Homework I: Modeling, Representation and Simulation of dynamical systems

Universidade Federal do Ceará

Eduardo Fernandes Montesuma - 374924

April 19, 2018

Contents

1 Modeling

1.1 Introduction

Throughout the next sections, and also through the next homeworks, we shall deal with the inverted control pendulum system, whose modeling, representation and simulation are discussed in this first chapter. However, before entering in the ideas behind the dynamic behavior of it, we shall make a few remarks.

First, as displayed in Figure 1, the inverted pendulum model finds applications in the control of segways and related devices. Second, it is a benchmark system, as stated in [1], for many control techniques. Being so, we shall use it to illustrate the concepts learned in class.

Figure 1: Example of a segway

Also, as a matter of clarity, we mean by,

- Modeling: the derivation and elucidation of dynamic equations of system's motion, through physics. After having the complete equations, we shall linearize them, to later on represent them.
- Representation: write the linearized equations through state-space models, and in frequency, with transfer functions.
- Simulation: to see how the system behaves in light of various inputs, in the time domain.

With those definitions, we intent to present a clear and objective guide, hoping to validate benchmark results. As tools to illustrate them, we shall employ programs in mainly two languages, Python and Matlab. You can also find out in the homework's website[1] the associated resources, as well as animations of the dynamic behavior of our system.

1.2 Physical Analysis

Systems like the already mentioned Segway device can be modeled as a cart (the basis, where one stands) with a pole attached to it (the person standing up). In such a representation, we have two bodies, the cart, with mass M, position x, velocity $\dot{x}$ and being under action of input force $u(t) = F(t)$ and friction, b_x and the pole, with a mass attached to its extrema, with length ℓ, angle deviation[2] θ, angular velocity ω and suffering from a friction between the cart and the junction b_θ.

This description is displayed Graphically in Figure 2,

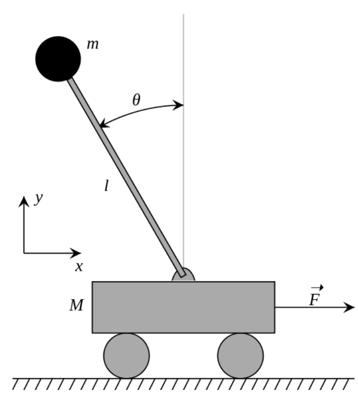

Figure 2: System's diagram

Also, we have some constraints over the defined variables, we assume the joint between the cart and the pole to be free to spin around, in such a way that, if we assume θ to be in radians, $0 \leq \theta \leq 2\pi$, where 0 is the upward position, and π, the downward[3], also, The total length of the rail is 2 meters. Assuming the 0 position, the cart's position satisfies the following constraint, $-1 \leq x \leq 1$ and finally, the input force is limited to a value of $120N$, that is, $|u(t)| \leq 120$.

[1] `https://github.com/eddardd/Control-Theory/tree/master/Advanced-Control`

[2] from the upward position

[3] Notice we are not imposing the segway angle's condition, for simplicity

Except from the first constraint, all other can be found in [1]. Also based on such paper, Table 1, bellow, shows a complete list of constants used in our models, and their description,

Symbol	Description	Value	Unit
θ	Angular position of the Pendulum	variable	radians
x	Cart's linear position	variable	meters
F	external force applied to the cart	variable	Newton
M	Cart's mass	4.8	kilogram
m	Point mass of pendulum	0.356	kilogram
ℓ	pole length	0.56	meters
b_{th}	joint's friction	0.035	Nms/rad
b_x	track's friction	4.9	Newton
g	Gravitational constant	9.81	meter/second^2
I	Pole's moment of inertia	0.006	$kilogram \cdot meter^2$
L	total length of rail	2	meters
F_{max}	Maximum input value	120	Newton

Table 1: System constants

Therefore, our system will be considered controlled if we stabilize the upward position, $\theta = 0$ with the additional constraint of not letting the cart go off the track. As we shall see through our simulations, it is often the case where such situation happens.

1.3 Dynamic Equations

We adopt as modeling strategy the Lagrangian formalism of classical mechanics, which is based on variational principles[4]. In order to take advantage of it, we recall the definition of Lagrangian, $\mathcal{L} = T - V$, where T is total kinetic energy in the system, and V, the total potential. Those later two quantities can be evaluated as,

$$T = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m||v_m||^2 + \frac{1}{2}I\omega^2 \tag{1}$$

$$V = mg\ell cos(\theta) \tag{2}$$

[4]the necessary theory for the curious reader may be found in Appendix 5

Here, we recognize that v_m, the velocity vector of the mass attached to the pole's extrema, is given by:

$$v_m = \left(\frac{d}{dt}\left(x - \ell sin(\theta)\right), \left(\frac{d}{dt}\left(\ell cos(\theta)\right)\right)\right) \tag{3}$$

$$= \left(x - \ell x\dot{\theta} cos(\theta), -\ell\dot{\theta} sin(\theta)\right) \tag{4}$$

By taking its squared euclidean norm, we conclude that,

$$||v_m||^2 = \dot{x}^2 - 2\dot{x}\dot{\theta}\ell cos(\theta) + \ell^2\dot{\theta}^2 \tag{5}$$

and so,

$$T = \frac{1}{2}(M+m)\dot{x}^2 - m\dot{x}\dot{\theta}\ell cos(\theta) + \frac{1}{2}(m\ell^2 + I)\dot{\theta}^2 \tag{6}$$

Therefore, we conclude that,

$$\mathcal{L} = \frac{1}{2}(M+m)\dot{x}^2 - m\dot{x}\dot{\theta}\ell cos(\theta) + \frac{1}{2}(m\ell^2 + I)\dot{\theta}^2 - mg\ell cos(\theta) \tag{7}$$

We shall now take our attention to the definition of action $\mathcal{S}$ within a physical system. The action can be formally defined as the integral of system's total energy through time[5],

$$\mathcal{S} = \int_{t_0}^{t} \mathcal{L}(\tau, x, \dot{x}, \theta, \dot{\theta}) d\tau \tag{8}$$

With such definition, we state the Principle of Least Action, which says: "The path taken by the system between times t_0 and t and configurations q_0 and q is the one for which the action is stationary to first order"[2] which here we take as an axiom. Under the Principle of Least Action, $\delta\mathcal{S} = 0$, we can conclude by appendix A[6], the functional $\mathcal{L}$ needs to satisfy the Euler-Lagrange equations, for variables x and θ,

[5]Which, in such case, is the Lagrangian itself.

[6]Appendix A contains a more formal discussion of those concepts, based on references [5] and [6]

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} - \frac{\partial \mathcal{L}}{\partial x} = F - b_x \dot{x} \tag{9}$$
$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\theta}} - \frac{\partial \mathcal{L}}{\partial \theta} = -b_\theta \dot{\theta} \tag{10}$$

Now, all we have to do is to take partial derivatives of $\mathcal{L}$, in order to use Equations 9 and 10 to describe the system's behavior. Being so,

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} - \frac{\partial \mathcal{L}}{\partial x} = (M+m)\ddot{x} - m\ell\ddot{\theta}cos(\theta) - m\ell\dot{\theta}^2 sin(\theta) \tag{11}$$
$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\theta}} - \frac{\partial \mathcal{L}}{\partial \theta} = -m\ell cos(\theta)\ddot{x} + m\ell\dot{\theta}^2 \dot{x} sin(\theta) + (m\ell^2 + I)\ddot{\theta} - m\ell\dot{\theta}^2\dot{x} sin(\theta) - mg\ell sin(\theta) \tag{12}$$

or, still,

$$(M+m)\ddot{x} - m\ell cos(\theta)\ddot{\theta} = F + m\ell\dot{\theta} sin(\theta) - b_x\dot{x} \tag{13}$$
$$-m\ell cos(\theta)\ddot{x} + (m\ell^2 + I)\ddot{\theta} = mg\ell sin(\theta) - b_\theta\dot{\theta} \tag{14}$$

These are two non-linear differential equations in terms of $\ddot{x}$ and $\ddot{\theta}$. To effectively represent the inverted pendulum dynamics, however, we need to solve 13 and 14 explicitly. To do so, we make use of Cramer"s Rule,

$$\begin{aligned}
\alpha &= \begin{vmatrix} M+m & -m\ell cos(\theta) \\ -m\ell cos(\theta) & (m\ell^2 + I) \end{vmatrix} \\
&= (M+m)(m\ell^2 + I) - m^2\ell^2 cos^2(\theta) \qquad (15)\\
N_x &= \begin{vmatrix} F + m\ell\dot{\theta} sin(\theta) - b_x\dot{x} & -m\ell cos(\theta) \\ mg\ell sin(\theta) - b_\theta\dot{\theta} & (m\ell^2 + I) \end{vmatrix} \\
&= (m\ell^2 + I)(F + m\ell\dot{\theta} sin(\theta) - b_x\dot{x}) + m\ell cos(\theta)(mg\ell sin(\theta) - b_\theta\dot{\theta}) \qquad (16)\\
N_\theta &= \begin{vmatrix} M+m & F + m\ell\dot{\theta} sin(\theta) - b_x\dot{x} \\ -m\ell cos(\theta) & mg\ell sin(\theta) - b_\theta\dot{\theta} \end{vmatrix} \\
&= (M+m)(mg\ell sin(\theta) - b_\theta\dot{\theta}) + m\ell cos(\theta)(F + m\ell\dot{\theta} sin(\theta) - b_x\dot{x}) \qquad (17)
\end{aligned}$$

With those expressions, $\ddot{x} = N_x/\alpha$ and $\ddot{\theta} = N_\theta/\alpha$, that is,

$$\ddot{x} = \frac{(m\ell^2 + I)(F + m\ell\dot{\theta}sin(\theta) - b_x\dot{x}) + m\ell cos(\theta)(mg\ell sin(\theta) - b_\theta\dot{\theta})}{(M+m)(m\ell^2 + I) - m^2\ell^2cos^2(\theta)} \tag{18}$$

$$\ddot{\theta} = \frac{(M+m)(mg\ell sin(\theta) - b_\theta\dot{\theta}) + m\ell cos(\theta)(F + m\ell\dot{\theta}sin(\theta) - b_x\dot{x})}{(M+m)(m\ell^2 + I) - m^2\ell^2cos^2(\theta)} \tag{19}$$

Equations 18 and 19 describe how the system evolve in time. They are non-linear, since they contain sines and cosines of θ, as well as α, which is itself a non-linear function of θ. The strong dependence of both equations in trigonometric functions reveals a heavy oscillatory response. Such claim shall be confirmed when we simulate the system to various inputs.

