbf{x}} = (\overline{x}_1, \ldots, \overline{x}_n)$ be an equilibrium point, then the Taylor Series Expansion evaluated at the point $\overline{\mathbf{x}}$ becomes

$$\begin{aligned}
\dot{x}_1 &\approx f_1(\overline{\mathbf{x}}) + \frac{\partial f_1(\overline{\mathbf{x}})}{\partial x_1}(x_1 - \overline{x}_1) + H.O.T \\
&\vdots \\
\dot{x}_n &\approx f_n(\overline{\mathbf{x}}) + \frac{\partial f_n(\overline{\mathbf{x}})}{\partial x_n}(x_n - \overline{x}_n) + H.O.T,
\end{aligned}$$

where the H.O.T (stands for higher-order terms) are non-linear terms. Therefore, they will be ignored. Since $\overline{\mathbf{x}}$ is an equilibrium point, $\mathbf{f}(\overline{\mathbf{x}}) = 0$. We also denote $w_i = (x_i - \overline{x}_i)$ for $i = 1, 2, \ldots, n$. The linearisation of the system then becomes

$$\begin{aligned}
\dot{x}_1 &= \frac{\partial f_1(\overline{\mathbf{x}})}{\partial x_1} w_1 \\
&\vdots \\
\dot{x}_n &= \frac{\partial f_n(\overline{\mathbf{x}})}{\partial x_n} w_n.
\end{aligned}$$

This can be written in matrix notation as

$$\begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1(\overline{\mathbf{x}})}{\partial x_1} & \cdots & \frac{\partial f_1(\overline{\mathbf{x}})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(\overline{\mathbf{x}})}{\partial x_1} & \cdots & \frac{\partial f_n(\overline{\mathbf{x}})}{\partial x_n} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}.$$

The matrix

$$J_{\mathbf{f}(\overline{\mathbf{x}})} = \begin{bmatrix} \frac{\partial f_1(\overline{\mathbf{x}})}{\partial x_1} & \cdots & \frac{\partial f_1(\overline{\mathbf{x}})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(\overline{\mathbf{x}})}{\partial x_1} & \cdots & \frac{\partial f_n(\overline{\mathbf{x}})}{\partial x_n} \end{bmatrix}$$

is the Jacobian matrix at the equilibrium $\overline{\mathbf{x}}$.

Chapter 7

Linear Control Systems

This chapter introduces control theory, which is used to determine how to stabilise and control an autonomous system with some input based on the state of the system. Such non-linear system can be described as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}). \tag{7.1}$$

If $\mathbf{w} = (\mathbf{x} - \overline{\mathbf{x}})$ and $\mathbf{v} = (\mathbf{u} - \overline{\mathbf{u}})$, a linearisation of the system (7.1) about the equilibrium point $(\overline{\mathbf{x}}, \overline{\mathbf{u}})$ results in the linear input-system

$$\dot{\mathbf{x}} = \frac{\partial \mathbf{f}(\overline{\mathbf{x}}, \overline{\mathbf{u}})}{\partial \mathbf{x}} \mathbf{w} + \frac{\partial \mathbf{f}(\overline{\mathbf{x}}, \overline{\mathbf{u}})}{\partial \mathbf{u}} \mathbf{v}.$$

We will examine an input-output system denoted Σ:

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B\mathbf{u}(t), \tag{7.2}$$

$$\mathbf{y}(t) = C\mathbf{x}(t). \tag{7.3}$$

A, B, and C are matrices where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$ and $C \in \mathbb{R}^{p \times n}$. The input is described by a vector $\mathbf{u}(t) \in \mathbb{R}^m$ and the set of all admissible continuous input functions to the system is denoted as $\mathbf{U}$. The state variable vector $\mathbf{x}(t)$ takes value in $\mathbb{R}^n$. The state variables are the minimum set of variables that fully describe the system, and therefore have enough information to predict future behaviour. Finally $\mathbf{y}(t)$ is the output vector of the system and it takes value in $\mathbb{R}^p$.

It is possible to model the input of the system as a controller

$$\mathbf{u}(t) = -K\mathbf{x}(t) \tag{7.4}$$

where $K \in \mathbb{R}^{m \times n}$, which let us control the system if the system is controllable. Using the controller $\mathbf{u}(t) = -K\mathbf{x}(t)$ we can rewrite the system as

$$\dot{\mathbf{x}}(t) = (A - BK)\mathbf{x}(t). \tag{7.5}$$

The rewriting gives us a closed loop system matrix A with feedback control. And with the matrix K, we can control and change the eigenvalues of $A - BK$. It is possible to increase or decrease control according to how displaced the system is compared to where you want it to be. The closed loop system will measure the output $\mathbf{y}(t)$ continuously and feed the information back to the input. The feedback compensates for the disturbance and improves the accuracy of the system. Figure 7.1 shows a closed loop system with a controller $-K$ inserted.

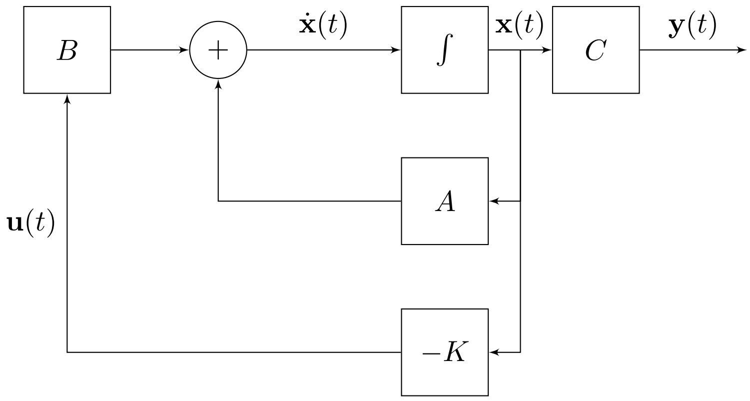


Figure 7.1: Closed loop system with controller $-K$.

The equations (7.2) and (7.3) for the system Σ can be solved as differential equations. The solution for $\dot{\mathbf{x}}(t)$ with the initial condition $\mathbf{x}(0) = \mathbf{x}_0$ is:

$$\mathbf{x}_{\mathbf{u}}(t, \mathbf{x}_0) = \mathrm{e}^{At}\mathbf{x}_0 + \int_0^t \mathrm{e}^{A(t-\tau)} B\mathbf{u}(\tau)\mathrm{d}\tau. \tag{7.6}$$

Hence the corresponding solution for $\mathbf{y}(t)$ is:

$$\mathbf{y}_{\mathbf{u}}(t, \mathbf{x}_0) = C\mathrm{e}^{At}\mathbf{x}_0 + \int_0^t C\mathrm{e}^{A(t-\tau)}\mathbf{u}(\tau)d\tau.$$

[6, pp. 37-38].

Definition 7.1 (Controllability)
The system $\dot{\mathbf{x}} = A\mathbf{x}(t) + B\mathbf{u}(t)$ is said to be controllable in time T if for any pair $(\mathbf{x}_0, \mathbf{x}_1)$ of states there exists an input $\mathbf{u} \in \mathbf{U}$ such that the corresponding solution satisfies $\mathbf{x}(0) = \mathbf{x}_0$ and $\mathbf{x}(T) = \mathbf{x}_1$. [10, p. 2]

Controllability can be characterised by the the reachable space. The reachable space is the set of all points that can be reached from the origin $\mathbf{x}(0) = 0$ in time T.

Definition 7.2 (Reachable Space)
The reachable space $\mathcal{W}_T$ at a time T is defined as the set of points $\mathbf{x}_1$ for which it is possible to find a control $\mathbf{u}(t)$ such that $\mathbf{x}_{\mathbf{u}}(T, 0) = \mathbf{x}_1$. The reachable space is defined as:

$$\mathcal{W}_T = \left\{ \int_0^T \mathrm{e}^{A(T-\tau)} B\mathbf{u}(\tau)\mathrm{d}\tau \mid \mathbf{u} \in \mathbf{U} \right\}.$$

[6, p. 39]

From (7.6) the following condition is derived:

$$\mathbf{x}_1 - \mathrm{e}^{AT}\mathbf{x}_0 \quad \in \mathcal{W}_T.$$

This means that a system is controllable in time T if it is reachable in time T. Reachability and controllability are equivalent properties, as long as the system is in continuous time. [6, p. 39] The matrix $\mathcal{C} = [B \quad AB \quad \cdots \quad A^{n-1}B]$ is called the controllability matrix and is used to determine if a system is controllable or not.

Definition 7.3 (Controllability Matrix)
The controllability matrix

$$\mathcal{C} = [B \quad AB \quad \cdots \quad A^{n-1}B].$$

[10, p. 3]

To better understand the term reachability and the use of the controllability matrix,see Example 7.1.

Example 7.1

1. In the first example we will see if a point $\mathbf{x}_1$ is reachable. We have to check if $\mathbf{x}_1$ is in the image of the controllability matrix $\mathcal{C}$ by finding a vector $\mathbf{v}$ that solves the equation $\mathcal{C}\mathbf{v} = \mathbf{x}_1$. We will look at the following A and B matrices:

$$A = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

   The point we want to reach $\mathbf{x}_1$ is

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

   We will first calculate AB:

$$AB = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}.$$

   Then we will look at the controllability matrix $\mathcal{C}$, which now yields

$$\mathcal{C} = \begin{bmatrix} B & AB \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}.$$

   Now we solve the equation $\mathcal{C}\mathbf{v} = \mathbf{x}_1$,

$$\begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

It is possible to solve the system of equations:

$$\begin{aligned} v_1 - 2v_2 &= 1 \quad \Rightarrow v_1 = 1 \\ -v_1 + 2v_2 &= -1 \Rightarrow v_2 = 0. \end{aligned}$$

The point $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is reachable!

2. In the second example we will look at how the span of the vectors can give us an understanding of whether the state space model $\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B\mathbf{u}(t)$ is controllable or not. We will look at the same A and B matrix but instead of looking at a specific point $\mathbf{x}_1$, we will look at the general equation

$$\begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \mathbf{x}_p.$$

Now we can write the linear combination of the system

$$v_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + v_2 \begin{bmatrix} -2 \\ 2 \end{bmatrix}.$$

Now we can look at the linear combination as a visual representation in Figure 7.2.

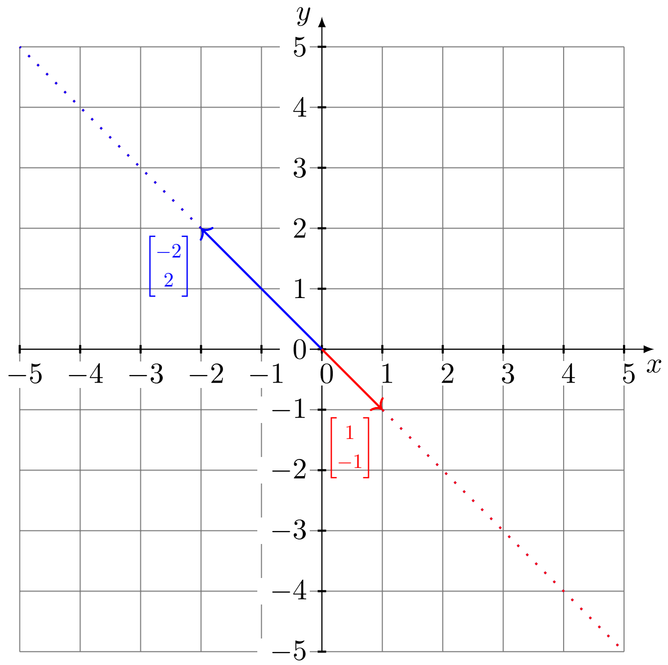


Figure 7.2: Visual representation of the span in Example 7.1.

The scalars v_1 and v_2 can only scale the vectors along the dotted line, and therefore it is impossible to reach points that are not on the line.

For larger matrices, more calculations are needed in order to do an example similar to Example 7.1. Therefore, some general equivalent statements that are used to determine controllability of a system can be seen in Corollary 7.1.

Corollary 7.1
The following statements are equivalent:

1. There exists $T > 0$ such that the system Σ is controllable at T.
2. Rank $\mathcal{C} = \text{rank}\ [B \quad AB \quad \cdots \quad A^{n-1}B] = n$.
3. The system Σ is controllable at T for all $T > 0$.

If one of these equivalent conditions is satisfied the matrix pair (A, B) is controllable.

[6, pp. 40-41]

We will now see an example of using Corollary 7.1.

Example 7.2
We want to see if the system $\dot{\mathbf{x}} = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \mathbf{u}(t)$ is controllable. We see that the A and B matrices are the same from Example 7.1. We already have the controllability matrix $\mathcal{C}$,

$$\mathcal{C} = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}.$$

To see if the system is controllable or not, we have to check if rank $\mathcal{C} = n = 2$. To find the rank, we need to reduce the matrix $\mathcal{C}$ to row echelon form,

$$\begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}.$$

We see that rank $\mathcal{C} \neq 2$ since there is only one pivot entry (c_{11}). Therefore the system is not controllable.

Constructing the matrix K in the controller $\mathbf{u}(t) = -K\mathbf{x}(t)$ leads to the following Theorem 7.1.

Theorem 7.1
The following conditions are equivalent:

1. The system $\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B\mathbf{u}(t)$ is controllable.
2. For an arbitrary polynomial $p(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n$ where $\lambda \in \mathbb{C}$ and $a_1, \cdots, a_n \in \mathbb{R}$, there exists a matrix K such that

   $$p(\lambda) = P_K(\lambda),$$

   where $P_K(\lambda)$ is the characteristic polynomial of $A - BK$.

[23, p. 44]

Chapter 8

Controlling the Inverted Pendulum

Having derived the equations of motion (5.17) and (5.18) (two second order differential equations), we can now write the state of the system, (which we denote as $\mathbf{z}$ as we use x for the cart's position)

$$\mathbf{z} = \begin{bmatrix} z_1 & z_2 & z_3 & z_4 \end{bmatrix}^\mathsf{T} = \begin{bmatrix} x(t) & \dot{x}(t) & \theta(t) & \dot{\theta}(t) \end{bmatrix}^\mathsf{T}$$

as a system of four first-order differential equations. As previously, $x(t), \dot{x}(t), \theta(t), \dot{\theta}(t)$ will only be denoted as $x, \dot{x}, \theta, \dot{\theta}$. This yields

$$\begin{aligned}
\dot{\mathbf{z}} &= \mathbf{f}(\mathbf{z}, \mathbf{u}) \\
&= \begin{bmatrix} \dot{z}_1 & \dot{z}_2 & \dot{z}_3 & \dot{z}_4 \end{bmatrix}^\mathsf{T} \\
&= \begin{bmatrix} \dot{x} & \ddot{x} & \dot{\theta} & \ddot{\theta} \end{bmatrix}^\mathsf{T} \\
&= \begin{bmatrix} \dot{x} \\ \frac{mg\cos(\theta)\sin(\theta) + \frac{\mu_2}{ml}\cos(\theta)\dot{\theta} - ml\sin(\theta)\dot{\theta}^2 + \mu_1\dot{x} - \mathbf{u}}{M+m-m\cos^2(\theta)} \\ \dot{\theta} \\ \frac{mg\cos^2(\theta)\sin(\theta) + \frac{\mu_2}{ml}\cos^2(\theta)\dot{\theta} - ml\cos(\theta)\sin(\theta)\dot{\theta}^2 + \cos(\theta)\mu_1\dot{x} - \mathbf{u}\cos(\theta)}{(M+m)l - ml\cos^2(\theta)} + \frac{g}{l}\sin(\theta) + \frac{\mu_2}{ml^2}\dot{\theta} \end{bmatrix} \\
&= \begin{bmatrix} z_2 \\ \frac{mg\cos(z_3)\sin(z_3) + \frac{\mu_2}{ml}\cos(z_3)z_4 - ml\sin(z_3)z_4^2 + \mu_1 z_2 - \mathbf{u}}{M+m-m\cos^2(z_3)} \\ z_4 \\ \frac{mg\cos^2(z_3)\sin(z_3) + \frac{\mu_2}{ml}\cos^2(z_3)z_4 - ml\cos(z_3)\sin(z_3)z_4^2 + \cos(z_3)\mu_1 z_2 - \mathbf{u}\cos(z_3)}{(M+m)l - ml\cos^2(z_3)} + \frac{g}{l}\sin(z_3) + \frac{\mu_2}{ml^2}z_4 \end{bmatrix}
\end{aligned} \tag{8.1}$$

by insertion of (5.17) and (5.18). Our next goal is to control the inverted pendulum such that it stays in equilibrium. However, we only know how to control a linear system and not a non-linear system. Therefore, we recall Theorem 6.2, wherein it states that the equilibrium point for the non-linear system is asymptotically stable if $\mathrm{Re}(\lambda_i) < 0$ for all eigenvalues of the Jacobian matrix of the linearised system. Per (7.5), this means that the specific K that controls the linear system will also control, and thereby lets us stabilise, the non-linear system. This is our motivation for linearising the system.

As Lyapunov's Indirect Theorem only accounts for initial states near an equilibrium point, we first need to examine the stability of the equilibrium point. If the equilibrium point is stable, we have no interest in controlling it. To examine if our non-linear system is stable, we look at the Jacobian matrix. We first need to find an equilibrium point for our system. First, we look at

$$\dot{\mathbf{z}} = \begin{bmatrix} z_2 \\ \frac{mg\cos(z_3)\sin(z_3)+\frac{\mu_2}{ml}\cos(z_3)z_4-ml\sin(z_3)z_4^2+\mu_1 z_2-\mathbf{u}}{M+m-m\cos^2(z_3)} \\ z_4 \\ \frac{mg\cos^2(z_3)\sin(z_3)+\frac{\mu_2}{ml}\cos^2(z_3)z_4-ml\cos(z_3)\sin(z_3)z_4^2+\cos(z_3)\mu_1 z_2-\mathbf{u}\cos(z_3)}{(M+m)l-ml\cos^2(z_3)} + \frac{g}{l}\sin(z_3) + \frac{\mu_2}{ml^2}z_4 \end{bmatrix} = \mathbf{0}. \tag{8.2}$$

It is quickly seen that z_2 and z_4 both need to equal zero to satisfy (8.2). Furthermore, we want to examine what the input $\mathbf{u}$ is at the equilibrium point.

We look at $\ddot{z}_2$,

$$\ddot{z}_2 = \frac{mg\cos(z_3)\sin(z_3) + \frac{\mu_2}{ml}\cos(z_3)z_4 - ml\sin(z_3)z_4^2 + \mu_1 z_2 - \mathbf{u}}{M+m-m\cos^2(z_3)} = 0.$$

As $z_2 = 0$ and $z_4 = 0$ we get

$$\ddot{z}_2 = \frac{mg\cos(z_3)\sin(z_3) - \mathbf{u}}{M+m-m\cos^2(z_3)} = 0. \tag{8.3}$$

We now see that (8.3) is equal to zero, when the numerator is equal to zero, so

$$mg\cos(z_3)\sin(z_3) - \mathbf{u} = 0.$$

We see that m and g are both positive constants and therefore these are irrelevant in this circumstance. Therefore we get

$$\cos(z_3)\sin(z_3) - \mathbf{u} = 0. \tag{8.4}$$

We see that (8.4) is equal to zero when $\mathbf{u} = \mathbf{0}$ and when

$$z_3 = \theta = \frac{a\pi}{2},\ a = 1, 3, \ldots,\ \text{ or } z_3 = \theta = n\pi,\ n \in \mathbb{N}_0. \tag{8.5}$$

Then we look at $\ddot{z}_4 = 0$

$$\ddot{z}_4 = \frac{mg\cos^2(z_3)\sin(z_3) + \frac{\mu_2}{ml}\cos^2(z_3)z_4 - ml\cos(z_3)\sin(z_3)z_4^2 + \cos(z_3)\mu_1 z_2 - \mathbf{u}\cos(z_3)}{(M+m)l - ml\cos^2(z_3)} \quad \cdots$$

$$\cdots + \frac{g}{l}\sin(z_3) + \frac{\mu_2}{ml^2}z_4 = 0.$$

As $z_2 = 0$ and $z_4 = 0$, and m, g and l are once again positive constants, we get

$$\ddot{z}_4 = \frac{mg\cos^2(z_3)\sin(z_3) - \mathbf{u}\cos(z_3)}{(M+m)l - ml\cos^2(z_3)} + \frac{g}{l}\sin(z_3) = 0. \tag{8.6}$$

Now we see that (8.6) is equal to zero if $\mathbf{u} = \mathbf{0}$ and $\sin(z_3) = 0$ i.e. when

$$z_3 = \theta = n\pi,\ n \in \mathbb{N}_0. \tag{8.7}$$

From (8.5) and (8.7) we get the equilibrium points for the system

$$\mathbf{z}_u = \begin{bmatrix} z_1 & 0 & n\theta & 0 \end{bmatrix}^\mathsf{T}, \quad \text{where } n = 0, 2, 4, 6, \ldots, \tag{8.8}$$

$$\mathbf{z}_d = \begin{bmatrix} z_1 & 0 & k\theta & 0 \end{bmatrix}^\mathsf{T}, \quad \text{where } k = 1, 3, 5, 7, \ldots, \tag{8.9}$$

where $\mathbf{z}_u$, $\mathbf{z}_d$ are the equilibrium points in the upright and downright positions, respectively. The first coordinates are set to z_1 because the cart's displacement is arbitrary in regards to stabilisation. In these equilibrium points, our input $\mathbf{u}$ is the zero vector, which means we do not influence the system in these equilibrium points. As we are interested in stabilising in the upright position we only look at $\mathbf{z}_u$. We can now find a Jacobian matrix for the equilibrium point $\mathbf{z}_u$. To do this we look at $\frac{\partial \dot{\mathbf{z}}}{\partial z_1}, \frac{\partial \dot{\mathbf{z}}}{\partial z_2}, \frac{\partial \dot{\mathbf{z}}}{\partial z_3}, \frac{\partial \dot{\mathbf{z}}}{\partial z_4}$:

$$\frac{\partial \dot{\mathbf{z}}}{\partial z_1} = \frac{\partial \dot{\mathbf{z}}}{\partial x} = \begin{bmatrix} \frac{\partial}{\partial x}\dot{x}, \\ \frac{\partial}{\partial x}\ddot{x}, \\ \frac{\partial}{\partial x}\dot{\theta}, \\ \frac{\partial}{\partial x}\ddot{\theta} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \frac{\partial \dot{\mathbf{z}}}{\partial z_2} = \frac{\partial \dot{\mathbf{z}}}{\partial \dot{x}} = \begin{bmatrix} \frac{\partial}{\partial \dot{x}}\dot{x} \\ \frac{\partial}{\partial \dot{x}}\ddot{x} \\ \frac{\partial}{\partial \dot{x}}\dot{\theta} \\ \frac{\partial}{\partial \dot{x}}\ddot{\theta} \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{\mu_1}{M} \\ 0 \\ \frac{\mu_1}{Ml} \end{bmatrix},$$

$$\frac{\partial \dot{\mathbf{z}}}{\partial z_3} = \frac{\partial \dot{\mathbf{z}}}{\partial \theta} = \begin{bmatrix} \frac{\partial \dot{x}}{\partial \theta} \\ \frac{\partial}{\partial \theta}\ddot{x} \\ \frac{\partial}{\partial \theta}\dot{\theta} \\ \frac{\partial}{\partial \theta}\ddot{\theta} \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{mg}{M} \\ 0 \\ \frac{g(M+m)}{Ml} \end{bmatrix}, \qquad \frac{\partial \dot{\mathbf{z}}}{\partial z_4} = \frac{\partial \dot{\mathbf{z}}}{\partial \dot{\theta}} = \begin{bmatrix} \frac{\partial}{\partial \dot{\theta}}\dot{x} \\ \frac{\partial}{\partial \dot{\theta}}\ddot{x} \\ \frac{\partial}{\partial \dot{\theta}}\dot{\theta} \\ \frac{\partial}{\partial \dot{\theta}}\ddot{\theta} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{\mu_2}{mlM} \\ 1 \\ \frac{\mu_2+\mu_2 M}{ml^2 M} \end{bmatrix}.$$

We can now write the Jacobian matrix $\frac{\partial}{\partial \mathbf{z}}\mathbf{f}(\mathbf{z}_u, \mathbf{u})$ with $\frac{\partial \dot{\mathbf{z}}}{\partial z_1}, \frac{\partial \dot{\mathbf{z}}}{\partial z_2}, \frac{\partial \dot{\mathbf{z}}}{\partial z_3}, \frac{\partial \dot{\mathbf{z}}}{\partial z_4}$ as its columns and the values from Table 8.1 inserted:

$$\frac{\partial}{\partial \mathbf{z}}\mathbf{f}(\mathbf{z}_u, \mathbf{u}) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{\mu_1}{M} & -\frac{mg}{M} & \frac{\mu_2}{mlM} \\ 0 & 0 & 0 & 1 \\ 0 & \frac{\mu_1}{Ml} & \frac{g(M+m)}{Ml} & \frac{\mu_2+\mu_2 M}{ml^2 M} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0.080 & -0.274 & 0.002 \\ 0 & 0 & 0 & 1 \\ 0 & 0.283 & 35.920 & 0.042 \end{bmatrix}. \tag{8.10}$$

If we now look at the determinant of the Jacobian matrix $\frac{\partial}{\partial \mathbf{z}}\mathbf{f}(\mathbf{z}_u, \mathbf{u})$ we get that

$$\det\left(\frac{\partial}{\partial \mathbf{z}}\mathbf{f}(\mathbf{z}_u, \mathbf{u})\right) = 0,$$

which confirms that our system has more than one equilibrium point as we derived in (8.8) and (8.9).

Name	Symbol	Value	Unit
Mass of cart	M	6.28	kg
Rod length	l	0.281	m
Mass of pendulum	m	0.175	kg
Gravitational acceleration	g	9.82	$\frac{\mathrm{m}}{\mathrm{s}^2}$
Cart viscous friction	μ_1	0.5	$\frac{\mathrm{Ns}}{\mathrm{m}}$
Pendulum viscous friction	μ_2	$0.5 \cdot 10^{-3}$	Nms

Table 8.1: Values for variables for the system. [21]

In the data sheet, the cart's viscous friction constant is set to 0. As we do not account for coulomb friction, but still wish to account for friction, we instead set the viscous friction constant to 0.5 $\frac{\mathrm{Ns}}{\mathrm{m}}$ From (8.10) we get the eigenvalues:

$$\lambda_1 = 0, \qquad \lambda_2 = 0.0822, \qquad \lambda_3 = 6.0133, \qquad \lambda_4 = -5.974.$$

We now see that $\mathrm{Re}(\lambda_2), \mathrm{Re}(\lambda_3) > 0$, and by Theorem 6.2 the equilibrium point is unstable. Therefore, we regain interest in controlling, and thereby stabilise, the pendulum. We now write the system (8.1) on the form (7.2):

$$\dot{\mathbf{z}}(t) = A\mathbf{z}(t) + B\mathbf{u}(t),$$

where $A = \frac{\partial}{\partial \mathbf{z}}\mathbf{f}(\mathbf{z}_u, \mathbf{u})$ is our Jacobian matrix found at the equilibrium point, $\mathbf{z}(t)$ is the state of the system, $\mathbf{u}(t)$ is the input and $B = \frac{\partial}{\partial \mathbf{u}}\mathbf{f}(\mathbf{z}_u, \mathbf{u})$ is another Jacobian matrix:

$$\frac{\partial}{\partial \mathbf{u}}\mathbf{f}(\mathbf{z}_u, \mathbf{u}) = \frac{\partial \dot{\mathbf{z}}}{\partial \mathbf{u}} = \begin{bmatrix} \frac{\partial}{\partial u}\dot{x} \\ \frac{\partial}{\partial u}\ddot{x} \\ \frac{\partial}{\partial u}\dot{\theta} \\ \frac{\partial}{\partial u}\ddot{\theta} \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{M} \\ 0 \\ -\frac{1}{Ml} \end{bmatrix} = \begin{bmatrix} 0 \\ -0.160 \\ 0 \\ -0.567 \end{bmatrix}. \tag{8.11}$$

We now have the following expression for the system:

$$\dot{\mathbf{z}}(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0.080 & -0.274 & 0.002 \\ 0 & 0 & 0 & 1 \\ 0 & 0.283 & 35.920 & 0.042 \end{bmatrix}\mathbf{z}(t) + \begin{bmatrix} 0 \\ -0.160 \\ 0 \\ -0.567 \end{bmatrix}\mathbf{u}(t)$$

As we wish to influence the system, we need to examine the controllability of the system. Per Definition 7.3 we determine the controllability matrix:

$$\mathcal{C} = \begin{bmatrix} B & AB & A^2B & A^3B \end{bmatrix} = \begin{bmatrix} 0 & -0.160 & -0.014 & 0.154 \\ -0.160 & -0.014 & 0.154 & -0.010 \\ 0 & -0.567 & -0.070 & -20.374 \\ -0.567 & -0.069 & -20.374 & -3.304 \end{bmatrix}. \tag{8.12}$$

Per Corollary 7.1, we look at the rank of (8.12) by reducing to row echelon form

$$\mathcal{C} = \begin{bmatrix} 0 & -0.160 & -0.014 & 0.154 \\ -0.160 & -0.014 & 0.154 & -0.010 \\ 0 & -0.567 & -0.070 & -20.374 \\ -0.567 & -0.069 & -20.374 & -3.304 \end{bmatrix} \sim \begin{bmatrix} -0.567 & -0.069 & -20.374 & -3.304 \\ 0 & -0.567 & -0.070 & -20.374 \\ 0 & 0 & 5.903 & 0.726 \\ 0 & 0 & 0 & 5.903 \end{bmatrix}. \tag{8.13}$$

From (8.13) we see that the rank of our controllability matrix $\mathcal{C}$ is $\mathrm{rank}(\mathcal{C}) = 4$. As this is full rank, we see that our system is controllable.

We now know that our system is controllable and we can write the system as the following per (7.5),

$$\dot{\mathbf{z}}(t) = (A - BK)\mathbf{z}(t). \tag{8.14}$$

In Example 8.1, we exemplify finding K for a given controllable system, (8.14).

Example 8.1 (Finding K using real eigenvalues)
We wish to find K such that our system (8.14) is stable and has the following eigenvalues:

$$\lambda_1 = -1,\ \lambda_2 = -2,\ \lambda_3 = -3,\ \lambda_4 = -4. \tag{8.15}$$

The eigenvalues (8.15) are chosen based on Theorem 6.2 as this states that a system is stable if the real parts of the eigenvalues are negative. We utilise the `Python` function `scipy.signal.place_poles()` from the `SciPy` module which yields the following code.

```
a,b,c,d=-1,-2,-3,-4
poles = np.array([a,b,c,d])
A = np.array([[0,1,0,0],
    [0,0.07961783439,-0.2736464968,0.001619071365],
    [0,0,0,1],
    [0,0.2833374889,35.92045018,0.04194604818]])
B = np.array([[0],
    [-0.1592356688],
    [0],
    [-0.5666749779]])
P = scipy.signal.place_poles(A, B, poles, method='YT', rtol=0.001, maxiter=100)
K = P.gain_matrix
print(K)
Q = A-np.matmul(B,K)

print(Q)
print(scipy.linalg.eig(Q))
```

The function `scipy.signal.place_poles()` yields a range of information, including the gain matrix `K`. The arguments `A,B` are the Jacobian matrices of the linearised system as seen in (8.10) and (8.11). The `poles` are the desired eigenvalues `a,b,c,d` of the system. The algorithm used is called `'YT'` ("Yang-Tits"). We are not going to explore this further, as this is out of bounds in regards to this project. After each iteration in the process of computing `K`, the determinant of the eigenvectors between `A-np.matmul(B,K)` gets compared to the value from the previous iteration. When the relative error between two iterations is lower than `rtol`, the iteration stops, and thus `K` has been determined. `maxiter` is the maximum number of iterations used to achieve `K`. `print(scipy.linalg.eig(Q))` prints the calculated eigenvalues of `A-np.matmul(B,K)`, such that we can compare them to the desired eigenvalues.

From this code we obtain

$$K = \begin{bmatrix} 4.09 & 8.01 & -126.21 & -20.11 \end{bmatrix}. \tag{8.16}$$

Inserting (8.16) into (8.14) yields

$$A - BK = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.651 & 1.355 & -20.371 & -3.201 \\ 0 & 0 & 0 & 1 \\ 2.315 & 4.821 & -35.601 & -11.355 \end{bmatrix}. \tag{8.17}$$

The computed eigenvalues of (8.17) are

$$\lambda_4 = -4. + 0j,\ \lambda_3 = -3. + 0j,\ \lambda_2 = -2. + 0j,\ \lambda_1 = -1. + 0j.$$

We see that the eigenvalues for this system, with the estimated controller K, gives us the desired eigenvalues, (8.15). However, in the code used for the simulation of the equations of motion, K is a variable matrix as opposed to fixed as seen in this specific example.

If you seek further information regarding `scipy.signal.place_poles()`, see the following reference: [3].

Determining K for the system (8.14) enables us to control and thereby stabilise the linearised system. Controlling the linearised system is equivalent to controlling the non-linear system, per Theorem 6.2. Therefore, we insert the relation $\mathbf{u}(t) = -K\mathbf{z}(t)$ from (7.4) in (8.1) and get

$$\dot{\mathbf{z}}(t) = \begin{bmatrix} z_2 \\ \frac{mg\cos(z_3)\sin(z_3)+\frac{\mu_2}{ml}\cos(z_3)z_4-ml\sin(z_3)z_4^2+\mu_1 z_2+K\mathbf{z}}{M+m-m\cos^2(z_3)} \\ z_4 \\ \frac{mg\cos^2(z_3)\sin(z_3)+\frac{\mu_2}{ml}\cos^2(z_3)z_4-ml\cos(z_3)\sin(z_3)z_4^2+\cos(z_3)\mu_1 z_2+K\mathbf{z}\cos(z_3)}{(M+m)l-ml\cos^2(z_3)}+\frac{g}{l}\sin(z_3)+\frac{\mu_2}{ml^2}z_4 \end{bmatrix}. \tag{8.18}$$

The controllable system (8.18), with K inserted, is the system that we simulate.

Chapter 9

Simulation of the Equations of Motion

In this chapter, we introduce the numerical method used to determine the solutions to the equations of motion. Furthermore, we will assess the restraints before executing the simulation. We examine the effect of different sets of eigenvalues on the system. In a trial and error process of testing different sets of eigenvalues, we find two sets of eigenvalues given the system specifications in Table 8.1.

9.1 Prerequisites for Simulation

The foundation of our simulation consists of a range of definitions and concepts from applied scientific computing and classical mechanics: the Runge-Kutta 4 Method, work, moment of inertia, and torque.

9.1.1 The Runge-Kutta 4 Method

We solve our non-linear system (8.18) by numerical approximation, by using the `scipy.integrate.solve_ivp` function from the `SciPy` module in `Python`. This function uses a fourth order Runge-Kutta Method to approximate the IVP. The fourth order Runge-Kutta method is:

$$z_{n+1} \approx z_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

where

$$\begin{aligned}
k_1 &= f\left(t_n, z_n\right), \\
k_2 &= f\left(t_n + \frac{h}{2}, z_n + \frac{h}{2}k_1\right), \\
k_3 &= f\left(t_n + \frac{h}{2}, z_n + \frac{h}{2}k_2\right), \\
k_4 &= f\left(t_n + h, z_n + hk_3\right),
\end{aligned}$$

where z_n and z_{n+1} are the function values for our system and h is the step size (for time dependent functions, the time difference) between z_{n+1} and z_n. [16, p. 316] Application of the Runge-Kutta 4 Method is explicitly exemplified in Example 9.1.