2 Representation

2.1 Nonlinear state-space equations

We recall Equations 18 and 19 which are two independently second-order ODE. As consequence, our system has two equations, each of which has two degrees of freedom, that is, our system has to be described by four state variables. A natural choice for such variables is, indeed, to pick $\mathbf{x} = (x, \dot{x}, \theta, \dot{\theta})$, hence,

$$\dot{x}_1 = \dot{x} = x_2 \tag{20}$$

$$\dot{x}_2 = \ddot{x} = \text{Equation 1.18} \tag{21}$$

$$\dot{x}_3 = \dot{\theta} = x_4 \tag{22}$$

$$\dot{x}_4 = \ddot{\theta} = \text{Equation 1.19} \tag{23}$$

Being $\mathcal{U} = \mathbb{R}$, the input space, our (nonlinear) state space model can be described by a function, $\boldsymbol{\phi} : \mathbb{R} \times \mathbb{R}^4 \times \mathbb{R}$ such that $\dot{\mathbf{x}} = \boldsymbol{\phi}(t, \mathbf{x}, u)$. This latter expression can be used to simulate the dynamics of our system, through numeric integration. To do that, we have used the following hard-coded function in python,

```python
def ODEeuler(f, xo, *args, to = 0, tf = 10, n = 1000, inp = 'Free'):

        """
        Euler methods for solving non-linear
        systems of ODEs
        """

        step = (tf-to)/n
        _, p = xo.shape

        T = np.zeros((n,));  T[0] = to
        x = np.zeros((n,p)); x[0,:] = xo
        t = np.linspace(to, tf, n)

        if inp == 'Free':
                F = np.zeros((n,))
        elif inp == 'Impulse':
                F = np.zeros((n,))
                F[0] = 1
        elif inp == 'Step':
                F = np.ones((n,))
        elif inp == 'Sinusoid':
                F = args[0]*np.sin(args[1]*t)

        for i in range(n-1):
                xm        = x[i,:] + f(x[i,:], F[i]) * (step / 2)
                x[i+1,:] = x[i,:] + f(xm, F[i]) * step

        return (T,x)
```

Listing 1: Numeric Integration Routine

As a matter of illustration, we have displayed the free-response to our system being at initial state $\mathbf{x}_0 = (0, 0, 0, 0.1)$, that is, it is at the upward position, but moving to in the anti-clockwise direction. Our results are displayed in Figure 3

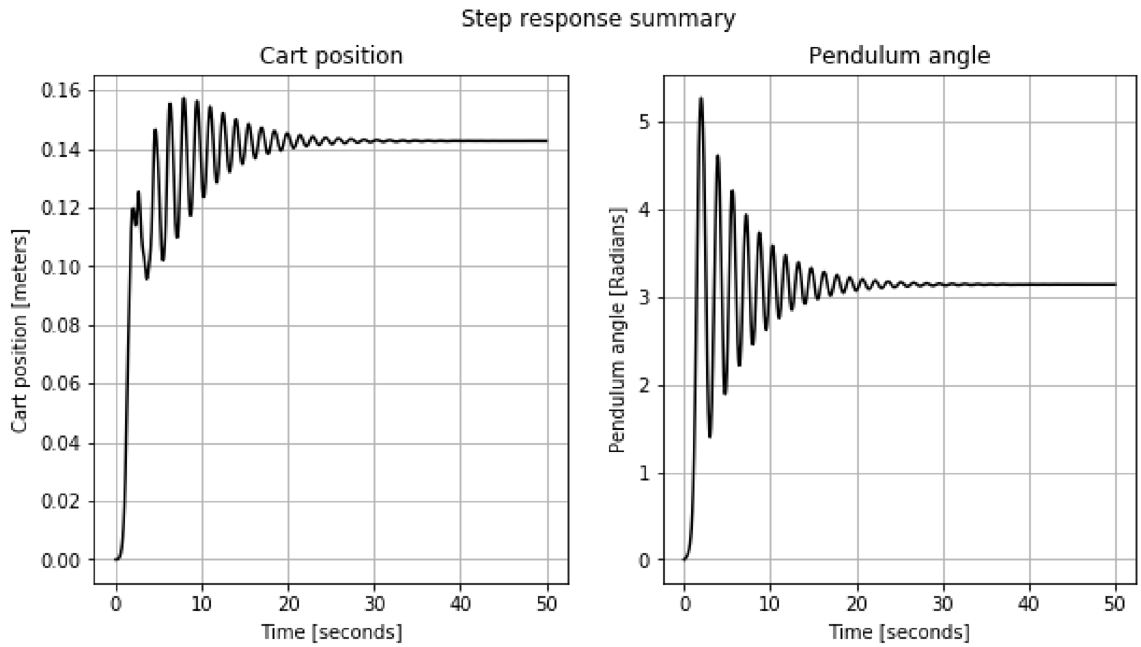

Figure 3: Nonlinear simulation for initial conditions

As we have mentioned, the responses should be oscillatory, since there are periodic functions in $\boldsymbol{\phi}$. Moreover, something interesting happens, as the system evolves, the pendulum falls from the initial position, $\theta = 0$, to the downward position, $\theta = \pi$. This indeed agrees with our experience, since any disturbance in the dynamic equilibrium of an well adjusted inverted pendulum can cause it to fall.

2.2 Linearization

Although useful, the non-linear model described by $\boldsymbol{\phi}$ have the necessity of a numeric integration procedure. A much simpler representation would be if each coordinate of our state vector derivative $\dot{\mathbf{x}}_i$ depended linearly on the states, and the output, that is,

$$\dot{\mathbf{x}}_i = \sum_{j=1}^{n} \mathbf{A}_{ij}\mathbf{x}_j + \mathbf{B}_i u$$

also, we define the system's output as a linear combination of our states and inputs. In practice, we were measuring and displaying outputs in the form $\mathbf{y}_i = \mathbf{x}_i$, that is, we observe our system's state. However, to keep it general enough,

$$\mathbf{y}_i = \sum_{j=1}^{n} \mathbf{C}_{ij}\mathbf{x}_j + \mathbf{D}_i u$$

As consequence of those later equations, we can express them in a compacter form using matrices. Being so, let $\mathbf{A}, \mathbf{C} \in \mathbb{R}^{n\times n}$ and $\mathbf{B}, \mathbf{D} \in \mathbb{R}^{n\times 1}$, our linear state space model can be described as,

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u \\ \mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}u \end{cases} \tag{24}$$

A question remains: how can we describe our linear system $\boldsymbol{\phi}$ in such framework? This question can be answered through the concept of linearization. In simple terms, linearizing the vector field $\boldsymbol{\phi}$ means to take its first order Taylor approximation. This procedure is detailed through Appendix 6. Moreover, let $\mathbf{x}_0, u_0$ be a state vector and input such that, $\boldsymbol{\phi}(t, \mathbf{x}_0, u_0) = 0$, we have:

$$\boldsymbol{\phi}(t, \mathbf{x}_0 + \mathbf{x}_d, u) = \boldsymbol{\phi}(t, \mathbf{x}_0, u_0) + \left.\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}}\right|_{\mathbf{x}_0, u_0} \mathbf{x}_d + \left.\frac{\partial \boldsymbol{\phi}}{\partial u}\right|_{\mathbf{x}_0, u_0} u_d \tag{25}$$

Being $\mathbf{x}_0$ and u_0 stationary points, $\boldsymbol{\phi}(t, \mathbf{x}_0, u_0) = 0$, and $\mathbf{x}_d, u_d$ small perturbations in the state vector, and input. Also, we recall that $\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}}$ is the Jacobian matrix of our field, with entries given by $\left(\left.\frac{\partial \boldsymbol{\phi}_i}{\partial \mathbf{x}_j}\right|_{\mathbf{x}_0, u_0}\right)_{i,j=1}^{n}$. In a similar way, $\frac{\partial \boldsymbol{\phi}}{\partial u}$ is a vector, whose entries are given by $\left(\left.\frac{\partial \boldsymbol{\phi}_j}{\partial u}\right|_{\mathbf{x}_0, u_0}\right)_{j=1}^{n}$. Therefore, it makes sense to associate $A = \frac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}}$ and $B = \frac{\partial \boldsymbol{\phi}}{u}$, hence,

$$\frac{d}{dt}\left(\mathbf{x}_0 + \mathbf{x}_d\right) = \boldsymbol{\phi}(t, \mathbf{x}_0 + \mathbf{x}_d, u_0 + u_d) \tag{26}$$

$$\dot{\mathbf{x}}_d = \left.\frac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}}\right|_{\mathbf{x}_0, u_0} \mathbf{x}_d + \left.\frac{\partial \boldsymbol{\phi}}{\partial u}\right|_{\mathbf{x}_0, u_0} u_d \tag{27}$$

where x_d is a new variable, corresponding to $\mathbf{x}$ translated $\mathbf{x}_0$ unities. Now, we need to identify the points $\mathbf{x}_0, u_0$ such that $\boldsymbol{\phi} = 0$. By inspection, it is easy to see that $\mathbf{x}_0 = (0,0,0,0)$ and $\mathbf{x}_0 = (0,0,\pi,0)$ along with $u_0 = 0$ are stationary points. Those indeed have a special physical meaning: $u_0 = 0$ implies that, initially, no force is being applied to the cart. Also, $\mathbf{x}_0 = (0,0,0,0)$ corresponds to the upward pendulum position, while $\mathbf{x}_0 = (0,0,\pi,0)$, to the downward.