Example 9.1 (Application of Runge-Kutta 4)
We seek to approximate the solution for

$$\frac{\mathrm{d}}{\mathrm{d}t}z = 2tz, \quad \text{with initial condition} \quad z(1) = 1 \tag{9.1}$$

and step size $h = 0.1$. The IVP (9.1) has the exact solution

$$\mathrm{e}^{t^2-1}.$$

For the first step, the k's and z_{n+1} are calculated:

$$\begin{aligned}
k_1 &= 2 \cdot 1 \cdot 1 = 2, \\
k_2 &= 2\left(1+\frac{0.1}{2}\right)\left(1+2\frac{0.1}{2}\right) = 2.31, \\
k_3 &= 2\left(1+\frac{0.1}{2}\right)\left(1+2.31\frac{0.1}{2}\right) = 2.34255, \\
k_4 &= 2\left(1+0.1\right)\left(1+0.1 \cdot 2.34255\right) = 2.715361, \\
z_{n+1} &\approx 1 + \frac{0.1}{6}\left(k_1+2k_2+2k_3+k_4\right) = 1.23367435.
\end{aligned}$$

We insert our results in Table 9.1.

z	t	h	k_1	k_2	k_3	k_4	Actual value of e^{t^2-1}
1	1	0.1	2	2.31	2.34255	2.715361	1
1.23367435	1.1						1.23367806

Table 9.1: Approximated values for RK4 example.

As made obvious by Table 9.1, the method is very precise. For the first step, the error is $3.710 \cdot 10^{-6}$.

9.1.2 Work

In our simulation, we want to assess the effect of scaling our chosen sets of eigenvalues (see Table 9.3) in regards to the time it takes the system to reach equilibrium. Therefore, it is relevant to determine the work exerted by the motor on the car, applying each of the scaled sets of eigenvalues.

Definition 9.1 (Work)
Assume that a force $\mathbf{F}(t)$ acts on a particle in time $[t_1, t_2]$. The work $W = W(t_1, t_2)$ done on the particle by the force is defined as

$$W = \int_{t_1}^{t_2} \langle \mathbf{F}(t) | \mathbf{v}(t) \rangle \, \mathrm{d}t,$$

where $\mathbf{v}(t)$ is the velocity of the particle.

[19, p. 22]

Remark for Definition 9.1: To apply work in our simulation, we regard our system as a particle.

When work is applied in `Python`, we utilise the `quad` function from the `scipy.integrate` package. $\mathbf{F}(t)$ in our simulation is our input $\mathbf{u}(t) = -K\mathbf{z}(t)$, and $\mathbf{v}(t)$ is the velocity of the cart $\dot{x}$ obtained through solving our system of equations of motion numerically. Applying work in `Python` yields the following code:

```python
def integrand(x):
    return np.matmul(F,v)
print(f'W = {quad(integrand,t_1,t_2)}')
```

In the code, the function `integrand(x)` defines the function you seek to integrate, and `np.matmul(F,v)` is the syntax for multiplying matrices (we regard the vectors $\mathbf{F}$(t) and $\mathbf{v}$(t) as $n \times 1$ matrices). The `quad` function then takes the integrand and the time limits, `t_1`,`t_2`, as arguments and approximates the integral. The code outputs the work as well as the numerical error obtained by the approximation. For now, assume that the error is minor. The specific error sizes will be assessed in Section 9.3.

9.1.3 Moment of Inertia

The moment of inertia is introduced in Definition 9.2 and will be used later in this project to determine the torque.

Definition 9.2 (Moment of Inertia of a Particle)
The quantity

$$I_z = mr^2$$

is called the moment of inertia of a particle with mass m performing a circular motion with radius r in the xy-plane around the z-axis.

[19, p. 225]

Moment of inertia is important in the derivations of maximum exerted force in Section 9.2.

9.1.4 Torque

We are interested in the force needed to perform the rotational motion of the pulley (see Figure 1.1). This force is called torque. We define this motion as a circular motion of a mass. Before we can define torque, we first need to define angular momentum for a particle, which is defined as

$$L = mr^2\dot{\theta} = I_z\dot{\theta}.$$

Torque is then defined as

$$\tau = \frac{\mathrm{d}}{\mathrm{d}t}L = I_z\frac{\mathrm{d}}{\mathrm{d}t}\dot{\theta} = I_z\ddot{\theta} \tag{9.2}$$

per [20]. We use this in Section 9.2.

9.2 Restraints of Simulation

The restraints for the simulation is based on the setup seen on Figure 1.1. The essential restraints for our simulation is the length of the rail, which is 0.89 m, the maximum exerted force by the motor on the cart, and the eigenvalues of $A - BK$.

The cart's position can only be in the interval $x \in [-0.445 \text{ m}, 0.445 \text{ m}]$ as we set $x = 0$ as the midpoint of the rail.

The motor used in the system is a Maxon 370356 DC motor [4], (specifically 578298/618572).

Name	Symbol	Value	Unit
Nominal torque	τ_n	0.420	Nm
Stall torque	τ_s	7.370	Nm

Table 9.2: Data sheet for the Maxon 370356 DC motor. [4]

In Table 9.2, the two values nominal torque and stall torque describe the maximum continuous torque for the motor and the torque when the rotational speed for the motor is zero, respectively. As we wish to find the maximum force exerted by the motor we look at the stall torque and divide it by the radius of the pulley $r_{Pulley} = 0.028$ m [21]. We then get,

$$F_{max} = \frac{\tau_s}{r_{Pulley}} = \frac{7.370 \text{ Nm}}{0.028\text{m}} = 263.2142857 \text{ N}. \tag{9.3}$$

The maximum force $F_{max} \approx 263$ N on the cart induced by the motor is based on the stall torque of the motor, at which it is only recommended to run for about $2 - 3$ seconds [22].

We divide the nominal torque from Table 9.2 by the same radius to determine the continuous force exerted by the motor and get,

$$F_{continuous} = \frac{\tau_n}{r_{Pulley}} = \frac{0.420 \text{ Nm}}{0.028 \text{ m}} = 15.00 \text{ N}. \tag{9.4}$$

We can, according to (9.4), expect a continous force of 15 N on our system.

Furthermore, we must consider the eigenvalues used. The set of eigenvalues used in Example 8.1 ($\lambda = \{-1, -2, -3, -4\}$) yields Figure 9.1 when using the initial condition

$$\mathbf{z}_0 = \begin{bmatrix} x_0 & \dot{x}_0 & \theta_0 & \dot{\theta}_0 \end{bmatrix}^\intercal = \begin{bmatrix} 0 & 0 & 0.2 & 0 \end{bmatrix}^\intercal . \tag{9.5}$$

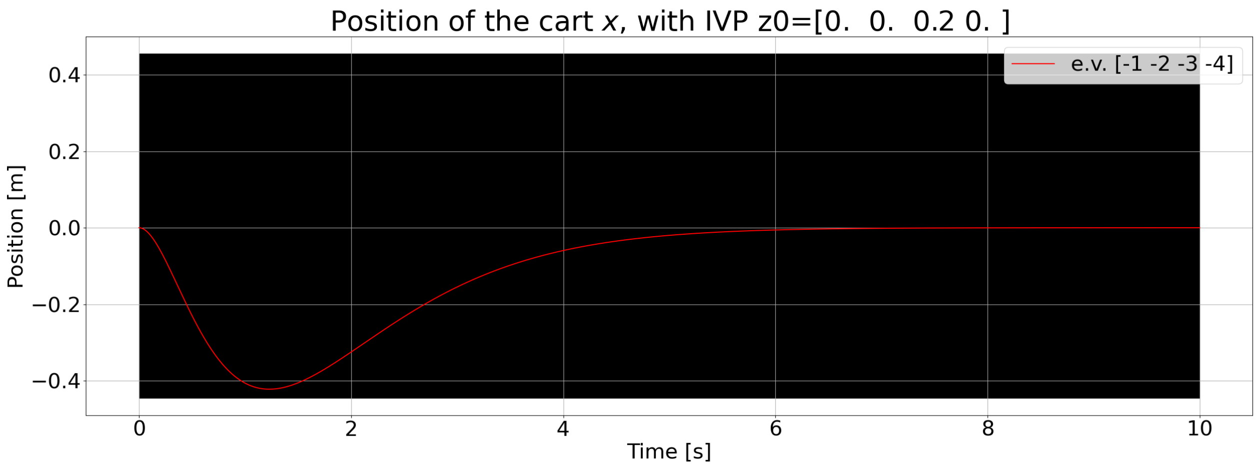


Figure 9.1: Position of the cart. [Made using the `matplotlib.pyplot` module in `Python`.]

We see on Figure 9.1 that the position of the cart goes from 0.00 m to ~ -0.42 m and then back to 0.00 m without going to the other side of the midpoint. The black area marks the area in which the graph for x is allowed. Breaking these limits will result in the cart falling off the rail.

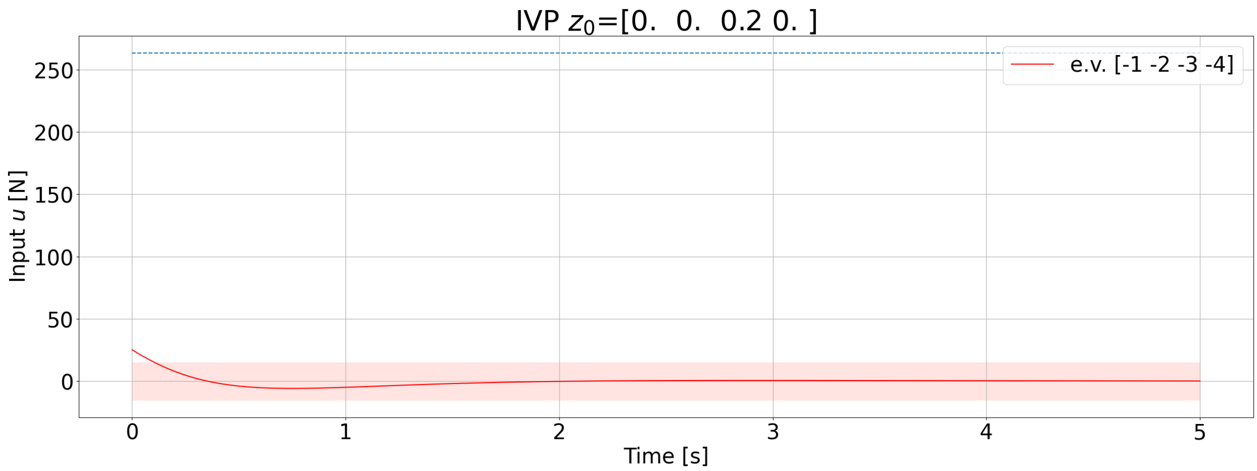


Figure 9.2: Input force exerted on the cart. [Made using the `matplotlib.pyplot` module in `Python`.]

From Figure 9.2 we see the input force exerted by the motor on the cart. Clearly, the maximum exerted force is never exceeded. Thus, the eigenvalues from Example 8.1 are promising candidates for scaling and simulation.

Through trial and error, we decide on using complex eigenvalues.

9.3 Simulation

Through trial and error, we have chosen two different sets of eigenvalues as seen in Table 9.3.

	Set 1	Set 2
Template set	$\{-2 \pm 3j, -3 \pm 3j\}$	$\{-1 \pm 1j, -2 \pm 2j\}$
Scaled by 1.5	$\{-3 \pm 4.5j, -4.5 \pm 4.5j\}$	$\{-1.5 \pm 1.5j, -3 \pm 3j\}$
Scaled by 2	$\{-4 \pm 6j, -6 \pm 6j\}$	$\{-2 \pm 2j, -4 \pm 4j\}$
Scaled by 2.5	$\{-5 \pm 7.5j, -7.5 \pm 7.5j\}$	$\{-2.5 \pm 2.5j, -5 \pm 5j\}$

Table 9.3: Sets of eigenvalues used in the simulation.

Furthermore, we will only assess the stabilisation of the inverted pendulum using the initial condition (9.5) as exemplified in Section 9.2. The initial angle θ_0 in radians is equivalent to 11.46 degrees.

9.3.1 Simulation Using Set 1 of Eigenvalues

Simulating using Set 1 of the eigenvalues from Table 9.3 provides Figures 9.3 - 9.6. As in Section 9.2, the black area on Figure 9.3 defines the area in which the cart is allowed to move.

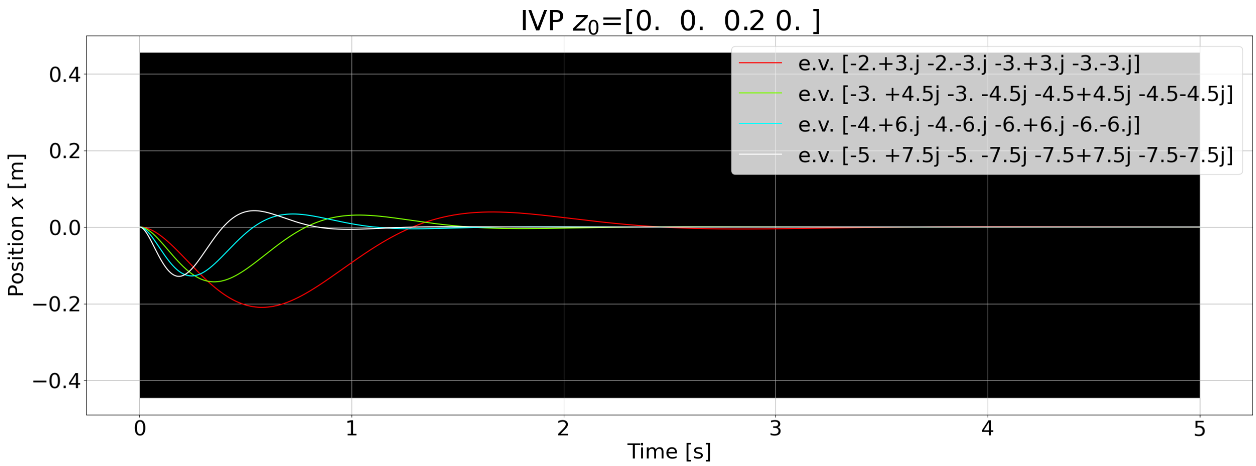


Figure 9.3: Position x of the cart. [Made using the `matplotlib.pyplot` module in `Python`.]

From Figure 9.3, we see that all four scaled sets of eigenvalues are able to stabilise the pendulum without breaking the limits set by the rail with the given initial condition (9.5).
Figure 9.4 shows how the angle θ of the pendulum changes over time. With Set 1 of eigenvalues (see Table 9.3), the initial condition θ_0 cannot exceed $\sim \pm 0.4$ radians. As this restraint is adhered to, all sets of eigenvalues are able to stabilise the system.

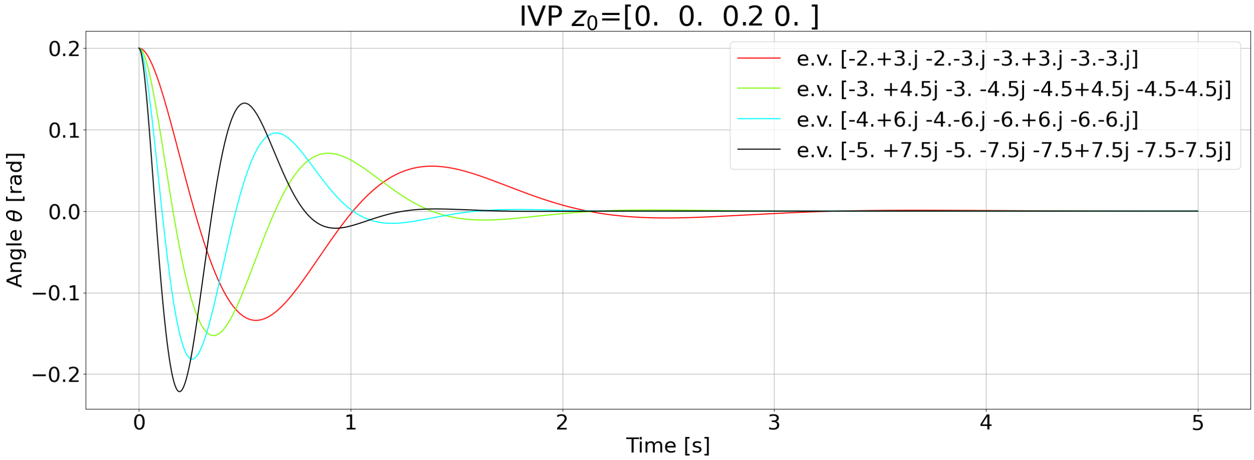
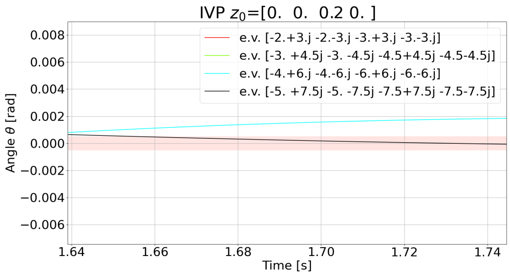


Figure 9.4: Angle θ of the pendulum. [Made using the `matplotlib.pyplot` module in `Python`.]

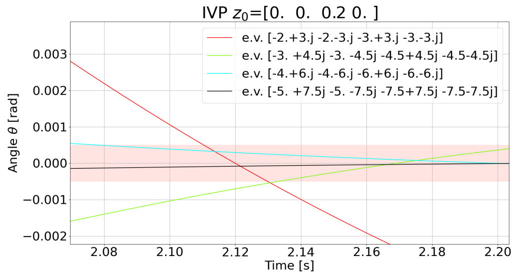
We set a region of convergence to $[-0.0005 \text{ rad}, 0.0005 \text{ rad}]$ (0.0286 degrees) around the time axis. This region is colored using `Python`'s colour scheme `misty rose`. We accept this as stability. We zoom in on the four solutions to assess their times of convergence. This is seen on Figure 9.5, and the results are seen in Table 9.4. For Figures 9.5a - 9.5d in full size, see Appendix C. We see that the template set of eigenvalues has the slowest conversion rate towards equilibrium. For each scaling, the conversion happens faster. As seen in Table 9.4, the template set of eigenvalues converges ~ 2.5 slower than the eigenvalues scaled by 2.5.