We now need to calculate the Jacobian matrix with respect to $\mathbf{x}$ of $\boldsymbol{\phi}$, and the derivative with respect to u, both applied in the stationary points of our field. After some calculations, we achieve the following results,

For $\mathbf{x}_0 = (0,0,0,0)$:

$$\frac{\partial \boldsymbol{\phi}_1}{\partial \mathbf{x}_1} = 0 \quad \frac{\partial \boldsymbol{\phi}_1}{\partial \mathbf{x}_2} = 1 \quad \frac{\partial \boldsymbol{\phi}_1}{\partial \mathbf{x}_3} = 0 \quad \frac{\partial \boldsymbol{\phi}_1}{\partial \mathbf{x}_4} = 0 \tag{28}$$

$$\frac{\partial \boldsymbol{\phi}_2}{\partial \mathbf{x}_1} = 0 \quad \frac{\partial \boldsymbol{\phi}_2}{\partial \mathbf{x}_2} = \frac{-(m\ell^2 + I)b_x}{\alpha} \quad \frac{\partial \boldsymbol{\phi}_2}{\partial \mathbf{x}_3} = \frac{(m\ell)^2 g}{\alpha} \quad \frac{\partial \boldsymbol{\phi}_2}{\partial \mathbf{x}_4} = \frac{-m\ell b_\theta}{\alpha} \tag{29}$$

$$\frac{\partial \boldsymbol{\phi}_3}{\partial \mathbf{x}_1} = 0 \quad \frac{\partial \boldsymbol{\phi}_3}{\partial \mathbf{x}_2} = 0 \quad \frac{\partial \boldsymbol{\phi}_3}{\partial \mathbf{x}_3} = 0 \quad \frac{\partial \boldsymbol{\phi}_3}{\partial \mathbf{x}_4} = 1 \tag{30}$$

$$\frac{\partial \boldsymbol{\phi}_4}{\partial \mathbf{x}_1} = 0 \quad \frac{\partial \boldsymbol{\phi}_4}{\partial \mathbf{x}_2} = \frac{-m\ell b_x}{\alpha} \quad \frac{\partial \boldsymbol{\phi}_4}{\partial \mathbf{x}_3} = \frac{(M+m)mg\ell}{\alpha} \quad \frac{\partial \boldsymbol{\phi}_4}{\partial \mathbf{x}_4} = \frac{-(M+m)b_\theta}{\alpha} \tag{31}$$

$$\frac{\partial \boldsymbol{\phi}_1}{\partial u} = 0 \quad \frac{\partial \boldsymbol{\phi}_2}{\partial u} = \frac{m\ell^2 + I}{\alpha} \quad \frac{\partial \boldsymbol{\phi}_3}{\partial u} = 0 \quad \frac{\partial \boldsymbol{\phi}_4}{\partial u} = \frac{m\ell}{\alpha} \tag{32}$$

Also, for $\mathbf{x}_0 = (0,0,\pi,0)$:

$$\frac{\partial \boldsymbol{\phi}_1}{\partial \mathbf{x}_1} = 0 \quad \frac{\partial \boldsymbol{\phi}_1}{\partial \mathbf{x}_2} = 1 \quad \frac{\partial \boldsymbol{\phi}_1}{\partial \mathbf{x}_3} = 0 \quad \frac{\partial \boldsymbol{\phi}_1}{\partial \mathbf{x}_4} = 0 \tag{33}$$

$$\frac{\partial \boldsymbol{\phi}_2}{\partial \mathbf{x}_1} = 0 \quad \frac{\partial \boldsymbol{\phi}_2}{\partial \mathbf{x}_2} = \frac{-(m\ell^2 + I)b_x}{\alpha} \quad \frac{\partial \boldsymbol{\phi}_2}{\partial \mathbf{x}_3} = \frac{-(m\ell)^2 g}{\alpha} \quad \frac{\partial \boldsymbol{\phi}_2}{\partial \mathbf{x}_4} = \frac{m\ell b_\theta}{\alpha} \tag{34}$$

$$\frac{\partial \boldsymbol{\phi}_3}{\partial \mathbf{x}_1} = 0 \quad \frac{\partial \boldsymbol{\phi}_3}{\partial \mathbf{x}_2} = 0 \quad \frac{\partial \boldsymbol{\phi}_3}{\partial \mathbf{x}_3} = 0 \quad \frac{\partial \boldsymbol{\phi}_3}{\partial \mathbf{x}_4} = 1 \tag{35}$$

$$\frac{\partial \boldsymbol{\phi}_4}{\partial \mathbf{x}_1} = 0 \quad \frac{\partial \boldsymbol{\phi}_4}{\partial \mathbf{x}_2} = \frac{m\ell b_x}{\alpha} \quad \frac{\partial \boldsymbol{\phi}_4}{\partial \mathbf{x}_3} = \frac{-(M+m)mg\ell}{\alpha} \quad \frac{\partial \boldsymbol{\phi}_4}{\partial \mathbf{x}_4} = \frac{-(M+m)b_\theta}{\alpha} \tag{36}$$

$$\frac{\partial \boldsymbol{\phi}_1}{\partial u} = 0 \quad \frac{\partial \boldsymbol{\phi}_2}{\partial u} = \frac{m\ell^2 + I}{\alpha} \quad \frac{\partial \boldsymbol{\phi}_3}{\partial u} = 0 \quad \frac{\partial \boldsymbol{\phi}_4}{\partial u} = \frac{-m\ell}{\alpha} \tag{37}$$

Equations 28-37 define two linear systems, one for each stationary point. Moreover, as we shall study in the stability section, $\mathbf{x}_0 = (0,0,0,0)$ is an unstable point, while $\mathbf{x}_0 = (0,0,\pi,0)$ is a stable one. This is evidenced in Figure 4, where we plot the simulation for the two responses,

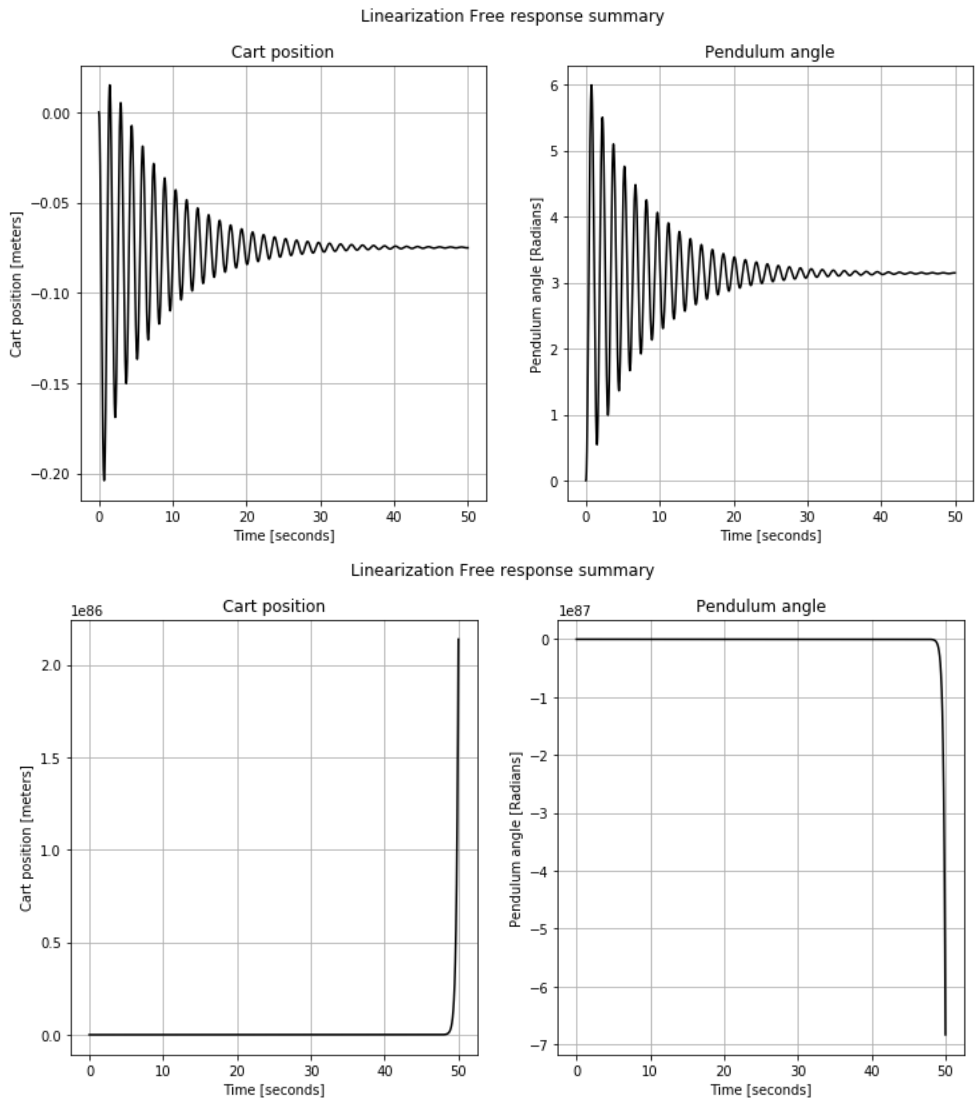

Figure 4: The response plot for the linearization results. Notice how, for $\mathbf{x}_0 = (0,0,0,0)$, the response is unstable.