Set of Eigenvalues	**Converges After**	**See Figure**
$\lambda = \{-5 \pm 7.5j, -7.5 \pm 7.5j\}$	$\sim$ 1.66 seconds	9.5a
$\lambda = \{-4 \pm 6j, -6 \pm 6j\}$	$\sim$ 2.08 seconds	9.5b
$\lambda = \{-3 \pm 4.5j, -4.5 \pm 4.5j\}$	$\sim$ 2.75 seconds	9.5c
$\lambda = \{-2 \pm 3j, -3 \pm 3j\}$	$\sim$ 4.125 seconds	9.5d

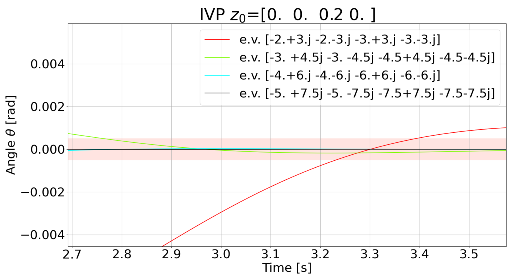
Table 9.4: Sets of eigenvalues and their time of convergence.



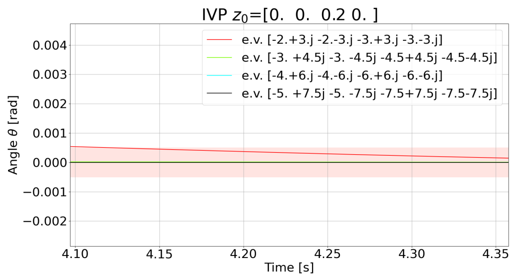
(a) Convergence with $\lambda = \{-5 \pm 7.5j, -7.5 \pm 7.5j\}$.



(b) Convergence with $\lambda = \{-4 \pm 6j, -6 \pm 6j\}$.



(c) Convergence with $\lambda = \{-3 \pm 4.5j, -4.5 \pm 4.5j\}$.



(d) Convergence with $\lambda = \{-2 \pm 3j, -3 \pm 3j\}$.

Figure 9.5: Convergence of the eigenvalues in Set 1. [Made using the `matplotlib.pyplot` module in `Python`.]

Figure 9.6 shows the exerted force on the system (the input $\mathbf{u}$) for each scaled set of eigenvalues. The light pink area marks the area for continuous force (± 15.00 N, per (9.4)), and the dotted line marks the maximum exerted force (263.21 N, per (9.3)). It is clearly seen that the exerted force on the system never exceeds the maximum limit, and it only exceeds the continuous force for short periods of time over the course of 5 seconds.

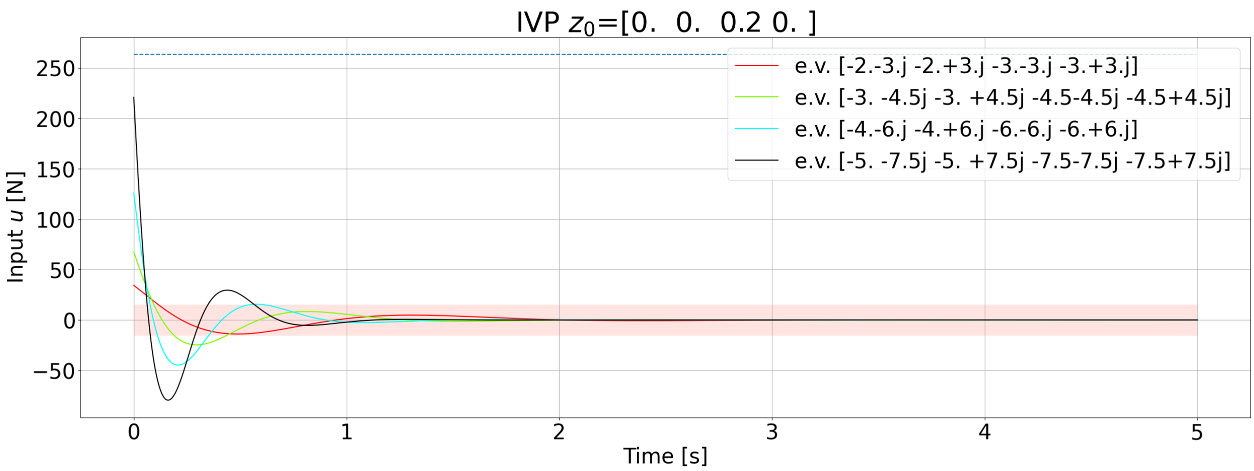


Figure 9.6: Input force $\mathbf{u}$ on the system. [Made using the `matplotlib.pyplot` module in `Python`.]

In order to compute the work the motor exerts on the cart, we look at the absolute values of our input ($\mathbf{F}$) and cart velocity ($\mathbf{v}$), as the total energy is direction invariant. Hence, we compute the work for the four scaled sets of eigenvalues using a code analogue to the code in Section 9.2, which yields

$$\begin{aligned}
W &= 31,489.62 \text{ J}, \quad \text{with error } 3.496 \cdot 10^{-10}, \quad \text{for eigenvalues } \lambda = \{-2 \pm 3j, -3 \pm 3j\},\\
W &= 39,088.39 \text{ J}, \quad \text{with error } 4.340 \cdot 10^{-10}, \quad \text{for eigenvalues } \lambda = \{-3 \pm 4.5j, -4.5 \pm 4.5j\},\\
W &= 64,873.40 \text{ J}, \quad \text{with error } 7.202 \cdot 10^{-10}, \quad \text{for eigenvalues } \lambda = \{-4 \pm 6j, -6 \pm 6j\},\\
W &= 120,039.63 \text{ J}, \quad \text{with error } 1.333 \cdot 10^{-09}, \quad \text{for eigenvalues } \lambda = \{-5 \pm 7.5j, -7.5 \pm 7.5j\}.
\end{aligned} \tag{9.6}$$

Per our calculations (9.6), the scaled set of eigenvalues that converges the fastest ($\lambda = \{-5 \pm 7.5j, -7.5 \pm 7.5j\}$) is also the set for which the cart exerts the largest amount of work to reach equilibrium. This is a relevant observation in regards to implementation in real life: fast stabilisation requires large amounts of work. As previously assumed, the error *is* minor.

9.3.2 Simulation Using Set 2 of Eigenvalues

Simulating using Set 2 of the eigenvalues from Table 9.3 provides Figures 9.7 - 9.10. As in Section 9.2, the black area on Figure 9.7 defines the area in which the cart is allowed to move.

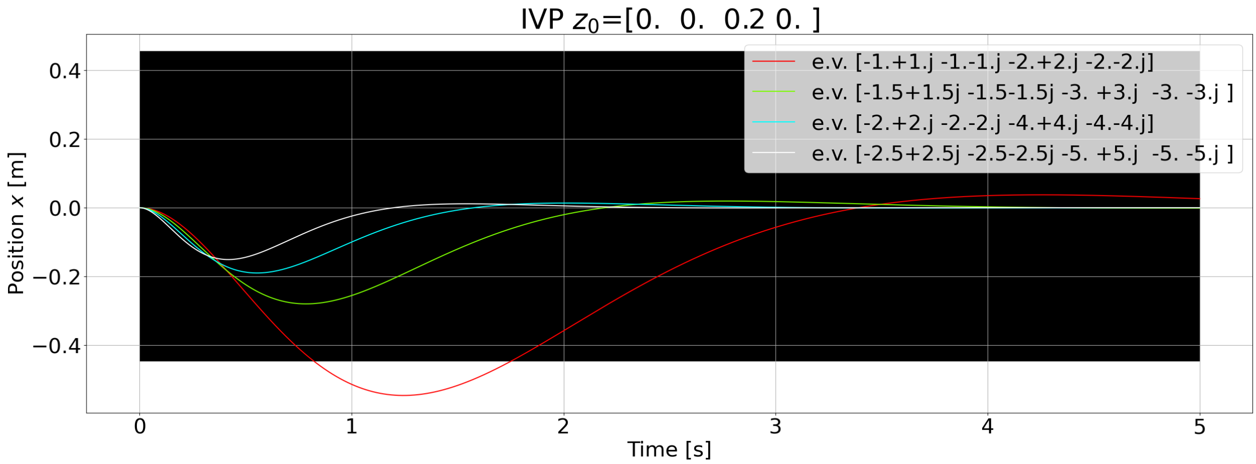


Figure 9.7: Position of the cart x. [Made using the `matplotlib.pyplot` module in `Python`.]

As seen on Figure 9.7, the template set of eigenvalues is not able to stabilise the pendulum with the given initial condition (9.5), as the cart will cross the limits set by the length of the rail. For the template set of eigenvalues to stabilise the system, the initial condition cannot exceed $\theta_0 = \pm 0.16$ radians. However, the scaled sets of eigenvalues are able to stabilise the pendulum within the limits and initial condition (9.5). Furthermore, we see that the largest scaled set, as with Set 1, converges towards stability the fastest.

Figure 9.8 shows how the angle θ of the pendulum changes over time. With Set 2 of eigenvalues (see Table 9.3), the initial condition θ_0 cannot exceed $\sim \pm 0.3$. As this restraint is adhered to, the scaled sets of eigenvalues are able to stabilise the system.

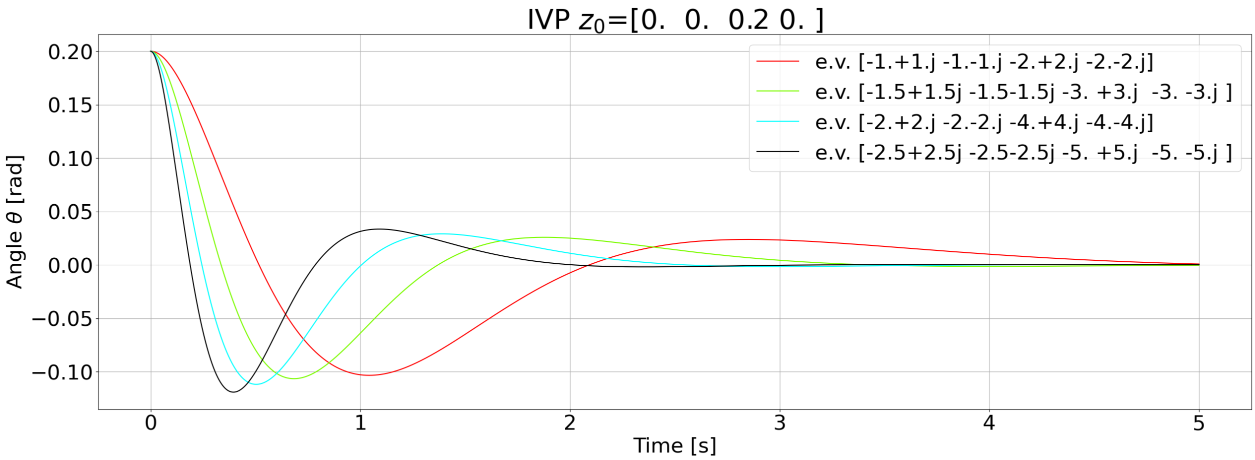
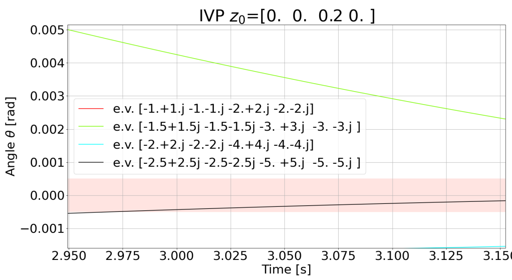


Figure 9.8: Angle of the pendulum θ. [Made using the `matplotlib.pyplot` module in `Python`.]

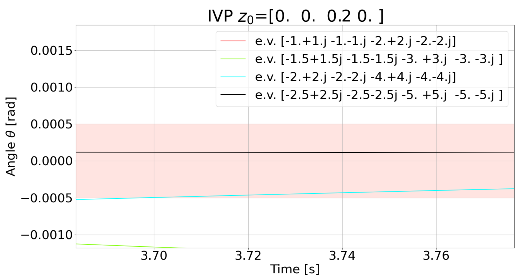
We zoom in on the four solutions to assess their times of convergence. This is seen on Figure 9.9, and the results are seen in Table 9.5. For Figures 9.9a - 9.9c in full size, see Appendix D. We ignore the template set, as previously mentioned. For each scaling, the conversion happens faster.

Set of Eigenvalues	Converges After	See Figure
$\lambda = \{-2.5 \pm 2.5j, -5 \pm 5j\}$	~ 2.975 seconds	9.9a
$\lambda = \{-2 \pm 2j, -4 \pm 4j\}$	~ 3.70 seconds	9.9b
$\lambda = \{-1.5 \pm 1.5j, -3 \pm 3j\}$	~ 4.90 seconds	9.9c

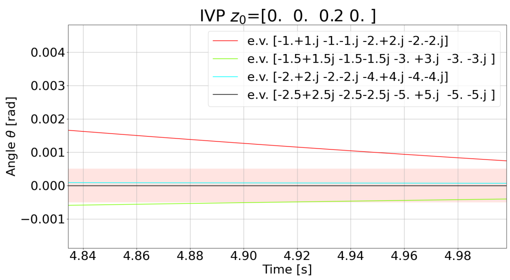
Table 9.5: Set 2 of eigenvalues and their time of convergence.



(a) Convergence with $\lambda = \{-2.5 \pm 2.5j, -5 \pm 5j\}$.



(b) Convergence with $\lambda = \{-2 \pm 2j, -4 \pm 4j\}$.



(c) Convergence with $\lambda = \{-1.5 \pm 1.5j, -3 \pm 3j\}$.

Figure 9.9: Convergence of the eigenvalues in Set 2. [Made using the `matplotlib.pyplot` module in `Python`.]

Figure 9.10 shows the exerted force on the system (the input $\mathbf{u}$) for all four scaled sets of eigenvalues. The light pink area marks the area for continuous force (± 15.00 N, per (9.4)). It is clearly seen that the exerted force on the system never exceeds the maximum limit (263.21 N, per (9.3)), and it only exceeds the continuous force for short periods of time over the course of 5 seconds for $\lambda = \{-2.5 \pm 2.5, -5 \pm 5j\}$.

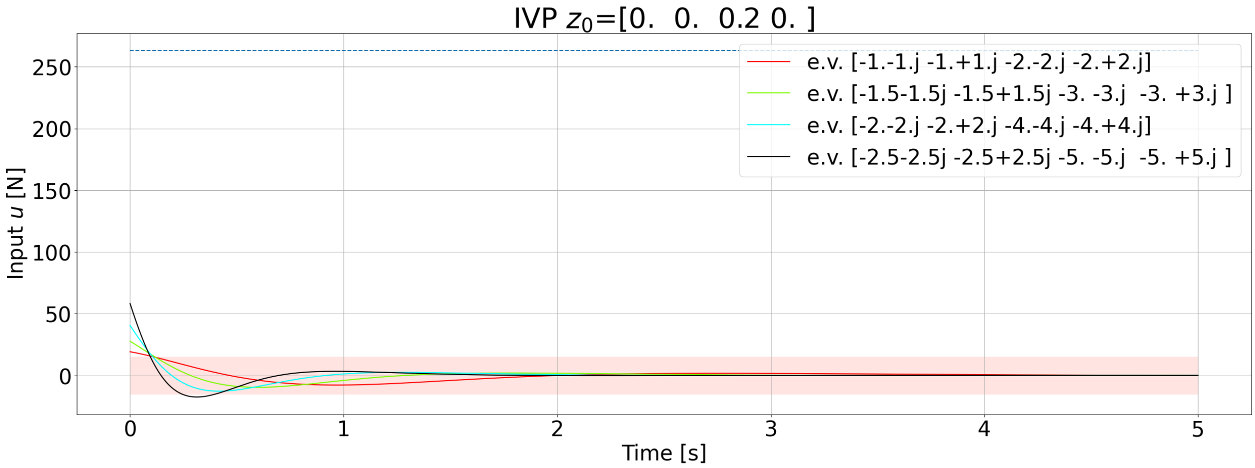


Figure 9.10: Input force $\mathbf{u}$ on the system. [Made using the `matplotlib.pyplot` module in `Python`.]

In order to compute the work the motor exerts on the cart, we look at the absolute values of our input ($\mathbf{F}$) and cart velocity ($\mathbf{v}$), analogously to (9.6), which yields

$$\begin{aligned} W &= 27{,}057.908 \text{ J}, \text{ with error } 3.004 \cdot 10^{-10}, \quad \text{for eigenvalues } \lambda = \{-1.5 \pm 1.5j, -3 \pm 3j\}\ , \\ W &= 24{,}697.058 \text{ J}, \text{ with error } 2.742 \cdot 10^{-10}, \quad \text{for eigenvalues } \lambda = \{-2 \pm 2j, -4 \pm 4j\}, \\ W &= 27{,}120.930 \text{ J}, \text{ with error } 3.011 \cdot 10^{-10}, \quad \text{for eigenvalues } \lambda = \{-2.5 \pm 2.5j, 5 \pm 5j\}. \end{aligned} \tag{9.7}$$

Note that the work calculated using our template eigenvalues is ignored as the template eigenvalues cannot stabilise the cart. Per our calculations (9.7), the scaled set of eigenvalues that converges the fastest ($\lambda = \{-2.5 \pm 2.5j, -5 \pm 5j\}$) is also the set for which the cart exerts the largest amount of work to reach equilibrium (albeit by a small margin), again analogously to (9.6). Once again, the errors are minor.

9.4 Choosing Eigenvalues for Real-life Applications

In the simulation, we have examined the effect of utilising two different sets of eigenvalues on the system (see Table 9.3), by scaling both sets by 1.5, 2, and 2.5, respectively. In both Set 1 and Set 2, the set scaled by 2.5 are the eigenvalues that stabilise the system the fastest but also require the most energy in order to obtain stability. With the inverted pendulum, we have an unlimited power supply. Therefore, the amount of energy needed to obtain stability is not crucial. Hence, we are only interested in stabilising the system as fast as possible. However, in a hypothetical case where the system is battery-driven, i.e. a drone or a Segway, it is crucial to consider the amount of work needed to obtain stability when choosing the eigenvalues. Hence, you may have to compromise on the convergence rate in order to prolong battery life when control and stability is applied in real-life situations. But as seen in Table 9.4 and (9.6), the difference in convergence time is 0.42 seconds when comparing the eigenvalues scaled by 2.5 and 2, respectively, albeit you can save more than 55 kJ of energy. So by choosing your eigenvalues cleverly, you can achieve fast stabilisation while saving a lot of energy.

Chapter 10

Conclusion

To analyse the inverted pendulum, we derived the equations of motion using the Euler-Lagrange Equations. We linearised the non-linear system, since it is not possible to use linear control theory on a non-linear system. The system was linearised using the Jacobian Matrix and by Lyapunov's Indirect Theorem it is possible to determine stability of an equilibrium point for the non-linear system. The system was unstable in the equilibrium point but we conclude that the system is controllable. Therefore we stabilized it by inserting a controller K resulting in a closed loop feedback system. The system was simulated using the numerical method Runge-Kutta 4. We chose different sets of eigenvalues for the system with the controller inserted, to stabilise the system, and thereby keep the inverted pendulum in an upright position. We conclude that the set of eigenvalues from Set 1, scaled by 2.5:

$$\lambda = \{-5 \pm 7.5j, -7.5 \pm 7.5j\}$$

was the fastest set to stabilize the inverted pendulum after $\sim$ 1.66 seconds, which was $\sim$ 0.42 seconds faster than the second-fastest set. However this set also requires the largest amount of work $W = 120,039.63$ J to reach the equilibrium.

Bibliography

[1] Sheldon Axler. *Linear Algebra Done Right*. 3rd edition. ISBN 978-3-319-11079-0: Springer, 2015.

[2] David Betounes. *Differential Equations: Theory and Applications*. 2nd edition. ISBN 978-1-4419-1162-9: Springer, 2010.

[3] The SciPy community. *scipy.signal.place_poles*. Accessed on 9th December 2021. `https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.place_poles.html`, 2021.

[4] Maxon group. *Maxon 370356 DC motor*. Accessed on 10th December 2021. `https://maxongroup.com/maxon/view/product/motor/dcmotor/re/re50/370356`, 2021.

[5] Kristian Thostrup & Kim Hansen. *Formelsamling Fysik A*. Accessed on 18th October 2021. `http://www.formelsamlingen.com/bogform/FysikFormelsamlingRC1.pdf`, 2008.

[6] Harry L. Trentelman & Anton A. Stoorvogel & Malo Hautus. *Control theory for linear systems*. 1st edition. ISBN 978-1-4471-0339-4: Springer-Verlag London, 2002.

[7] Arne Jensen. *Matematikuddannelsernes 3. semester: Introduktion til sædvanlige differentialligning*. Updated on September 3rd 2021. 2021.

[8] Hassan K. Khalil. *Nonlinear Systems*. 3rd edition. ISBN 0-13-067389-7: Prentice Hall, 2002.

[9] David C. Lay. *Linear Algebra and Its Applications*. 5th edition. ISBN 978-1-292-09223-2: Lisbeth Fajstrup, 2016.

[10] John Leth. *A crash course in linear system theory*. 2021.

[11] L. D. Landau & E. M. Lifschitz. *Mechanics*. 3rd edition. ISBN 0-7506-2896-0: Butterworth-Heinemann, 1976.

[12] Oliver Matte. *Notes on differential equations*. 2021.

[13] Jacob Schach Møller. *Ordinary Differential Equations*. 2014.

[14] Ravi P. Agarwal & Donal O'Regan. *An Introduction to Ordinary Differential Equations*. 1st edition. ISBN 978-0-387-71275-8: Springer, 2008.

[15] Ebbe Thue Poulsen. *Funktioner af en og flere variable*. 1st edition. ISBN 8712038261: Cph.:Gad, 2001.

[16] Timothy Sauer. *Numerical Analysis*. 2nd edition. ISBN 978-0-321-78367-7: Pearson Education, Inc, 2012.

[17] Florian Scheck. *Mechanics*. 6th edition. ISBN 978-3-662-55490-6: Springer, 2018.

[18] Poul Seeburger. *Taylor Polynomials of Functions of Two Variables*. Accessed on 7th December 2021. `https://math.libretexts.org/Bookshelves/Calculus/Supplemental_Modules_(Calculus)/Multivariable_Calculus/3%3A_Topics_in_Partial_Derivatives/Taylor__Polynomials_of_Functions_of_Two_Variables`, 2020.

[19] Keith R. Symon. *Mechanics*. 3rd edition. ISBN 0-2010-7392-7: Menlo Park, CA:Addison-Wesley, 1971.

[20] *Torque and Angular Momentum in Circular Motion.* Accessed on 11th December 2021. `http://www.physnet.org/modules/pdf_modules/m34.pdf`, 2002.

[21] Niels Skov Vestergaard. *Swing-Up and Stabilization of a Cart Pendulum and Twin Pendulum System.* `https://projekter.aau.dk/projekter/files/294585861/thesisNielsSkovVestergaard.pdf`, 2018.

[22] *What Even is Stall Torque.* Accessed on 10th December 2021. `https://icrservices.com/2019/12/12/what-even-is-stall-torque/`, 2019.

[23] Jerzy Zabczyk. *Mathematical Control Theory.* 2nd edition. ISBN 978-0-8176-4732-2: Birkhäuser, 2008.

Appendix A

Computing e^{tA} for $\mathbb{R}^{2\times 2}$

After deriving Putzer's Algorithm (see Chapter 4), we here introduce a range of generalisations in regards to determining e^{tA} for $\mathbb{R}^{2\times 2}$ matrices.

The characteristic polynomial $P_A(\lambda)$ of a matrix $A \in \mathbb{R}^{2\times 2}$ can be expressed using trace (Definition 2.6) and the determinant of A:

$$P_A(\lambda) = \det(A - \lambda I) = \lambda^2 - \operatorname{tr}(A)\lambda + \det A. \tag{A.1}$$

To show (A.1), the matrices A and λI will be written:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}.$$

Then the determinant of the resulting matrix can be expressed as:

$$\begin{aligned}
\det(A - \lambda I) &= \det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix}\right) \\
&= (a_{11} - \lambda)(a_{22} - \lambda) - a_{21}a_{12} \\
&= a_{11}a_{22} - a_{11}\lambda - \lambda a_{22} + \lambda^2 - a_{21}a_{12} \\
&= \underbrace{a_{11}a_{22} - a_{21}a_{12}}_{\det A} - a_{11}\lambda - \lambda a_{22} + \lambda^2 \\
&= \lambda^2 - \underbrace{(a_{11} + a_{22})}_{\operatorname{tr} A}\lambda + \det A \\
&= \lambda^2 - \operatorname{tr}(A)\lambda + \det A.
\end{aligned}$$

If $\operatorname{tr} A = 0$, Theorem A.1 is true.

Theorem A.1 (Eigenvalues of a Matrix $A \in \mathbb{R}^{2\times 2}$ if tr $A = 0$)
Let a matrix $A \in \mathbb{R}^{2\times 2}$ have $\operatorname{tr} A = 0$. Then the following statements are true in regards to the eigenvalues λ_1,λ_2 of A:

1. If $\det A = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = 0.$
2. If $\det A < 0 \quad \Rightarrow \quad \lambda_1 = +\sqrt{-\det A}, \quad \lambda_2 = -\sqrt{-\det A}.$
3. If $\det A > 0 \quad \Rightarrow \quad \lambda_1 = +j\sqrt{\det A}, \quad \lambda_2 = -j\sqrt{\det A}.$

[13, p. 170] and [7, p. 11]

Proof

1. Assume $\det A = 0$. The eigenvalues λ can be found by solving the characteristic equation $P_A(\lambda) = 0$ per Definition 2.4:

$$0 = \det(A - \lambda I) = \lambda^2 - \underbrace{\operatorname{tr} A}_{0}\lambda + \underbrace{\det A}_{0} = \lambda^2 \quad \Rightarrow \quad \lambda = 0.$$

   Then $\lambda = 0$ with the algebraic multiplicity 2. ◇

2. Assume $\det A < 0$. The characteristic equation $P_A(\lambda) = 0$ with $\operatorname{tr} A = 0$ is:

$$\begin{aligned} \det(A - \lambda I) = \lambda^2 - \underbrace{\operatorname{tr} A}_{0}\lambda + \det A = \lambda^2 + \det A &= 0 \\ \lambda^2 &= -\det A \\ \lambda &= \pm\sqrt{-\det A}. \end{aligned}$$

   Therefore the eigenvalues are $\lambda = \pm\sqrt{-\det A}$ when $\det A < 0$. ◇

3. Assume $\det A > 0$. The characteristic equation $P_A(\lambda) = 0$ with $\operatorname{tr} A = 0$ is:

$$\begin{aligned} \det(A - \lambda I) = \lambda^2 - \underbrace{\operatorname{tr} A}_{0}\lambda + \det A = \lambda^2 + \det A &= 0 \\ \lambda^2 &= -\det A \\ \lambda &= \pm\sqrt{-\det A}. \end{aligned}$$

   However since $\det A > 0$ it is necessary to multiply the expression with the imaginary unit j. Therefore the eigenvalues when $\det A > 0$ are $\lambda = \pm\mathrm{j}\sqrt{\det A}$. ■

Theorem A.2 (Matrix Exponential Function e^{tA} of a Matrix $A \in \mathbb{R}^{2\times 2}$)
Let a matrix $A \in \mathbb{R}^{2\times 2}$ have $\operatorname{tr} A = 0$. Then the following statements of the exponential function e^{tA} are true.

1. If $\det A = 0$ then

$$\mathrm{e}^{tA} = I + tA.$$

2. If $\det A < 0$ then

$$\mathrm{e}^{tA} = \cosh\left(t\sqrt{-\det A}\right) I + \frac{\sinh\left(t\sqrt{-\det A}\right)}{\sqrt{-\det A}} A.$$

3. If $\det A > 0$ then

$$\mathrm{e}^{tA} = \cos\left(t\sqrt{\det A}\right) I + \frac{\sin\left(t\sqrt{\det A}\right)}{\sqrt{\det A}} A.$$

[13, p. 147] and [7, p. 11]

Proof

1. Assume $\det A = 0$ and $\operatorname{tr} A = 0$. Then the eigenvalues of A are $\lambda_1 = \lambda_2 = 0$, per Theorem A.1. From (4.3) it is known that:

$$\begin{aligned} P_0 &= I \\ P_1 &= A - \lambda_1 I = A - 0I = A. \end{aligned}$$

   Applying Putzer's Algorithm, $r_1(t)$ and $r_2(t)$ are:

$$\begin{aligned} r_1(t) &= \mathrm{e}^{\lambda_1 t} = \mathrm{e}^{0\cdot t} = \mathrm{e}^0 = 1, \\ r_2(t) &= \mathrm{e}^{t\lambda_2} \int_0^t \mathrm{e}^{-s\lambda_2} r_1(s)\mathrm{d}s, \\ &= 1 \int_0^t \mathrm{e}^{-s\cdot 0} 1\mathrm{d}s, \\ &= 1 \int_0^t 1 \cdot 1\mathrm{d}s \\ &= 1[s]_0^t = 1 \cdot (t - 0) \\ &= 1t \\ r_2(t) &= t. \end{aligned}$$

   Now inserting the values into the general equation (4.2):

$$\mathrm{e}^{tA} = r_1(t)P_0 + r_2(t)P_1 = 1 \cdot I + t \cdot A = I + tA. \qquad \diamond$$

2. Assume $\det A < 0$ and $\operatorname{tr} A = 0$, then by Theorem A.1 the eigenvalues $\lambda = \pm\sqrt{-\det A}$. We write the known equations:

$$\begin{aligned} P_0 &= I, \\ \lambda_1 &= +\sqrt{-\det A}, \\ \lambda_2 &= -\sqrt{-\det A}, \\ P_1 &= A - \lambda_1 I = A - \sqrt{-\det(A)} I, \\ r_1(t) &= \mathrm{e}^{\lambda_1 t} = \mathrm{e}^{\sqrt{-\det(A)} t}, \\ r_2(t) &= \mathrm{e}^{\lambda_2 t} \int_0^t \mathrm{e}^{-s\lambda_2} r_1(s)\mathrm{d}s. \end{aligned}$$

Then calculate $r_2(t)$:

$$\begin{aligned}
r_2(t) &= \mathrm{e}^{\lambda_2 t}\int_0^t \mathrm{e}^{-s\lambda_2} r_1(s)\mathrm{d}s,\\
&= \mathrm{e}^{-\sqrt{-\det(A)}t}\int_0^t \underbrace{\mathrm{e}^{-s\left(-\sqrt{-\det A}\right)}\mathrm{e}^{\sqrt{-\det(A)}s}}_{\mathrm{e}^{2s\sqrt{-\det A}}}\mathrm{d}s,\\
&= \mathrm{e}^{-\sqrt{-\det(A)}t}\int_0^t \mathrm{e}^{2s\sqrt{-\det A}}\mathrm{d}s,\\
&= \mathrm{e}^{-\sqrt{-\det(A)}t}\left[\frac{1}{2\sqrt{-\det A}}\mathrm{e}^{2s\sqrt{-\det A}}\right]_0^t,\\
&= \mathrm{e}^{-\sqrt{-\det(A)}t}\left(\frac{1}{2\sqrt{-\det A}}\mathrm{e}^{2t\sqrt{-\det A}} - \frac{1}{2\sqrt{-\det A}}\right),\\
&= \frac{1}{2\sqrt{-\det A}}\mathrm{e}^{\sqrt{-\det(A)}t} - \frac{1}{2\sqrt{-\det A}}\mathrm{e}^{-\sqrt{-\det(A)}t},\\
&= \frac{1}{\sqrt{-\det A}}\frac{1}{2}\left(\mathrm{e}^{\sqrt{-\det(A)}t} - \mathrm{e}^{-\sqrt{-\det(A)}t}\right).
\end{aligned}$$

Recalling that $\sinh(t) = \frac{1}{2}(\mathrm{e}^t - \mathrm{e}^{-t})$ then $r_2(t)$ can be rewritten as:

$$r_2(t) = \frac{\sinh\left(t\sqrt{-\det A}\right)}{\sqrt{-\det A}}.$$

Now inserting the values into the general equation (4.2):

$$\begin{aligned}
\mathrm{e}^{tA} &= r_1(t)P_0 + r_2(t)P_1\\
&= \mathrm{e}^{\sqrt{-\det(A)}t}I + \frac{\sinh\left(t\sqrt{-\det A}\right)}{\sqrt{-\det A}}\left(A - \sqrt{-\det(A)}I\right)\\
&= \mathrm{e}^{\sqrt{-\det(A)}t}I + \frac{\sinh\left(t\sqrt{-\det A}\right)}{\sqrt{-\det A}}A - \sinh\left(t\sqrt{-\det(A)}\right)I.
\end{aligned}$$

Now put I outside the parentheses:

$$= \left(\mathrm{e}^{\sqrt{-\det(A)}t} - \sinh\left(t\sqrt{-\det A}\right)\right)I + \frac{\sinh\left(t\sqrt{-\det A}\right)}{\sqrt{-\det A}}A.$$

Then we rewrite $\sinh\left(t\sqrt{-\det A}\right) = \frac{1}{2}\left(\mathrm{e}^{\sqrt{-\det(A)}t} - \mathrm{e}^{-\sqrt{-\det(A)}t}\right)$ but only inside the parentheses:

$$\begin{aligned}
&= \left(\mathrm{e}^{\sqrt{-\det(A)}t} - \frac{1}{2}\left(\mathrm{e}^{\sqrt{-\det(A)}t} - \mathrm{e}^{-\sqrt{-\det(A)}t}\right)\right)I + \frac{\sinh\left(t\sqrt{-\det A}\right)}{\sqrt{-\det A}}A\\
&= \left(\mathrm{e}^{\sqrt{-\det(A)}t} - \frac{1}{2}\mathrm{e}^{\sqrt{-\det(A)}t} + \frac{1}{2}\mathrm{e}^{-\sqrt{-\det(A)}t}\right)I + \frac{\sinh\left(t\sqrt{-\det A}\right)}{\sqrt{-\det A}}A\\
&= \left(\frac{1}{2}\mathrm{e}^{\sqrt{-\det(A)}t} + \frac{1}{2}\mathrm{e}^{-\sqrt{-\det(A)}t}\right)I + \frac{\sinh\left(t\sqrt{-\det A}\right)}{\sqrt{-\det A}}A\\
&= \frac{1}{2}\left(\mathrm{e}^{\sqrt{-\det(A)}t} + \mathrm{e}^{-\sqrt{-\det(A)}t}\right)I + \frac{\sinh\left(t\sqrt{-\det A}\right)}{\sqrt{-\det A}}A.
\end{aligned}$$

Since $\cosh(t) = \frac{1}{2}(\mathrm{e}^t + \mathrm{e}^{-t})$ the expression can be rewritten as:

$$\mathrm{e}^{tA} = \cosh\left(t\sqrt{-\det A}\right)I + \frac{\sinh\left(t\sqrt{-\det A}\right)}{\sqrt{-\det A}}A. \qquad \diamond$$

3. Assume $\det A > 0$ and $\operatorname{tr} A = 0$, then by Theorem A.1 the eigenvalues $\lambda = \pm j\sqrt{\det A}$. We will write the known equations:

$$\begin{aligned} P_0 &= I, \\ P_1 &= A - \lambda_1 I = A - j\sqrt{\det(A)}I, \\ r_1(t) &= \mathrm{e}^{tj\sqrt{\det A}}. \end{aligned}$$