As consequence, for a matter of comparison, we shall use the linearization around $\mathbf{x}_0 = (0, 0, \pi, 0)$. In this notation, since we subtract π from $\mathbf{x}$, the upward position changes to π, while the downward to 0[7].

2.3 Transfer Function derivation

To express the system in terms of transfer functions, we perform a similar procedure we have done in the linearization section, but, at this time, in Equations 13 and 14. The procedure come from noticing that the only coordinate that change between the stationary points is $\theta = 0$ and $\theta = \pi$. Effectively,

Function	$\theta = 0$	$\theta = \pi$
$sin(\theta)$	θ	$-\theta$
$cos(\theta)$	1	-1
$\dot{\theta}\theta$	0	0

Table 2: Approximation table

Working first for $\theta = 0$,

$$(M+m)\ddot{x} - m\ell\ddot{\theta} = F - b_x\dot{x} \tag{38}$$

$$-m\ell\ddot{x} + (m\ell^2 + I)\ddot{\theta} = mg\ell\theta - b_\theta\dot{\theta} \tag{39}$$

now, in the frequency domain,

$$\Big((M+m)s^2 + b_x s\Big)X - m\ell s^2\Theta = F \tag{40}$$

$$-m\ell s^2 X + \Big((m\ell^2 + I)s^2 + b_\theta s - mg\ell\Big)\Theta = 0 \tag{41}$$

Solving for X and Θ results in calculating the following determinants,

[7]To not cause any confusion, the plot displayed in Figure 4 were converted to the previous notation.

$$
\begin{aligned}
D(s) &= \begin{vmatrix} (M+m)s^2 + b_x s & -m\ell s^2 \\ -m\ell s^2 & (m\ell^2 + I)s^2 + b_\theta s - mg\ell \end{vmatrix} \\
&= \alpha s^4 + \Big((M+m)b_\theta + (m\ell^2 + I)b_x \Big) s^3 + \Big(b_\theta b_x - (M+m)mg\ell \Big) s^2 - mg\ell b_x s \\
N_x(s) &= \begin{vmatrix} F & -m\ell s^2 \\ 0 & (m\ell^2 + I)s^2 + b_\theta s - mg\ell \end{vmatrix} \\
&= F\Big((m\ell^2 + I)s^2 + b_\theta s - mg\ell \Big) \\
N_\theta(s) &= \begin{vmatrix} (M+m)s^2 + b_x s & F \\ -m\ell s^2 & 0 \end{vmatrix} \\
&= F\Big(m\ell s^2 \Big)
\end{aligned}
$$

Then, we can write, for $\mathbf{x}_0 = (0, 0, 0, 0)$

$$
T_x = \frac{X(s)}{U(s)} \quad = \frac{(m\ell^2 + I)s^2 + b_\theta s - mg\ell}{\alpha s^4 + \Big((M+m)b_\theta + (m\ell^2 + I)b_x \Big) s^3 + \Big(b_\theta b_x - (M+m)mg\ell \Big) s^2 - mg\ell b_x s} \tag{42}
$$

$$
T_\theta = \frac{\Theta(s)}{U(s)} \quad = \frac{m\ell s^2}{\alpha s^4 + \Big((M+m)b_\theta + (m\ell^2 + I)b_x \Big) s^3 + \Big(b_\theta b_x - (M+m)mg\ell \Big) s^2 - mg\ell b_x s} \tag{43}
$$

Now, for $\mathbf{x}_0 = (0, 0, \pi, 0)$, we can repeat the above procedure to find,

$$
T_x = \frac{X(s)}{U(s)} \quad = \frac{(m\ell^2 + I)s^2 + b_\theta s + mg\ell}{\alpha s^4 + \Big((M+m)b_\theta + (m\ell^2 + I)b_x \Big) s^3 + \Big(b_\theta b_x + (M+m)mg\ell \Big) s^2 + mg\ell b_x s} \tag{44}
$$

$$
T_\theta = \frac{\Theta(s)}{U(s)} \quad = \frac{-m\ell s^2}{\alpha s^4 + \Big((M+m)b_\theta + (m\ell^2 + I)b_x \Big) s^3 + \Big(b_\theta b_x + (M+m)mg\ell \Big) s^2 + mg\ell b_x s} \tag{45}
$$

These equations are the counterpart in frequency of the linear state-space representation. It is relatively easy to see that $D(s) = det(s\mathbf{I} - \mathbf{A})$, implying that the poles of both transfer functions are the eigenvalues of $\mathbf{A}$, allowing us to make the same inference we will do using the matrix spectra, for the denominator's polynomial of T_x and T_θ.

Equations T_x along with T_θ determines an I/O model (in frequency) for the inverted pendulum system. Indeed, we can use them to retrieve differential equations for them:

$$\begin{cases} \alpha \dfrac{d^4 x}{dt^4} + \Big((M+m)b_\theta \quad + (m\ell^2 + I)b_x \Big) \dfrac{d^3 x}{dt^3} + \\ \qquad \Big(b_\theta b_x + (M+m)mg\ell \Big) \dfrac{d^2 x}{dt^2} + mg\ell b_x \dfrac{dx}{dt} = (m\ell^2 + I)\dfrac{d^2 u}{dt^2} + b_\theta \dfrac{du}{dt} + mg\ell u \\ \alpha \dfrac{d^3 \theta}{dt^3} + \Big((M+m)b_\theta \quad + (m\ell^2 + I)b_x \Big) \dfrac{d^3 \theta}{dt^2} + \\ \qquad \Big(b_\theta b_x + (M+m)mg\ell \Big) \dfrac{d\theta}{dt} + mg\ell b_x \theta = -m\ell \dfrac{du}{dt} \end{cases}$$

If those were isolated SISO systems, the role that the applied force has in the control would be clear. However, our input affects both the angle, and the position simultaneously, that is why we have a MIMO system. Also important to mention, the other state variables - $\dot{x}$ and $\dot{\theta}$ - transfer functions can be obtained by multiplying T_x and T_θ by s. Further comments in which state variables we shall observe and manipulate to stabilize the pendulum around $(0,0,0,0)$ shall be made in the second homework.

3 Stability analysis

In this section, we discuss further results about the stability of our system. Up to now, we have claimed that $\mathbf{x}_0 = (x, 0, 0, 0)$ is a unstable stationary point, and $\mathbf{x}_0 = (x, 0, \pi, 0)$. First, we need to formalize the concepts of stability,

Definition 1. *A stationary point $\mathbf{x}^* = (x_1^*, x_2^*, x_3^*, x_4^*) \in \mathbb{R}^4$ is said to be stable if, and only if, for each $\epsilon > 0$ there exists a $\delta > 0$ such that for every solution $\mathbf{x}(t) = \boldsymbol{\phi}(t, \mathbf{x})$,*

$$|\boldsymbol{\phi}(0, \mathbf{x}) - \mathbf{x}^*| < \delta \overset{\forall t \geq 0}{\rightarrow} |\boldsymbol{\phi}(t, \mathbf{x}) - \mathbf{x}^*| < \epsilon$$

Following the common-sense, stability says that trajectories which initiates δ-close to stationary points $\mathbf{x}^*$ stay ϵ-close, as time goes by. We say a point is unstable if it is not stable. Another kind of stability is asymptotically stability, which may be defined as,

Definition 2. *A stationary point* $\mathbf{x}^* = (x_1^*, x_2^*, x_3^*, x_4^*) \in \mathbb{R}^4$ *is said to be asymptotically stable if, and only if, there exists a* δ *such that each trajectory* $\boldsymbol{\phi}(t, \mathbf{x})$,

$$|\boldsymbol{\phi}(0, \mathbf{x}) - \mathbf{x}^*| < \delta \to \lim_{t \to \infty} \boldsymbol{\phi}(t, \mathbf{x}) = \mathbf{x}^*$$

This has the same sense as before, but with a little bit more of flexibility, in the sense that solutions are only guaranteed to approach $\mathbf{x}^*$ in the infinity.

In addition to those concepts, we shall define the sinks of a given field $\boldsymbol{\phi}$ as the points $\mathbf{x}^*$ whose Jacobian matrix of $\boldsymbol{\phi}$ is negative definite, that is,

Definition 3. *A stationary point* $\mathbf{x}^*$ *of a field* $\boldsymbol{\phi}$ *is a sink, if and only if all eigenvalues of* $\dfrac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}}$ *are negative.*

An important theorem gives us the connection between sinks and asymptotically stable points, which is the Lyapunov-Perron theorem, whose enunciate may be found bellow. With such theorem, we substitute the stability analysis of the field $\mathbf{f}$, by the stability analysis of the Jacobian, which is easier.