Then calculate $r_2(t)$:

$$\begin{aligned} r_2(t) &= \mathrm{e}^{\lambda_2 t} + \int_0^t \mathrm{e}^{-s\lambda_2} r_1(s)\mathrm{d}s \\ &= \mathrm{e}^{-j\sqrt{\det(A)}t} \int_0^t \mathrm{e}^{-s\left(-j\sqrt{\det A}\right)} \mathrm{e}^{j\sqrt{\det A}}\mathrm{d}s \\ &= \mathrm{e}^{-j\sqrt{\det(A)}t} \int_0^t \mathrm{e}^{2sj\sqrt{\det A}}\mathrm{d}s \\ &= \mathrm{e}^{-i\sqrt{\det(A)}t} \left[\frac{1}{2j\sqrt{\det A}} \cdot \mathrm{e}^{2sj\sqrt{\det A}}\right]_0^t \\ &= \mathrm{e}^{-j\sqrt{\det(A)}t} \left(\frac{1}{2j\sqrt{\det A}} \cdot \mathrm{e}^{2tj\sqrt{\det A}} - \frac{1}{2j\sqrt{\det A}}\right) \\ &= \frac{1}{2j\sqrt{\det A}} \left(\mathrm{e}^{tj\sqrt{\det A}} - \mathrm{e}^{-tj\sqrt{\det A}}\right). \end{aligned}$$

We will recall the relation between sine and the exponential function:

$$\sin(x) = \frac{\mathrm{e}^{jx} - \mathrm{e}^{-jx}}{2j}$$

and use that to write $r_2(t)$ as

$$r_2(t) = \frac{\sin\left(t\sqrt{\det A}\right)}{\sqrt{\det A}}.$$

Now inserting the values into the general equation (4.2):

$$\begin{aligned} \mathrm{e}^{tA} &= r_1(t)P_0 + r_2(t)P_1 \\ &= \underbrace{\mathrm{e}^{tj\sqrt{\det A}}}_{r_1(t)} \underbrace{I}_{P_0} + \underbrace{\frac{\sin\left(t\sqrt{\det A}\right)}{\sqrt{\det A}}}_{r_2(t)} \cdot \underbrace{\left(A - j\sqrt{\det(A)}I\right)}_{P_1} \\ &= \mathrm{e}^{tj\sqrt{\det A}}I + \frac{\sin\left(t\sqrt{\det A}\right)}{\sqrt{\det A}}A - j\sin\left(t\sqrt{\det A}\right)I \\ &= \left(\mathrm{e}^{tj\sqrt{\det A}} - j\sin\left(t\sqrt{\det A}\right)\right) I + \frac{\sin\left(t\sqrt{\det A}\right)}{\sqrt{\det A}}A. \end{aligned}$$

Now we will rewrite

$$\sin\left(t\sqrt{\det A}\right) = \frac{\mathrm{e}^{tj\sqrt{\det A}} - \mathrm{e}^{-tj\sqrt{\det A}}}{2j},$$

but only inside the parenthesis:

$$\begin{aligned}
\mathrm{e}^{tA} &= \left(\mathrm{e}^{tj\sqrt{\det A}} - j\frac{\mathrm{e}^{tj\sqrt{\det A}} - \mathrm{e}^{-tj\sqrt{\det A}}}{2j}\right) I + \frac{\sin\left(t\sqrt{\det A}\right)}{\sqrt{\det A}} A \\
&= \left(\mathrm{e}^{tj\sqrt{\det A}} - \frac{j\mathrm{e}^{tj\sqrt{\det A}} - j\mathrm{e}^{-tj\sqrt{\det A}}}{2j}\right) I + \frac{\sin\left(t\sqrt{\det A}\right)}{\sqrt{\det A}} A \\
&= \left(\mathrm{e}^{tj\sqrt{\det A}} - \frac{1}{2}\left(\mathrm{e}^{tj\sqrt{\det A}} - \mathrm{e}^{-tj\sqrt{\det A}}\right)\right) I + \frac{\sin\left(t\sqrt{\det A}\right)}{\sqrt{\det A}} A \\
&= \left(\mathrm{e}^{tj\sqrt{\det A}} - \frac{1}{2}\mathrm{e}^{tj\sqrt{\det A}} + \frac{1}{2}\mathrm{e}^{-tj\sqrt{\det A}}\right) I + \frac{\sin\left(t\sqrt{\det A}\right)}{\sqrt{\det A}} A \\
&= \left(\frac{1}{2}\mathrm{e}^{tj\sqrt{\det A}} + \frac{1}{2}\mathrm{e}^{-tj\sqrt{\det A}}\right) I + \frac{\sin\left(t\sqrt{\det A}\right)}{\sqrt{\det A}} A \\
&= \left(\frac{1}{2}\left(\mathrm{e}^{tj\sqrt{\det A}} + \mathrm{e}^{-tj\sqrt{\det A}}\right)\right) I + \frac{\sin\left(t\sqrt{\det A}\right)}{\sqrt{\det A}} A.
\end{aligned}$$

We will recall the relation between cosine and the exponential function:

$$\cos(x) = \frac{\mathrm{e}^{jx} + \mathrm{e}^{-jx}}{2}.$$

Using this leads to the equation:

$$\mathrm{e}^{tA} = \cos\left(\sqrt{\det(A)}t\right) I + \frac{\sin\left(t\sqrt{\det A}\right)}{t\sqrt{\det A}} A. \tag{A.2}$$

Notice that the expression (A.2) is real even though the eigenvalues in this instance are complex. In general, it applies that for a real matrix A, then e^{tA} is also a real matrix. ■

Lemma A.1
Let $c \in \mathbb{R}$, then:

$$\mathrm{e}^{tcI} = \mathrm{e}^{tc} I.$$

[13, p.147] and [7, p. 11]

Proof
Since A can be written as $A = cI$ then e^{tcI} is just another expression for e^{tA}. The eigenvalues are both c such that $\lambda_1 = \lambda_2 = c$. The following are also known:

$$\begin{aligned}
P_0 &= I, \\
P_1 &= A - cI, \\
r_1(t) &= \mathrm{e}^{tc}.
\end{aligned}$$

We calculate $r_2(t)$:

$$\begin{aligned} r_2(t) &= \mathrm{e}^{ct} \int_0^t \mathrm{e}^{-sc} \cdot \mathrm{e}^{sc} \mathrm{d}s \\ &= \mathrm{e}^{ct} t. \end{aligned}$$

Now using Putzer's Algorithm:

$$\mathrm{e}^{tA} = \mathrm{e}^{tc} I + \mathrm{e}^{ct} t(A - cI),$$

and since it is known that $A = cI$, the last part of the equation equals 0. This leads to:

$$\mathrm{e}^{tcI} = \mathrm{e}^{tc} I.$$

■

Lemma A.2
Let $A \in \mathbb{R}^{2\times 2}$ then:

$$\operatorname{tr}\left(A - \frac{1}{2}\operatorname{tr}(A) I\right) = 0.$$

[13, p.147] and [7, p. 12]

Proof
We will start by calculating $\frac{1}{2}\operatorname{tr}(A)I$:

$$\frac{\operatorname{tr} A \cdot I}{2} = \frac{(a_{11} + a_{22})I}{2} = \begin{bmatrix} \frac{a_{11}+a_{22}}{2} & 0 \\ 0 & \frac{a_{11}+a_{22}}{2} \end{bmatrix}.$$

Then subtract the matrix from A and calculate the trace:

$$\begin{aligned} \operatorname{tr}\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} - \begin{bmatrix} \frac{a_{11}+a_{22}}{2} & 0 \\ 0 & \frac{a_{11}+a_{22}}{2} \end{bmatrix}\right) &= \operatorname{tr}\left(\begin{bmatrix} \frac{a_{11}}{2} - \frac{a_{22}}{2} & a_{12} \\ a_{21} & -\frac{a_{11}}{2} + \frac{a_{22}}{2} \end{bmatrix}\right) \\ &= \frac{a_{11}}{2} - \frac{a_{22}}{2} - \frac{a_{11}}{2} + \frac{a_{22}}{2} \\ &= 0. \end{aligned}$$

■

Lemma A.3
Let $A \in \mathbb{R}^{2\times 2}$ and $c = \frac{1}{2}\operatorname{tr} A$ then:

$$\mathrm{e}^{tA} = \mathrm{e}^{\left(\frac{1}{2}\operatorname{tr} A\right)t} \mathrm{e}^{t\left(A - \frac{1}{2}\operatorname{tr}(A) I\right)}.$$

[7, p. 12]

Proof
It is true that:

$$e^{t(B+cI)} = e^{tc}e^{tB}.$$

This will not be proven. Therefore

$$\begin{aligned}\exp\left(\underbrace{\left(\frac{1}{2}\operatorname{tr} A\right)}_{c} t\right)\exp\left(t\underbrace{\left(A-\frac{1}{2}\operatorname{tr}(A)I\right)}_{B}\right) &= e^{t(B+cI)}\\ &= e^{\left(t\left(A\underbrace{-\frac{1}{2}\operatorname{tr}(A)I+\frac{1}{2}\operatorname{tr}(A)I}_{0}\right)\right)}\\ &= e^{tA}.\end{aligned}$$

■

Theorem A.3
Let $A \in \mathbb{R}^{2\times 2}$ and $B = A - \frac{1}{2}\operatorname{tr}(A)I$. Then the eigenvalues of B are:

1. If $\det B = 0 \Rightarrow \lambda = \frac{1}{2}\operatorname{tr} A$.
2. If $\det B < 0 \Rightarrow \lambda = \frac{1}{2}\operatorname{tr}(A) \pm \sqrt{-\det B}$.
3. If $\det B > 0 \Rightarrow \lambda = \frac{1}{2}\operatorname{tr}(A) \pm j\sqrt{\det B}$.

[7, p. 12]

Proof

1. If $\det B = 0$.
   We write the general equation:

$$B - \lambda_B I = \left(\underbrace{A - \frac{1}{2}\operatorname{tr}(A)I}_{B}\right) - \left(\underbrace{\lambda - \frac{1}{2}\operatorname{tr} A}_{\lambda_B}\right) I.$$

   It is now possible to find an expression of λ:

$$\lambda_B = \lambda - \frac{1}{2}\operatorname{tr} A.$$

   We isolate λ:

$$\lambda = \frac{1}{2}\operatorname{tr}(A) + \lambda_B.$$

Since Lemma A.2 states that $\operatorname{tr}\left(A - \frac{1}{2}\operatorname{tr}(A)I\right) = 0$. We can use Theorem A.1 to realise that $\lambda_B = 0$. Therefore:

$$\lambda = \frac{1}{2}\operatorname{tr} A. \qquad \diamond$$

2. If $\det B < 0$.
   We look at the equation for λ again:

$$\lambda = \frac{1}{2}\operatorname{tr}(A) + \lambda_B.$$

   Again since Lemma A.2 states that $\operatorname{tr}(A - \frac{1}{2}\operatorname{tr}(A)I) = 0$. We can use Theorem A.1 to realise that $\lambda_B = \pm\sqrt{-\det B}$. Therefore:

$$\lambda = \frac{1}{2}\operatorname{tr}(A) \pm \sqrt{-\det B}. \qquad \diamond$$

3. If $\det B > 0$.
   We look at the equation for λ again:

$$\lambda = \frac{1}{2}\operatorname{tr}(A) + \lambda_B.$$

   Again since Lemma A.2 states that $\operatorname{tr}(A - \frac{1}{2}\operatorname{tr}(A)I) = 0$, we can use Theorem A.1 to realise that $\lambda_B = \pm j\sqrt{\det B}$. Therefore:

$$\lambda = \frac{1}{2}\operatorname{tr}(A) \pm j\sqrt{\det B}. \qquad \blacksquare$$

Theorem A.4
Let $A \in \mathbb{R}^{2\times 2}$, $c = \frac{1}{2}\operatorname{tr} A \in \mathbb{R}$ and $B = A - cI \in \mathbb{R}^{n\times n}$. Then the expressions for the matrix exponential function are:

1. If A has exactly one eigenvalue:

$$\mathrm{e}^{tA} = \mathrm{e}^{tc}(I + tB).$$

2. If A has two different real eigenvalues:

$$\mathrm{e}^{tA} = \mathrm{e}^{tc}\left(\cosh\left(t\sqrt{-\det B}\right) I + \frac{\sinh\left(t\sqrt{-\det B}\right)}{\sqrt{-\det B}} B\right).$$

3. If A has a pair of complex conjugate eigenvalues:

$$\mathrm{e}^{tA} = \mathrm{e}^{tc}\left(\cos\left(t\sqrt{\det B}\right) I + \frac{\sin\left(t\sqrt{\det B}\right)}{\sqrt{\det B}} B\right).$$

Proof
Let $A \in \mathbb{R}^{2\times 2}$, $B = A - \frac{1}{2}\mathrm{tr}(A)I$ and $c = \frac{1}{2}\mathrm{tr}\, A$.

1. If $\det B = 0$, then by Theorem A.3, A has exactly one eigenvalue.
   We know from Lemma A.3 that the exponential matrix can be written as:
   $$\mathrm{e}^{tA} = \mathrm{e}^{tc}\mathrm{e}^{tB}.$$
   By Theorem A.2 we can rewrite the expression as:
   $$\mathrm{e}^{tA} = \mathrm{e}^{tc}(I + tB). \qquad \diamond$$

2. If $\det\, B < 0$ then by Theorem A.3 A has two different real eigenvalues.
   By Lemma A.3 and Theorem A.2 we can write:
   $$\mathrm{e}^{tA} = \mathrm{e}^{tc}\mathrm{e}^{tB} = \mathrm{e}^{tc}\left(\cosh\left(t\sqrt{-\det B}\right) I + \frac{\sinh\left(t\sqrt{-\det B}\right)}{\sqrt{-\det B}} B\right). \qquad \diamond$$

3. If $\det\, B > 0$ then by Theorem A.3 A has a pair of complex conjugate eigenvalues.
   By Lemma A.3 and Theorem A.2 we can write:
   $$\mathrm{e}^{tA} = \mathrm{e}^{tc}\mathrm{e}^{tB} = \mathrm{e}^{tc}\left(\cos\left(t\sqrt{\det B}\right) I + \frac{\sin\left(t\sqrt{\det B}\right)}{\sqrt{\det B}} B\right). \qquad \blacksquare$$

Now that we have introduced Putzer's Algorithm and its properties in the special case where $A \in \mathbb{R}^{2\times 2}$, we can make an example.

Example A.1
Given the matrix
$$A = \begin{bmatrix} 6 & 5 \\ 0 & 7 \end{bmatrix},$$
we can calculate e^{tA}.

From Definition 2.5, we know that the eigenvalues of an upper-triangular matrix, such as A in this example, are the diagonal entries of the matrix, hence the eigenvalues being
$$\lambda_1 = 6 \text{ and } \lambda_2 = 7.$$

We acknowledge that this problem can be solved in two different ways: either by applying Putzer's Algorithm or utilizing Theorem A.4. Therefore, we will demonstrate both in order to show that the results are the same regardless of the method.

1. Firstly, we apply Putzer's Algorithm. We recall

$$P_0 = I \text{ and } P_k = (A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_k I), k = 1, 2, \ldots, n, \tag{4.3}$$

from which we calculate P_0 and P_1.

$$P_0 = I \qquad\qquad P_1 = (A - \lambda_1 I) = \begin{bmatrix} 0 & 5 \\ 0 & 1 \end{bmatrix}.$$

We determine $r_1(t)$ and $r_2(t)$:

$$\begin{aligned} r_1(t) &= \mathrm{e}^{\lambda_1 t} = \mathrm{e}^{6t}, \\ r_2(t) &= \mathrm{e}^{\lambda_2 t} \int_0^t \mathrm{e}^{-s\lambda_2} r_1(s) \mathrm{d}s, \\ &= \mathrm{e}^{7t} \int_0^t \mathrm{e}^{-7s} \mathrm{e}^{6s} \mathrm{d}s, \\ &= \mathrm{e}^{7t} \int_0^t \mathrm{e}^{-s} \mathrm{d}s, \\ &= \mathrm{e}^{7t} \left[-\mathrm{e}^{-s} \right]_0^t, \\ &= \mathrm{e}^{7t} \left(-\mathrm{e}^{-t} - \left(-\mathrm{e}^{-0} \right) \right), \\ &= \mathrm{e}^{7t} \left(-\mathrm{e}^{-t} - (-1) \right), \\ &= \mathrm{e}^{7t} \left(-\mathrm{e}^{-t} + 1 \right), \\ &= \mathrm{e}^{7t} - \mathrm{e}^{7t} \mathrm{e}^{-t}, \\ &= \mathrm{e}^{7t} - \mathrm{e}^{6t}. \end{aligned}$$

We can now calculate e^{tA}:

$$\begin{aligned} \mathrm{e}^{tA} &= P_0 r_1(t) + P_1 r_2(t) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathrm{e}^{6t} + \begin{bmatrix} 0 & 5 \\ 0 & 1 \end{bmatrix} (\mathrm{e}^{7t} - \mathrm{e}^{6t}) \\ &= \begin{bmatrix} \mathrm{e}^{6t} & 0 \\ 0 & \mathrm{e}^{6t} \end{bmatrix} + \begin{bmatrix} 0 & 5(\mathrm{e}^{7t} - \mathrm{e}^{6t}) \\ 0 & \mathrm{e}^{7t} - \mathrm{e}^{6t} \end{bmatrix} \\ &= \begin{bmatrix} \mathrm{e}^{6t} & 5(\mathrm{e}^{7t} - \mathrm{e}^{6t}) \\ 0 & \mathrm{e}^{7t} \end{bmatrix}. \end{aligned} \tag{A.3}$$

2. Secondly, we utilise Theorem A.4, from which we recall that the matrix exponential function for any given matrix $A \in \mathbb{R}^{2 \times 2}$ with two real and distinct eigenvalues can be written as

$$\mathrm{e}^{tA} = \mathrm{e}^{tc} \left(\cosh\left(t\sqrt{-\det B} \right) I + \frac{\sinh\left(t\sqrt{-\det B} \right)}{\sqrt{-\det B}} B \right). \tag{A.4}$$

We now compute $c = \frac{1}{2}\mathrm{tr}(A)$ and $B = A - cI$:

$$\begin{aligned} c &= \frac{1}{2}(6+7) = \frac{13}{2}, \\ B &= \begin{bmatrix} 6 & 5 \\ 0 & 7 \end{bmatrix} - \frac{13}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\ &= \begin{bmatrix} 6 & 5 \\ 0 & 7 \end{bmatrix} - \begin{bmatrix} \frac{13}{2} & 0 \\ 0 & \frac{13}{2} \end{bmatrix}, \\ &= \begin{bmatrix} -\frac{1}{2} & 5 \\ 0 & \frac{1}{2} \end{bmatrix}. \end{aligned}$$

We now compute the determinant of B:

$$\mathrm{det}B = \left(-\frac{1}{2}\right)\left(\frac{1}{2}\right) - 0 \cdot 5 = \left(-\frac{1}{2}\right)\left(\frac{1}{2}\right) = -\frac{1}{4}.$$

Using Theorem A.3, we can confirm that the eigenvalues are in fact

$$\lambda = \frac{1}{2}\mathrm{tr}A \pm \sqrt{-\mathrm{det}B} = \frac{1}{2}(6+7) \pm \frac{1}{2} = \frac{13}{2} \pm \frac{1}{2} \quad \Rightarrow \quad \lambda_1 = 6, \quad \lambda_2 = 7$$

given that $\mathrm{det}B = -\frac{1}{4} < 0$.