Theorem 1. *Let* $\mathbf{x}^* \in E \subset \mathbb{R}^n$ *be a stationary point of the field* $\mathbf{f} : E \to \mathbb{R}^n$. *If* $\mathbf{x}^*$ *is a sink of* $\boldsymbol{\phi}$, *then it is an asymptotically stable point of* $\boldsymbol{\phi}$.

A more detailed discussion about the proof, and stability analysis of linearized models can be found in the appendix. For now, it is sufficient to say that the time response of the system relies on the spectra of matrix $A = \dfrac{\partial \boldsymbol{\phi}}{\partial \mathbf{x}}$. Indeed, let λ_i be the eigenvalue associated with eigenvector $\boldsymbol{\xi}_i$, then, being $\mathbf{x} = \sum_i \boldsymbol{\xi}_i e^{\lambda_i t}$,

$$\mathbf{A}\mathbf{x} = \sum_i e^{\lambda_i t} \boldsymbol{A}\boldsymbol{\xi}_i \tag{46}$$

$$= \sum_i \lambda_i e^{\lambda_i t} \xi_i \tag{47}$$

$$= \dot{\mathbf{x}} \tag{48}$$

As we wanted. Thus, the stability will be determined by the real part of each eigenvalue, λ_i. The proof of Theorem 1, indeed, only establishes the intuitive fact that, in a neighborhood of each domain's point, the non-linear solution resembles the linear one.

Therefore, we need to look into the spectra of A. A quick evaluation using the benchmark values, gives us the eigenvalues displayed in Table 3 and Figure 5

Eigenvalue	A_0	A_π
λ_1	0	0
λ_2	-4.423	-1.0819
λ_3	-0.9469	-0.126 + 4.218i
λ_4	4.0349	-0.126 - 4.218i

Table 3: System's Eigenvalues

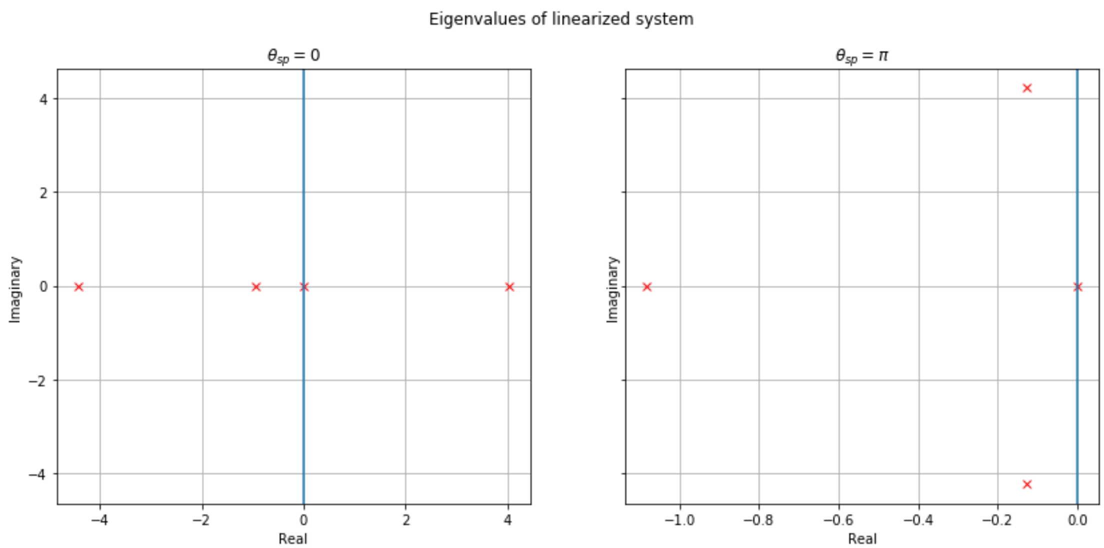

Figure 5: Eigenvalues Positions

Those results indeed confirm our conclusion that $\mathbf{x}_0 = (x, 0, 0, 0)$ is a source, or an unstable stability point, while $\mathbf{x}_0 = (x, 0, \pi, 0)$ is a stable one, and therefore a sink. The effort in the next chapters will be to analyze and build controllers to stabilize the system near $\mathbf{x}_0 = (x, 0, 0, 0)$, so we can have a pendulum with stable upward position.

To point out, we make a more formal discussion of those concepts in Appendix B. These formalizations are based on [3] and [4].

4 Simulation

In this section, we shall simulate our system due four kinds of inputs: 1) $u(t) = 0$, that is, we want the free response due to some initial conditions, 2) $u(t) = \delta(t)$, that is, we want the impulse response of our system, 3) $u(t) = \delta_{-1}$, the step response and 4) $u(t) = Asin(\omega t)$, the response to an sinusoidal signal with amplitude A, and frequency ω.

To access the performance of our linearization, we define our approximation error as being $e(t) = y(t) - \tilde{y}(t)$, where y is any of our outputs. Being so, we use the following integrals to measure how well the linearization approximated the actual response:

$$
\begin{aligned}
IAE &= \int_0^{+\infty} |e(t)| dt \\
ISE &= \int_0^{+\infty} e(t)^2 dt \\
ITAE &= \int_0^{+\infty} t|e(t)| dt
\end{aligned}
$$

Since we are dealing with computational simulations, the error function is rather discrete, in the sense that $e(t)$ becomes a vector in any computational language, with entries $e[k]$. Being so, the later equations become:

$$
\begin{aligned}
IAE &= \sum_{k=0}^{n} |e[k]| \\
ISE &= \sum_{k=0}^{n} e[k]^2 \\
ITAE &= \sum_{k=0}^{n} T[k]|e[k]| \\
RMSE &= \frac{1}{n}\sqrt{ISE}
\end{aligned}
$$

4.1 Free response simulation

We already presented both simulations throughout the paper. This time, we shall consider Figure 6, which displays both kinds of responses,

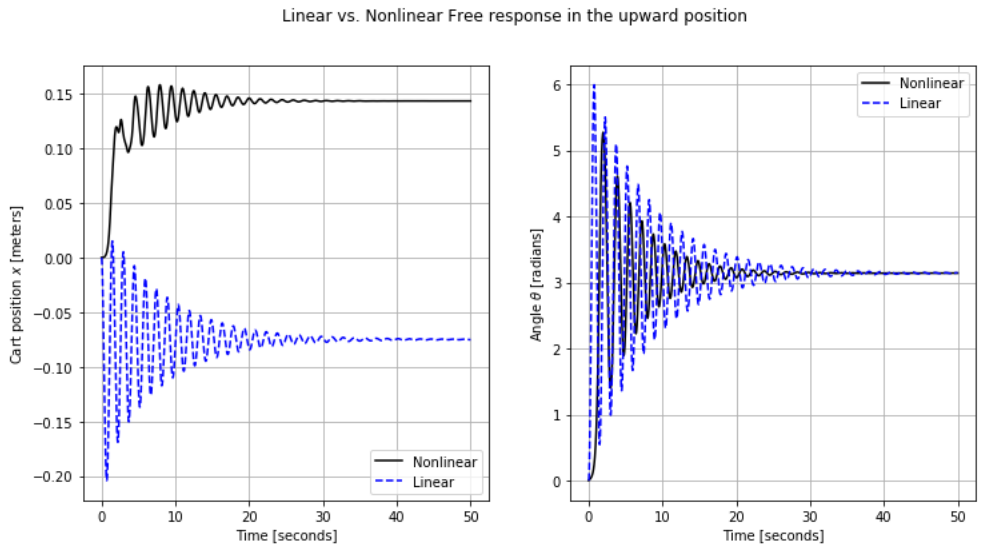

Figure 6: Comparison between the nonlinear and the linear response

In that case, it seems that the angle linearization approximates better the nonlinear response than the position linearization. However, using our assessment criterion, here displayed in Table 4, the situation is exactly the opposite,

Criteria	x	θ
IAE	2121.74	31227.18
ISE	459.49	103066.68
ITAE	54392.61	785445.18
RMSE	2.14×10^{-3}	3.21×10^{-2}

Table 4: Linearization performance

A closer look in figure 6 point out that, despite the resemblance of the linear and nonlinear responses, the nonlinear seems to have an delay, making the linearization simplistic in displaying the behavior of our system. For the cart's position, the linearization is obviously wrong, since its steady state value is different, as well as the response itself.