We now insert c, B, and $\mathrm{det}B$ into (A.4) and thereby yield:

$$\begin{aligned} \mathrm{e}^{tA} &= \mathrm{e}^{\frac{13}{2}t}\left(\frac{1}{2}\left(\mathrm{e}^{\frac{1}{2}t} + \mathrm{e}^{-\frac{1}{2}t}\right)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\frac{1}{2}\left(\mathrm{e}^{\frac{1}{2}t} - \mathrm{e}^{-\frac{1}{2}t}\right)}{\frac{1}{2}}\begin{bmatrix} -\frac{1}{2} & 5 \\ 0 & \frac{1}{2} \end{bmatrix}\right) \\ &= \mathrm{e}^{\frac{13}{2}t}\left(\frac{1}{2}\left(\mathrm{e}^{\frac{1}{2}t} + \mathrm{e}^{-\frac{1}{2}t}\right)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \left(\mathrm{e}^{\frac{1}{2}t} - \mathrm{e}^{-\frac{1}{2}t}\right)\begin{bmatrix} -\frac{1}{2} & 5 \\ 0 & \frac{1}{2} \end{bmatrix}\right) \\ &= \mathrm{e}^{\frac{13}{2}t}\frac{1}{2}\left(\mathrm{e}^{\frac{1}{2}t} + \mathrm{e}^{-\frac{1}{2}t}\right)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \mathrm{e}^{\frac{13}{2}t}\left(\mathrm{e}^{\frac{1}{2}t} - \mathrm{e}^{-\frac{1}{2}t}\right)\begin{bmatrix} -\frac{1}{2} & 5 \\ 0 & \frac{1}{2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\mathrm{e}^{\frac{13}{2}t}\left(\mathrm{e}^{\frac{1}{2}t}+\mathrm{e}^{-\frac{1}{2}t}\right)}{2} & 0 \\ 0 & \frac{\mathrm{e}^{\frac{13}{2}t}\left(\mathrm{e}^{\frac{1}{2}t}+\mathrm{e}^{-\frac{1}{2}t}\right)}{2} \end{bmatrix} + \begin{bmatrix} -\frac{\mathrm{e}^{\frac{13}{2}t}\left(\mathrm{e}^{\frac{1}{2}t}-\mathrm{e}^{-\frac{1}{2}t}\right)}{2} & 5\left(\mathrm{e}^{7t} - \mathrm{e}^{6t}\right) \\ 0 & \frac{\mathrm{e}^{\frac{13}{2}t}\left(\mathrm{e}^{\frac{1}{2}t}-\mathrm{e}^{-\frac{1}{2}t}\right)}{2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\mathrm{e}^{7t}}{2} + \frac{\mathrm{e}^{6t}}{2} - \frac{\mathrm{e}^{7t}}{2} + \frac{\mathrm{e}^{6t}}{2} & 5\left(\mathrm{e}^{7t} - \mathrm{e}^{6t}\right) \\ 0 & \frac{\mathrm{e}^{7t}}{2} + \frac{\mathrm{e}^{6t}}{2} + \frac{\mathrm{e}^{7t}}{2} - \frac{\mathrm{e}^{6t}}{2} \end{bmatrix} \\ &= \begin{bmatrix} \mathrm{e}^{6t} & 5\left(\mathrm{e}^{7t} - \mathrm{e}^{6t}\right) \\ 0 & \mathrm{e}^{7t} \end{bmatrix}. \end{aligned} \tag{A.5}$$

The two matrices (A.3) and (A.5) are equivalent.

Appendix B

Computing e^{tA} for $\mathbb{R}^{3\times 3}$

We can make a lot of generalizations for $\mathbb{R}^{2\times 2}$ (see Appendix A). However, this is not the case for $\mathbb{R}^{n\times n}, 2 < n \in \mathbb{N}$. The following Example B.1 illustrates the intricacy of calculating the matrix exponential function e^{tA} for a 3×3 matrix. As you can imagine, calculating e^{tA} for $A \in \mathbb{R}^{n\times n}$ only gets increasingly more difficult the higher $n \in \mathbb{N}$ is.

Example B.1
Let a matrix $A \in \mathbb{R}^{3\times 3}$ be given by

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

To calculate the matrix exponential function e^{tA} the eigenvalues are first determined. The characteristic polynomial is

$$\begin{aligned} P_A(\lambda) &= \det(A - \lambda I) \\ &= \det\left(\begin{bmatrix} 1-\lambda & 1 & 0 \\ 0 & 0-\lambda & 1 \\ 0 & 1 & 0-\lambda \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} 1-\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{bmatrix}\right) \\ &= (1-\lambda)\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} \underbrace{- 0\begin{vmatrix} 1 & 0 \\ 1 & -\lambda \end{vmatrix} + 0\begin{vmatrix} 1 & 0 \\ -\lambda & 1 \end{vmatrix}}_{0} \\ &= (1-\lambda)((-\lambda)^2 - 1) \\ &= (1-\lambda)(\lambda^2 - 1) \\ &= \lambda^2 - 1 - \lambda^3 + \lambda. \end{aligned}$$

Then the characteristic polynomial is set to zero, to find the eigenvalues

$$0 = \lambda^2 - 1 - \lambda^3 + \lambda \quad \Rightarrow \quad \lambda_1 = -1, \quad \lambda_2 = 1 \text{ (alg. mult. 2).}$$

Here the algebraic multiplicity is 2, therefore the three eigenvalues for A are:

$$\lambda_1 = -1, \quad \lambda_2 = 1, \quad \lambda_3 = 1.$$

Since there are three eigenvalues, the matrices P_0, P_1 and P_2 must be determined. The matrix P_0 is known:

$$P_0 = I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The matrices P_1 and P_2 must be calculated. First, we calculate P_1:

$$P_1 = (A - \lambda_1 I) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} - (-1) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Next, we calculate P_2:

$$\begin{aligned} P_2 &= (A - \lambda_1 I)(A - \lambda_2 I) \\ &= \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \\ &= \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

We compute $r_1(t)$:

$$r_1(t) = \mathrm{e}^{t\lambda_1} = \mathrm{e}^{-t}.$$

We compute $r_2(t)$:

$$\begin{aligned} r_2(t) &= \mathrm{e}^{t\lambda_2} \int_0^t \mathrm{e}^{-s\lambda_2} r_1(s) \mathrm{d}s \\ &= \mathrm{e}^t \int_0^t \mathrm{e}^{-s} \mathrm{e}^{-s} \mathrm{d}s \\ &= \mathrm{e}^t \int_0^t \mathrm{e}^{-2s} \mathrm{d}s \\ &= \mathrm{e}^t \left[-\frac{1}{2} \mathrm{e}^{-2s} \right]_0^t \\ &= \mathrm{e}^t \left(\left(-\frac{1}{2} \mathrm{e}^{-2t} \right) - \left(-\frac{1}{2} \mathrm{e}^{-2\cdot 0} \right) \right) \\ &= \mathrm{e}^t \left(-\frac{1}{2} \mathrm{e}^{-2t} + \frac{1}{2} \right) \\ &= -\frac{1}{2} \mathrm{e}^{-2t} \mathrm{e}^t + \frac{1}{2} \mathrm{e}^t \\ &= -\frac{1}{2} \mathrm{e}^{-t} + \frac{1}{2} \mathrm{e}^t \\ &= \frac{1}{2} \left(\mathrm{e}^t - \mathrm{e}^{-t} \right) \end{aligned} \tag{B.1}$$

We calculate $r_3(t)$ using (B.1):

$$
\begin{aligned}
r_3(t) &= \mathrm{e}^{t\lambda_3} \int_0^t \mathrm{e}^{-s\lambda_3} r_2(s)\mathrm{d}s \\
&= \mathrm{e}^t \int_0^t \mathrm{e}^{-s} \frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right)\mathrm{d}s \\
&= \mathrm{e}^t \left[-\mathrm{e}^{-s}\frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right)\right]_0^t \\
&= \mathrm{e}^t \left(\left(-\mathrm{e}^{-t}\frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right)\right) - \left(-\mathrm{e}^{-0}\frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right)\right)\right) \\
&= \mathrm{e}^t \left(\left(-\mathrm{e}^{-t}\frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right)\right) - \left(-\frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right)\right)\right) \\
&= \mathrm{e}^t \left(-\mathrm{e}^{-t}\frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right) + \frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right)\right) \\
&= \underbrace{-\mathrm{e}^{-t}\mathrm{e}^t}_{-1} \frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right) + \frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right)\mathrm{e}^t \\
&= -\frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right) + \frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right)\mathrm{e}^t.
\end{aligned}
$$

Now the matrix exponential function e^{tA} can be calculated:

$$
\begin{aligned}
\mathrm{e}^{tA} &= r_1(t)P_0 + r_2(t)P_1 + r_3(t)P_2 \\
&= \mathrm{e}^{-t}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} + \frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right)\begin{bmatrix}2 & 1 & 0\\ 0 & 1 & 1\\ 0 & 1 & 1\end{bmatrix} + \left(-\frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right) + \frac{1}{2}\left(\mathrm{e}^t - \mathrm{e}^{-t}\right)\mathrm{e}^t\right)\begin{bmatrix}0 & 1 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \\
&= \begin{bmatrix}\mathrm{e}^t & \mathrm{e}^t\sinh t & -\sinh t + \mathrm{e}^t \sinh t\\ 0 & \cosh t & \sinh t\\ 0 & \sinh t & \cosh t\end{bmatrix}.
\end{aligned}
$$

Appendix C

Simulation Using Set 1 of Eigenvalues

In this appendix, we see all the figures generated from simulation that are neither relevant in the context of the simulation (see Section 9.3.1), nor full size.

C.1 Velocity Graph

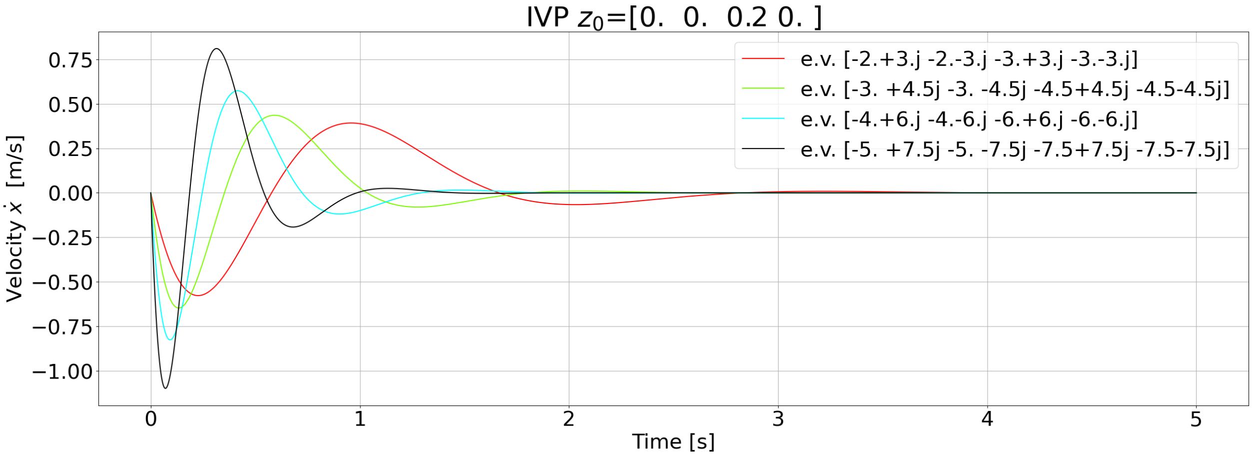


Figure C.1: Velocity $\dot{x}$ of the cart. [Made using the `matplotlib.pyplot` module in `Python`.]

C.2 Full Size Convergence Graphs (Figure 9.5)

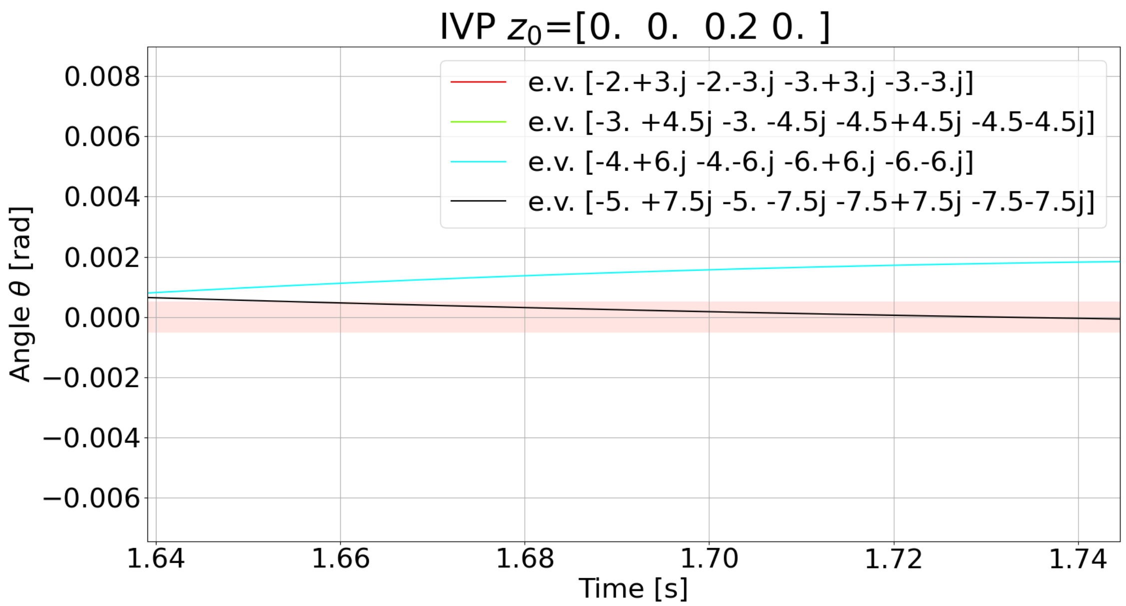


Figure C.2: Convergence with $\lambda = \{-5 \pm 7.5j, -7.5 \pm 7.5j\}$. [Made using the `matplotlib.pyplot` module in `Python`.]

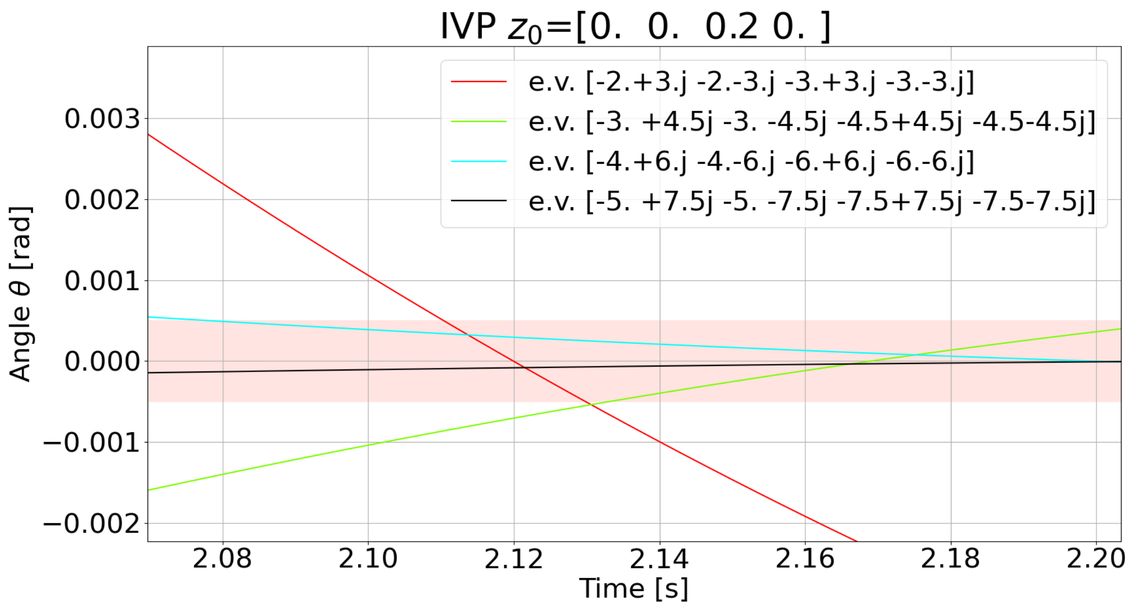


Figure C.3: Convergence with $\lambda = \{-4 \pm 6j, -6 \pm 6j\}$. [Made using the `matplotlib.pyplot` module in `Python`.]