4.2 Impulse response simulation

The impulse response is quite similar to the free response, due to its nature. To understand such similarity, we examine the impulse function: for $t = 0$, $\delta(t)$ is non-zero, therefore the ODE behaves like an non-homogeneous term. For $t > 0$, $\delta(t) = 0$, and therefore there is no input applied to the ODE. Hence, the ODE behaves like the homogeneous case, but rather with initial conditions imposed by the impulse input at $t = 0$. This is shown in Figure 7,

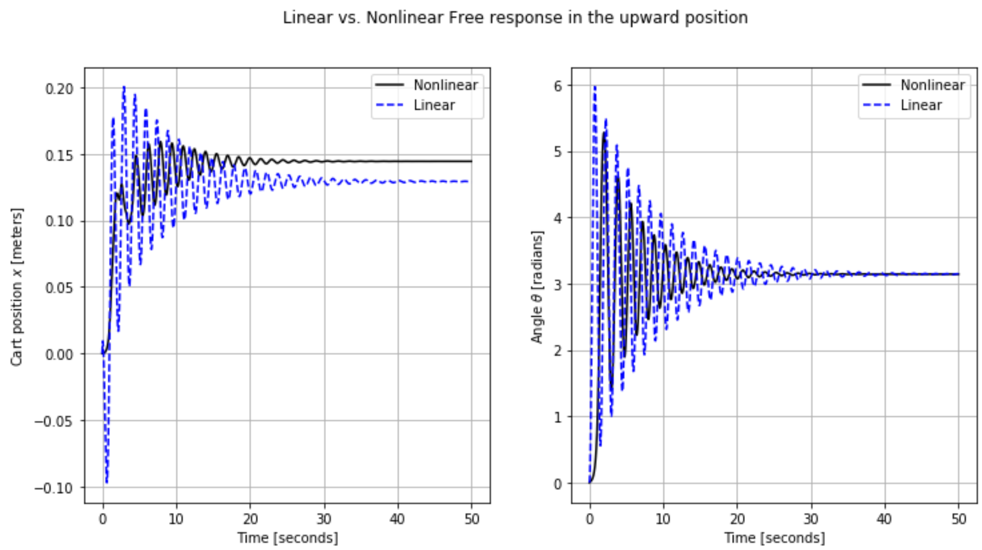

Figure 7: Comparison between the nonlinear and the linear response

As we can see, the non-linear response does not exhibit much change, except for the steady state value of x. Also, the linear behavior seems to gather more of the position behavior, but still, there is some offset. For the angle response, the same delay pattern appears. Our criterion for the impulse response are,

Criteria	x	θ
IAE	216.87	31221.00
ISE	7.69	102997.68
ITAE	4138.64	785437.07
RMSE	2.77×10^{-4}	3.21×10^{-2}

Table 5: Linearization performance

4.3 Step Response

The step response of our system has a more complicated scenario. Following our common sense, we would think that, by constantly pushing the cart to the right, its movement is unbounded in that direction. Also, due to the instability in the upward position, the pendulum will fall to the downward position. Therefore, the behavior of both variables is undesirable: the cart will go off the track, and the pendulum will go to the downward position. This is evidenced in Figure 8,

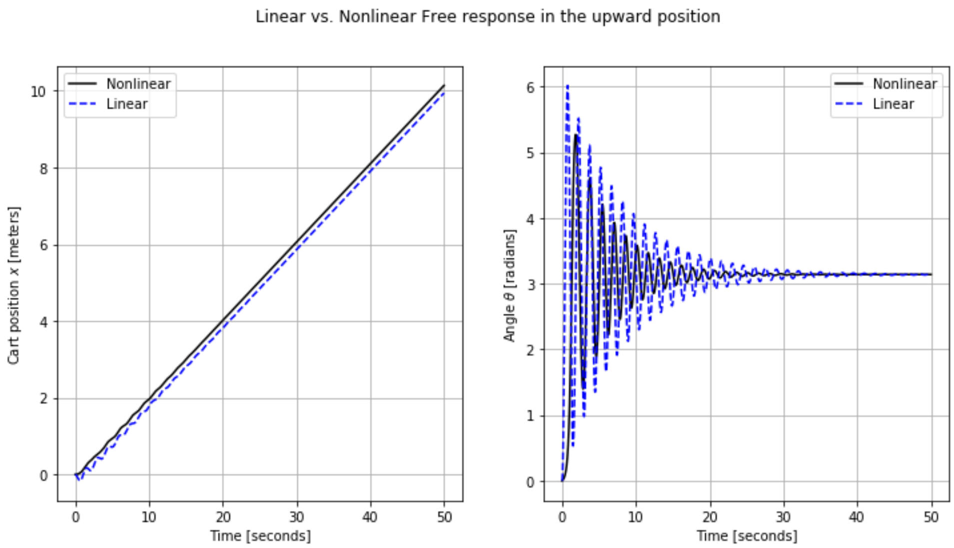

Figure 8: Comparison between the nonlinear and the linear response

That time, both linear responses appears to follow the behavior of our non-linear one. This fact, however, does no reflect itself in our assessment criterion, due to the magnitude of our output, that time,

Criteria	x	θ
IAE	1913.63	31296.08
ISE	373.37	103609.80
ITAE	48914.53	785475.00
RMSE	1.93×10^{-3}	3.22×10^{-2}

Table 6: Linearization performance

Now, we want to analyze the time-response of our system, approximating it as a second-order process. Since the position response is unbounded, it does not make sense to do such analysis to it. Now, we need to determine ξ and ω_n in order to gather: 1) The rise time of $\theta(t)$, 2) The overshoot, 3) The peak time and 4) The settling time. Being so,

Definition 4. *The rise time, t_r, of a system's response $y(t)$ is the time that y takes to go from 0.1 to 0.9 of its steady state value. The settling time, t_s, is the time required for $y(t)$ to be around 0.1* *of its steady state value. The overshoot, OS, is the ratio of how much the peak of $y(t)$ exceeds the steady state value. The peak time, t_p is the time which y have its maximum value.*

For a second-order system, we have:

$$\omega_n \approx \frac{1.8}{t_r}$$
$$\xi = \frac{-ln(OS)}{\sqrt{ln(OS)^2 + \pi^2}}$$

Now, we can evaluate t_r, OS and t_p empirically (from the response we have acquired), which are sufficient to calculate ξ and ω_n. Provided these last two values, we can calculate the settling time as $\frac{4}{\omega_n \xi}$. The results for both linear and nonlinear responses are displayed in Table 7,

Measure	Linear	Nonlinear
t_r (seconds)	0.25	0.60
t_s (seconds)	22.88	12.51
M_p	91.60%	67.83%
t_p (seconds)	0.74	2.06

Table 7: Time domain response information summary

5 Appendix A: Calculus of Variations

5.1 Introduction

We begin our discussion with the intent to define the stationary points of a functional. This, as we shall seen, gives the basis for the so-called Euler-Lagrange equations, which are used to model our physical problems.

Definition 5. *Consider a vector space $\mathcal{V}$, of arbitrary dimension. An function is any function $J : \mathcal{V} \to \mathbb{R}$.*

Henceforth, we shall pay attention to a special kind of functional, defined as:

$$J(y) = \int_{x_0}^{x_1} f(x, y, \dot{y}) dx \tag{49}$$

Specially, the domain in which J takes its values is $\mathcal{C}^2[x_0, x_1]$[8]. Although we have the definition of our functional, we still do not have the concept of maxima and minima for it:

Definition 6. *Let $(\mathcal{V}, || \cdot ||)$ be a metric space, and let $\mathcal{S} \subset \mathcal{V}$. We say that J attains a local maximum on $\mathcal{S}$, at $\hat{y} \in S$ if there exists $\epsilon > 0$ such that $J(\hat{y}) - J(y) \leq 0$ for all $\hat{y} \in S$, such that $||\hat{y} - y|| < \epsilon$. The definition for local minimum is analogous.*

Notice that, for functions $\hat{y} \in \mathcal{S}$, if y is in the neighborhood of $\hat{y}$, we represent $\hat{y}$ as a perturbation[9] of y:

$$\hat{y} = y + \epsilon\eta$$

This notion of perturbation indeed defines a topology in $\mathcal{C}^2$. We shall deal with special cases of perturbations, that is, we restrict η to suffice $\eta(x_0) = \eta(x_1) = 0$. Under this kind of assumption, our problem is called **fixed endpoint variation problem**. Graphically, we display this relation in Figure 9

Hence, we shall work within two sets of functions:

$$\mathcal{S} = \{y \in \mathcal{C}^2[x_0, x_1] : y(x_0) = y_0, y(x_1) = y_1\} \tag{50}$$
$$\mathcal{H} = \{\eta \in \mathcal{C}^2[x_0, x_1] : \eta(x_0) = \eta(x_1) = 0\} \tag{51}$$

[8]The space of twice differentiable functions

[9]Indeed, we could also talk about a intrinsic and uncontrollable noise

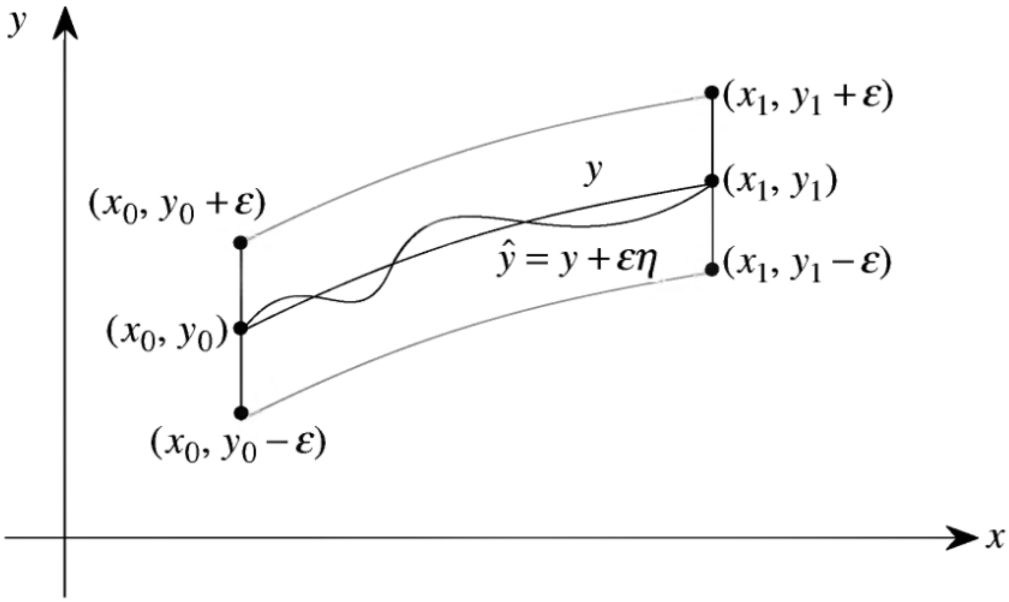

Figure 9: Region within $\mathbb{R}^2$ delimited by the perturbation in y

5.2 First Variation

Let us consider $f(x, y, \hat{y}')$ for small perturbations in $\hat{y}$:

$$
\begin{aligned}
f(x, \hat{y}, \hat{y}') &= f(x, y + \epsilon\eta, y' + \epsilon\eta') && (52)\\
&= f(x, y, y') + \epsilon(\eta\frac{\partial f}{\partial y} + \frac{\partial f}{\partial y'}) + \mathcal{O}(\epsilon^2) && (53)
\end{aligned}
$$

In which, from (4) to (5), we have used Taylor's approximation. We want to investigate $\Delta J(y) = J(\hat{y}) - J(y)$. This quantity can be expressed as:

$$
\begin{aligned}
\Delta J(y) &= \int_{x_0}^{x_1} f(x, \hat{y}, \hat{y}')dx - \int_{x_0}^{x_1} f(x, y, y')dx && (54)\\
&= \int_{x_0}^{x_1} \{f(x, y, y') + \epsilon(\eta\frac{\partial f}{\partial y} + \frac{\partial f}{\partial y'}) + \mathcal{O}(\epsilon^2) - f(x, y, y')\}dx && (55)\\
&= \epsilon \int_{x_0}^{x_1} (\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'})dx + \mathcal{O}(\epsilon^2) && (56)\\
&= \epsilon\delta J(\eta, y) + \mathcal{O}(\epsilon^2) && (57)
\end{aligned}
$$

Where $\delta J(\eta, y) = \int_{x_0}^{x_1}(\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'})dx$ is called the first variation of J. Now, since the boundary values of η are zero, $\eta \in \mathcal{H} \to -\eta \in \mathcal{H}$, and $\delta J(\eta, y) = -\delta J(-\eta, y)$. For small values of ϵ, the sign of $\Delta J(y)$ is determined by $\delta J(\eta, y)$, thus, if it is supposed to have a local maximum in $\mathcal{S}$, $J(\hat{y}) - J(y)$ does not change sign for any $\hat{y} \in S, ||\hat{y} - y|| < \epsilon$. Thus:

$$\delta J(\eta, y) = \int_{x_0}^{x_1}(\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'})dx = 0 \tag{58}$$

For all $\eta \in \mathcal{H}$. We could use similar arguments for the case in which J attains a local minima in $\mathcal{S}$. Equation 58, indeed, stablish the infinite-dimensional case for stationary points of J. In order to make this expression more tractable, we use integration by parts:

$$\int_{x_0}^{x_1}(\eta\frac{\partial f}{\partial y} + \eta'\frac{\partial f}{\partial y'})dx = \eta\frac{\partial f}{\partial y}\bigg|_{x_0}^{x_1} - \int_{x_0}^{x_1}\eta\frac{d}{dx}\frac{\partial f}{\partial y'}dx \tag{59}$$

$$= -\int_{x_0}^{x_1}\eta\frac{d}{dx}\frac{\partial f}{\partial y'}dx \tag{60}$$

Where, from Equation 59 to Equation 60 we have used the fact that η is zero at boundary values. With this result, the first variation takes the form:

$$\int_{x_0}^{x_1}\eta\left\{\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right\}dx \tag{61}$$

Now, defining $E(x) = \frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}$, we can understand Equation 61 as an inner product within a Hilbert Space[10], that is:

$$\langle \eta, E \rangle = \int_{x_0}^{x_1}\eta(x)E(x)dx \tag{62}$$

Indeed, $\langle \eta, E \rangle = 0$ stablish the orthogonality between η and E. We are now interested in the proposition that, is those two functions are indeed orthogonal, and η non-zero inside an open interval of $\mathbb{R}$, then $E(x) = 0$ for all x.

[10]Indeed, $\mathcal{C}^2[x_0, x_1]$ is a complete space with a inner product, so it is, by definition, a Hilber Space

5.3 The Euler-Lagrange equations

Proposition 1. *Suppose that $\langle \eta, g \rangle = 0$, for all $\eta \in \mathcal{H}$. If $g : [x_0, x_1] \to \mathbb{R}$ is continuous, then $g = 0$ on the interval $[x_0, x_1]$.*

Proof. Suppose that $g \neq 0$ for some $c \in [x_0, x_1]$. Without loss of generality, assume $g(c) > 0$, and by continuity $c \in (x_0, x_1)$. Then, there exists a subinterval $(\alpha, \beta) \subset (x_0, x_1)$, such that $c \in (\alpha, \beta)$, this implies that $g(x) > 0$, in (α, β). Notice that there exists a function[11] $\nu \in \mathcal{C}^2[x_0, x_1]$ such that $\nu > 0, \forall x \in (\alpha, \beta)$, and $\nu = 0, \forall x \in [x_0, x_1] - (\alpha, \beta)$. Therefore, since $\nu \in \mathcal{H}$:

$$\langle \nu, g \rangle = \int_{x_0}^{x_1} \nu(x) g(x) dx \tag{63}$$

$$= \int_{\alpha}^{\beta} \nu(x) g(x) dx \tag{64}$$

$$= 0 \tag{65}$$

Which contradicts the fact that $\langle \eta, g \rangle = 0, \forall \eta \in \mathcal{H}$. Thus, $g(x) = 0, \forall x \in [x_0, x_1]$. ☐

This proposition stablish the conditions in which we can conclude that $E(x) = 0, \forall x \in [x_0, x_1]$. This implies in the following corollary:

Corollary 1. *Let $J : \mathcal{C}^2 \to \mathbb{R}$ be a functional of the form,*

$$J(y) = \int_{x_0, x_1} f(x, y, y') dx \tag{66}$$

where f has continuous partial derivatives of second order with respect to x, y and y'. Let,

$$\mathcal{S} = \{y \in \mathcal{C}^2[x_0, x_1] : y(x_0) = y_0, y(x_1) = y_1\}$$

if $y \in \mathcal{S}$ is a extremum point of J, then:

$$\frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} = 0 \tag{67}$$

This latter equation is called **The Euler-Lagrange Equation**. It is the infinite-dimensional analogue for the conditions $\nabla f = 0$ and $\frac{d}{dx} f = 0$.

For now on, we shall be interested in describing physical systems through this theory[12]. Specifically, we shall define a quantity,

$$L = T - V \tag{68}$$

[11] We can choose $\nu(x) = (x - \alpha)^3 (\beta - x)^3$, for instance

[12] This is called Lagrangian Mechanics

Where T is the total kinetic energy of the system, and V, the total potential energy of the system. Depending on the model we use for whichever system we want to control, we shall employ different technique in order to adequate L to be f, in Euler-Lagrange equations.

6 Appendix B: Nonlinear ordinary differential equations

The concepts presented here are a synthesis from references [3] and [4]. For a further explanation of those concepts, the readers are invited to look more careful inside those books.

6.1 Introductory Notions

In order to discuss the stability of nonlinear systems, we need to establish some notions. We start with the definitions about nonlinear state-space representation. We begin with the definition of a vector field,

Definition 7. *Let $E \subset \mathbb{R}^n$ and $F \subset \mathbb{R}^m$ be euclidean vectorial spaces of dimensions n and m, respectively. We call* ***vectorial field*** *any function $f : E \to F$.*

In practice, vectorial fields assigns, to each point in $E \subset \mathbb{R}^n$, a vector $\mathbf{f} \subset \mathbb{R}^m$. Being so, they can be represented by m functions of n variables, that is,

$$\mathbf{f}(\mathbf{x}) = \left[f_1(\mathbf{x}), \cdots, f_m(\mathbf{x})\right]^T$$

Such definition provide us a mathematical tool to analyze non-linear differential equations. For instance, let $x_1(t), \cdots, x_n(t)$ be n functions of time, each of them satisfying its non-linear differential equation, $f_j(t, x_1, \cdots, x_n)$. We can create a system, in the following way,

$$\begin{cases} \dot{x}_1 & = f_1(t, x_1, \cdots, x_n) \\ \dot{x}_2 & = f_2(t, x_1, \cdots, x_n) \\ & \vdots \\ \dot{x}_n & = f_n(t, x_1, \cdots, x_n) \end{cases}$$

This can be represented, in a more compact form, by denoting,

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

indeed, this was the approach we have done through the linearization section, and also through the system's representation as a vector field, from $\mathbb{R}^4$ to $\mathbb{R}^4$. Indeed, as time goes by, the vectors $\dot{\mathbf{x}} \in \mathbb{R}^n$ attached to the point $\mathbf{x}$ through the time index t changes. Therefore, it makes sense to give the following definition,

Definition 8. *Consider the non-linear differential equations,*

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

*Any solution is a curve $\mathbf{x}$ in $\mathbb{R}^n$ with $\mathbf{x}(0) = \mathbf{x}_0$, and for each time t, $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t))$ is called the trajectory of f, by $\mathbf{x}_0$. Also, the sets defined as $\{\mathbf{x}(t) : t \in \mathbb{R}^n\} \subset \mathbb{R}^n$ are called **the orbits** of $\mathbf{f}$ by $\mathbf{x}_0$.*