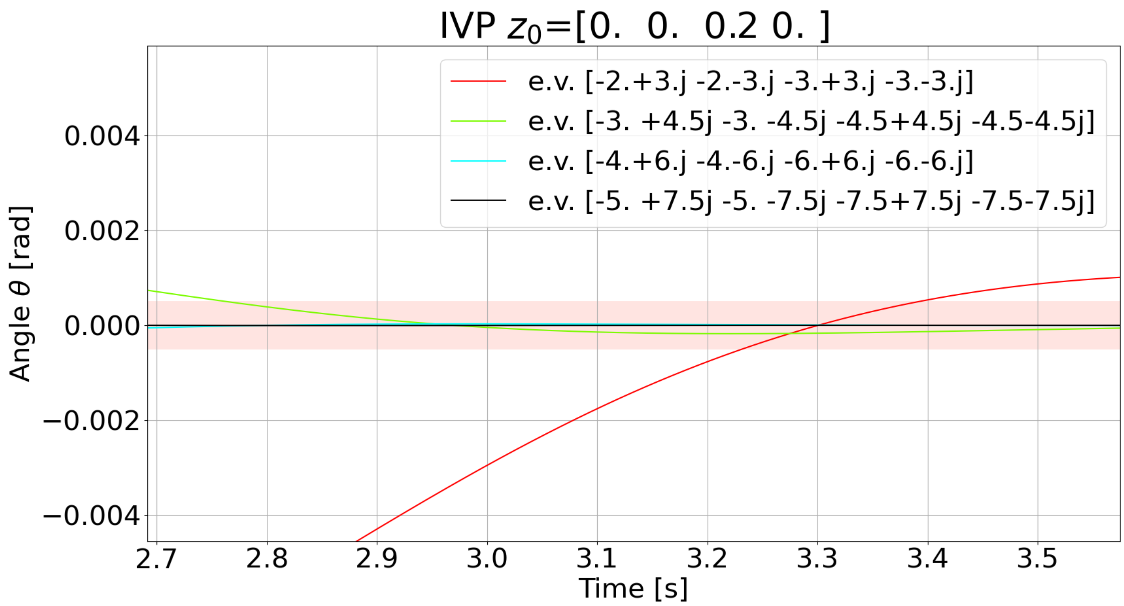


Figure C.4: Convergence with $\lambda = \{-3 \pm 4.5j, -4.5 \pm 4.5j\}$. [Made using the `matplotlib.pyplot` module in `Python`.]

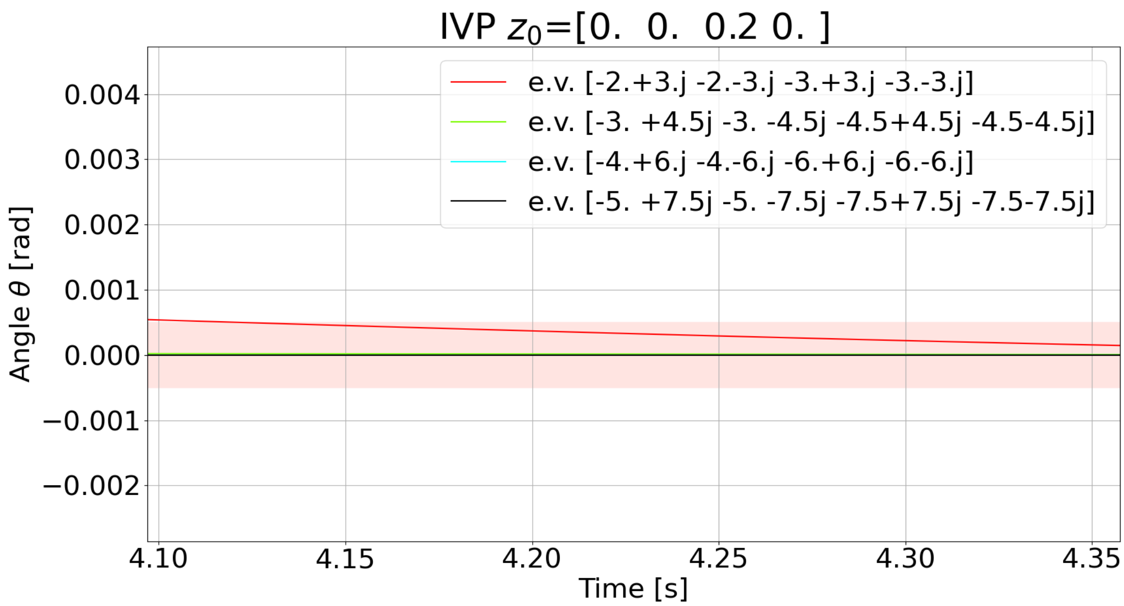


Figure C.5: Convergence with $\lambda = \{-2 \pm 3j, -3 \pm 3j\}$. [Made using the `matplotlib.pyplot` module in `Python`.]

C.3 Angular Velocity Graph

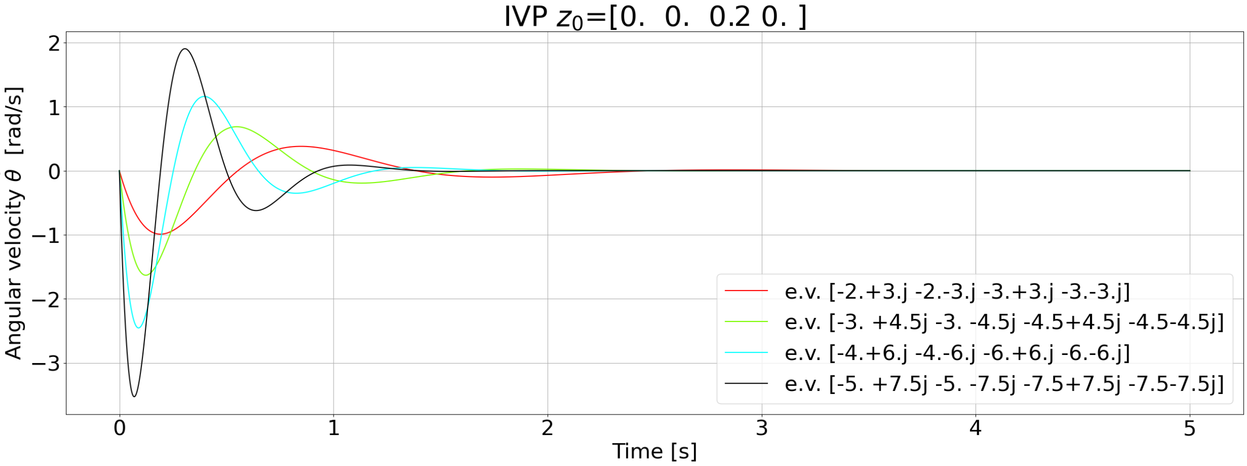


Figure C.6: Angular velocity $\dot{\theta}$ of the pendulum. [Made using the `matplotlib.pyplot` module in `Python`.]

Appendix D

Simulation Using Set 2 of Eigenvalues

In this appendix, we see all the figures generated from simulation that are neither relevant in the context of the simulation (see Section 9.3.2), nor full size.

D.1 Velocity Graph

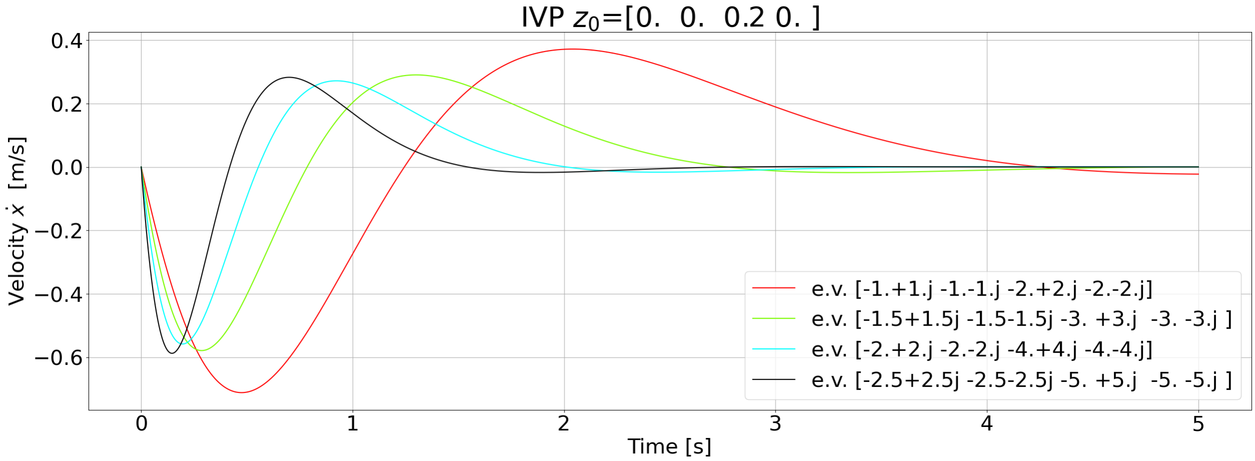


Figure D.1: Velocity $\dot{x}$ of the cart. [Made using the `matplotlib.pyplot` module in `Python`.]

D.2 Full Size Convergence Graphs (Figure 9.9)

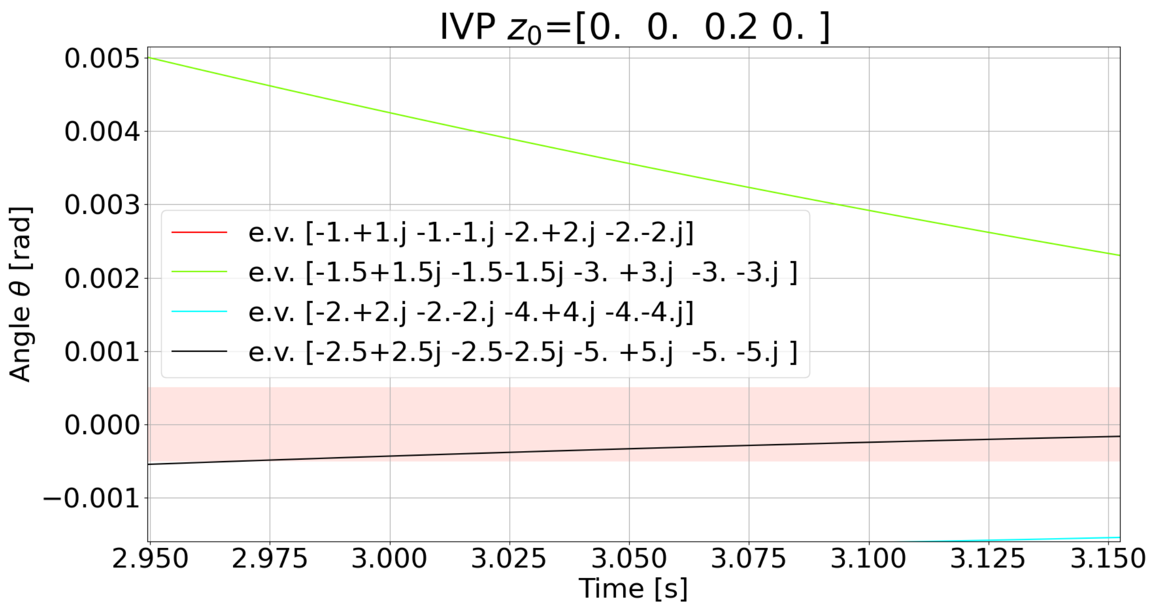


Figure D.2: Convergence with $\lambda = \{-2.5 \pm 2.5j, -5 \pm 5j\}$. [Made using the `matplotlib.pyplot` module in `Python`.]

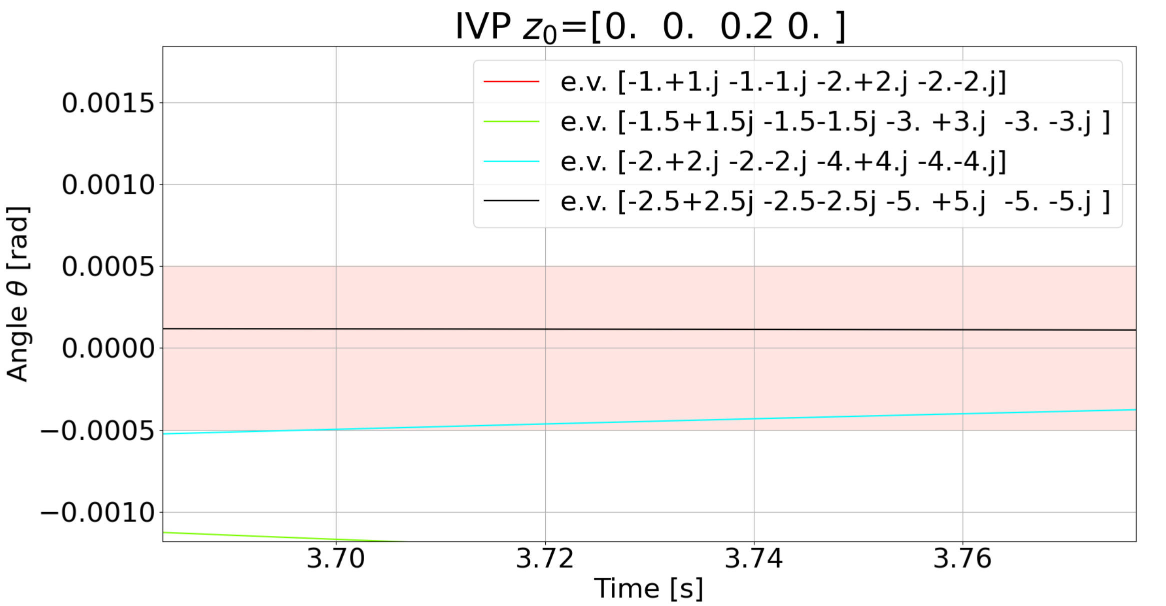


Figure D.3: Convergence with $\lambda = \{-2 \pm 2j, -4 \pm 4j\}$. [Made using the `matplotlib.pyplot` module in `Python`.]

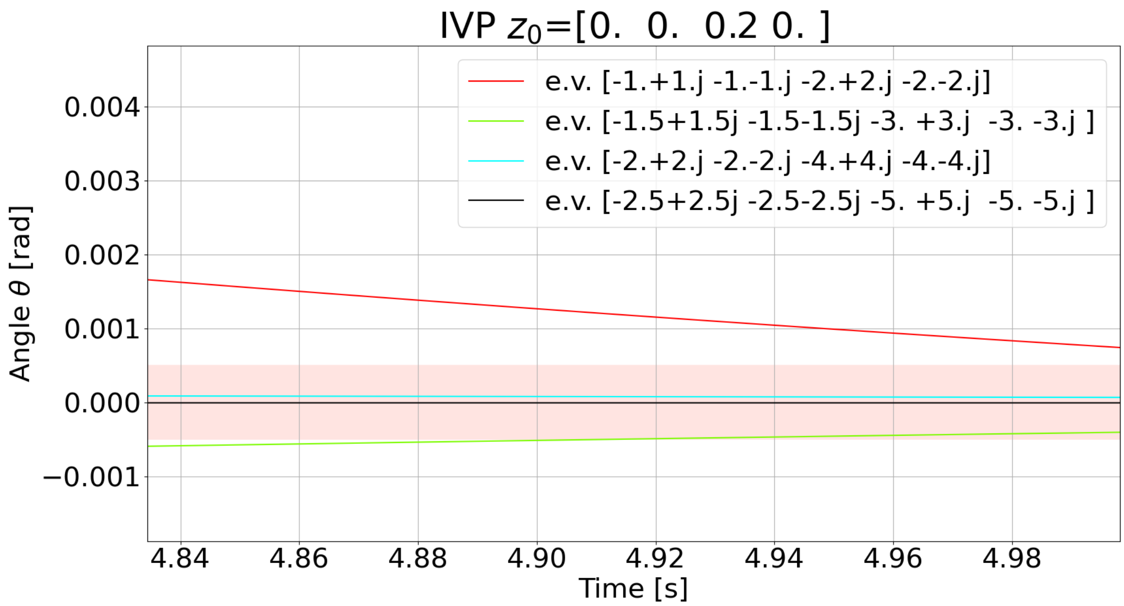


Figure D.4: Convergence with $\lambda = \{-1.5 \pm 1.5j, -3 \pm 3j\}$. [Made using the `matplotlib.pyplot` module in `Python`.]

D.3 Angular Velocity Graph

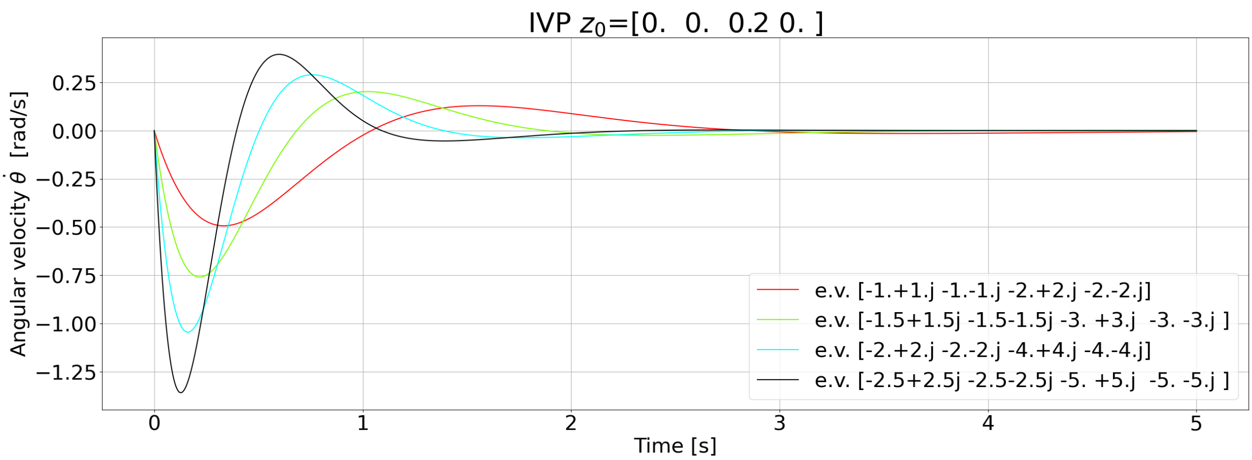


Figure D.5: Angular velocity $\dot{\theta}$ of the pendulum. [Made using the `matplotlib.pyplot` module in `Python`.]