Also important to mention, we shall adopt the notation $\phi_t(\mathbf{x}) = \mathbf{x}(t)$, which is known as **the flux** of f. Our last definition gives us precisely what are the singularities of the vector field,

Definition 9. *Given a non-linear differential equation, $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, a singularity of f, or stationary point $\mathbf{x}_0$, is a point such that $\mathbf{f}(\mathbf{x}_0) = 0$.*

The notion of stationarity is given by the fact that, if $\mathbf{f}$ is zero in $\mathbf{x}_0$, then the trajectory $\mathbf{x}(t)$, passing through $\mathbf{x}_0$ is constant after reaching $\mathbf{x}_0$. Now, we turn to the stability in dynamic systems,

Definition 10. *Given a non-linear differential equation, $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ and a stationary point $\mathbf{x}_0$. We say $\mathbf{x}_0$ is stable if, for every $\epsilon > 0$ there exists a δ such that*

$$|\mathbf{x} - \mathbf{x}_0| < \delta \to |\phi_t(\mathbf{x}) - \mathbf{x}_0| < \epsilon, \forall t.$$

Moreover, $\mathbf{x}_0$ is asymptotically stable if, for every $\epsilon > 0$,

$$|\phi_t(\mathbf{x}) - \mathbf{x}_0| < \epsilon \to \lim_{t\to\infty} \phi_t(\mathbf{x}) = \mathbf{x}_0$$

The main difference from stability to asymptotically stability is that in the first, we are guaranteed to never leave an ϵ-neighborhood of $\mathbf{x}_0$, provided we are δ-close to it. In the second, provided that we have started ϵ close, we may leave this neighborhood, but we are guaranteed that, somewhere in the future, we will return.

The stationary points are indeed the more valuable points in any differential equation, since they provide information about the behavior of our system (namely, from where we start, to where we go). As we shall see,

- Unstable stationary points are equivalent to sources, that is, the dynamics of the system has origin there,
- Since the last ones are unstable, they are repelled from there, in the direction of stable points - the sinks -. The stable stationary points are, thus, equivalent to sinks.

Now, we shall work with linearizations. From elementary calculus, we know the concept of Taylor series expansion, that is, given an analytic function[13], it can be expanded as a power series of x, that is,

$$f(x) = \sum_{n=0}^{\infty} \frac{(x-x_0)^n}{n!} \frac{d^n}{dt^n} f(x_0)$$

This series can be truncated, in order to generate the best n-th order polynomial approximation of f. For example, the best first-order approximation of f is,

$$f(x) = f(x_0) + \left.\frac{df}{dt}\right|_{x=x_0} (x - x_0) + \mathcal{O}((x-x_0)^2)$$

If x_0 is a stationary point of f, then $f(x_0) = 0$, and the approximation is linear, that is, provided that $\left.\frac{df}{dt}\right|_{x=x_0} = m$, $f(x) = m(x - x_0)$. This, although, only guarantees linearizations for the univariate case, which needs to be generalized to vectorial fields, $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$. Before doing so, let us suppose proved the result for surfaces[14], that is, given a function $f : \mathbb{R}^n \to \mathbb{R}$,

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f|^T_{\mathbf{x}=\mathbf{x}_0} (\mathbf{x} - \mathbf{x}_0)$$

where $\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}^T$ is the gradient vector of f. To begin with, we notice $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ is a vector composed by n surfaces, that is, $\mathbf{f} = (f_j)_{j=1}^n$, each with its gradient, ∇f_j. We notice that,

$$\begin{aligned} \dot{x}_1 &= f_1(\mathbf{x}) \\ &\vdots \\ \dot{x}_n &= f_n(\mathbf{x}) \end{aligned}$$

[13] Infinitely many differentiable functions

[14] Most calculus textbooks will employ what we are doing to prove this result for surfaces: they suppose the result is valid for the univariate case, and then they construct the Taylor series for most complex functions. We repeat, here, the idea of the argument.

Each f_j can be linearized as $f_j(\mathbf{x}) = f_j(\mathbf{x}_0) + \nabla f_j|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x} - \mathbf{x}_0)$, therefore,

$$\begin{aligned}\dot{x}_1 &= f_1(\mathbf{x}_0) + \nabla f_1^T|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x} - \mathbf{x}_0)\\ &\vdots\\ \dot{x}_n &= f_n(\mathbf{x}_0) + \nabla f_n^T|_{\mathbf{x}=\mathbf{x}_0}(\mathbf{x} - \mathbf{x}_0)\end{aligned}$$

gathering those together,

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}_0) + \mathbf{J}|_{\mathbf{x}=\mathbf{x}_0}(x - \mathbf{x}_0)$$

in which $\mathbf{J}$ is the Jacobian matrix, given by,

$$\mathbf{J} = \begin{bmatrix}\nabla f_1^T\\ \nabla f_2^T\\ \vdots\\ \nabla f_n^T\end{bmatrix} = \begin{bmatrix}\dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n}\\ \vdots & \vdots & \vdots\\ \dfrac{\partial f_n}{\partial x_1} & \cdots & \dfrac{\partial f_n}{\partial x_n}\end{bmatrix}$$

6.2 Stability of dynamic systems

6.2.1 Stability in linear fields

A linear field is a linear transform from $\mathbf{R}^n$ to itself. Thus, it makes sense to use the terminology to denote a matrix, which we will call here, in reference to state-space equations, as A. Moreover, let us assume that our differential equation does not have inputs, thus,

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

We want to show that the solutions of such equation relies on the spectra of A. Indeed, assume A is diagonalizable, that is, A has eigenvalues $\lambda_1, \cdots, \lambda_n$, associated with eigenvectors $\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_n$. By building functions $\mathbf{g}_i = \boldsymbol{\xi}_i e^{\lambda_i t}$, each of them is a solution of our ODE, since,

$$\begin{aligned}\dot{\mathbf{g}}_i &= \boldsymbol{\xi}_i \frac{d}{dt} e^{\lambda_i t}\\ &= \boldsymbol{\xi}_i \lambda_i e^{\lambda_i t}\\ &= \mathbf{A}\boldsymbol{\xi}_i e^{\lambda_i t}\\ &= \mathbf{A}\mathbf{g}_i\end{aligned}$$

Moreover, since the vectors $\boldsymbol{\xi}_i$ are linear independent, the solutions $\mathbf{g}_i$ are linearly independent. We conclude that the solution of our ODE is given by $\mathbf{x}(t) = \sum_{i=1}^{n} \boldsymbol{\xi}_i e^{\lambda_i t}$. We have just proven the following theorem,

Theorem 2. *Let $A \in \mathbb{R}^{n\times n}$ be the linear diagonalizable field of the ODE $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. Then,*

$$\mathbf{x}(t) = \sum_{i=1}^{n} \boldsymbol{\xi}_i e^{\lambda_i t}$$

Although not in its general form (we have assumed that A is diagonalizable), this theorem indeeds proves that the solutions of a linear ODE relies on the spectra of $\mathbf{A}$. It can be proven even if the field is not diagonalizable, using Jordan's canonical form. We then, state the following corollary,

Corollary 2. *If $\mathbf{A}$ have an eigenvalue with positive real part, then the response is asymptotically unstable.*

A proof of such corollary comes from noticing that, if any $\lambda_i > 0$, then the limit $\lim_{t\to\infty} \mathbf{x}(t) = \infty$, in the sense that at least one of its coordinates grows without bounds. We conclude that a full characterization of stable linear differential equations is to identify those as having negative definite fields.

6.2.2 Stability of nonlinear dynamic systems

The intents of this section is to characterize the equivalence between stationary points which are asymptotically stable (according to definition 4) and sink points, discussed in the corollary 1. To that end, we need a theorem to establish this equivalence.

Before presenting the theorem without its proof, it is noteworthy that, at the bottom, what the theorem really says is that in a sufficiently small neighborhood of each singularity, the nonlinear field $\mathbf{f}$ resembles the linear one, $\mathbf{A}$. This is exactly the procedure of linearization we have just described.

Being so, we have the famous,

Theorem 3. *Let $\mathbf{x}_0 \in \mathbb{R}^n$ be a singularity of the field $f : \mathbb{R}^n \to \mathbb{R}^n$. If x_0 is a sink of f, then x_0 is a asymptotically stable point of f.*

We do not present a proof here, because we believe it is out of the scope of the homework, although [3] has a detailed proof and discussion about such result. As a final remark, this theorem allows us to study the behavior of system's stability in the singularities of it.

References

[1] Adreas Kroll, Horst Schulte *Benchmark problems for nonlinear systems identification and control using Soft Computing methods: Need and overview*. Applied Soft Computing, 2014.

[2] Roger Penrose *The Road to Reality*. Vintage Books, 2007.

[3] Claus I. Doering, Artur O. Lopes *Equações Diferenciais Ordinárias*. IMPA, coleção Matemática Universitária, 2010.

[4] J. H. Hubbard, B. H. West *Differential Equations: A Dynamical Systems approach*. Springer-Verlag, Texts in Applied Mathematics, 1998.

[5] Bruce van Brunt *The Calculus of Variations*. Springer, Universitext, 2004.

[6] John L. Troutman *Variational Calculus and Optimal Control*. Springer, 1995.

[7] Gene F. Franklin, J. David Powell, Abbas Emami-Naeni *Feedback control of dynamic systems*. Seventh Edition, Pearson Higher Education, 2015.