

Classifying and Analysing Dynamical Regimes of the Double Pendulum

Authors:

Besty Huang,
Dariyan Khan,
Deepak Malhotra,
Dhruv Gupta,
Elly Hung,
Wilson Huang

Supervisor:
Dr Phil Ramsden

Github Link: https://github.com/Zhuoyue-Huang/M2R_DP

June 2022

Abstract

This paper investigates the dynamical properties of a double pendulum, ignoring the effects of air resistance and friction. We prove the existence of equilibrium points and carry out equilibrium point analysis using linearisation, showing that the double pendulum can have periodic or quasiperiodic motion for small initial angle conditions.

We then use a variety of tools to investigate how sensitive the double pendulum is to initial conditions on both a local and global scale, restricting the investigation to double pendulums with stiff massless unit length rods and unit point masses as the weights. Locally, patterns in the phase space can be visualised using Poincaré sections and bifurcation diagrams. Considering the system as a time series also then allows us to employ more tools such as Fast Fourier Transforms, recurrence plots, and statistical results via auto-correlation coefficients. In the global setting, we look at the maximal Lyapunov exponent, the time taken to double the initial difference, and the time taken for the pendulum to flip to analyse the transitions from non-chaotic to chaotic motion.

Contents

1	Introduction	1
2	Derivation of Master Equations	2
2.1	Notations	2
2.2	Lagrange's equations	2
2.3	First order system	3
3	Equilibrium Point Analysis	5
3.1	Existence	5
3.2	Stability by linearisation	7
3.2.1	Case 1: stable equilibrium point $\theta_1 = \theta_2 = 0$	7
3.2.2	Case 2: unstable equilibrium point $\theta_1 = 0, \theta_2 = \pi$	8
3.3	The stable equilibrium point	9
3.3.1	Plots against time	12
3.4	The unstable equilibrium points	14
4	Dynamical Analysis	15
4.1	Non-chaotic and chaotic motion	15
4.2	Time series	16
4.2.1	FFT method	16
4.2.2	Recurrence plots	20
4.3	Poincaré sections	22
4.3.1	Theory	22
4.3.2	Numerical results	22
4.4	Bifurcation diagrams	25
4.4.1	Theory	25
4.4.2	Numerical results	25
4.5	Lyapunov exponents	26
4.5.1	Theory	26
4.5.2	Numerical results	28
4.6	Sensitive dependence globally	29
4.6.1	Time for error to double	29

4.6.2	Time for pendulum to flip	31
4.6.3	Auto-Correlation Coefficients	34
5	Conclusion	39
A	Further stability by linearisation	43
A.0.1	Case 3: unstable equilibrium point $\theta_1 = \pi, \theta_2 = 0$	43
A.0.2	Case 4: unstable equilibrium point $\theta_1 = \theta_2 = \pi$	44
B	Varying parameters for a periodic linearisation result	45
C	Definition of Lyapunov exponent	48
D	More about recurrence quantification analysis	49
E	Figures referred to in Chapter 4	50
F	Explanation of the applicability of the Poincaré recurrence theorem	54
G	Proof of concept of obtaining the numerical solution via a neural network	57

Chapter 1

Introduction

Figure 1.1 is an example of a double pendulum. It is a physical system consisting of two pendulums attached end to end. A double pendulum is used often as an example or a teaching tool for the study of chaos in dynamical systems. This is because some regimes show a high sensitivity to initial conditions: i.e. two double pendulum systems starting from very close initial conditions can take very different trajectories after a period of time.

Specifically, this paper is an exploration on double pendulums where we ignore air resistance and friction, as well as reduce the balls of the pendulum to point masses and have the rods be without mass. The aim of this study is to find solutions to the equations governing the motion of the double pendulum through various methods and examine the dynamical properties in close detail.

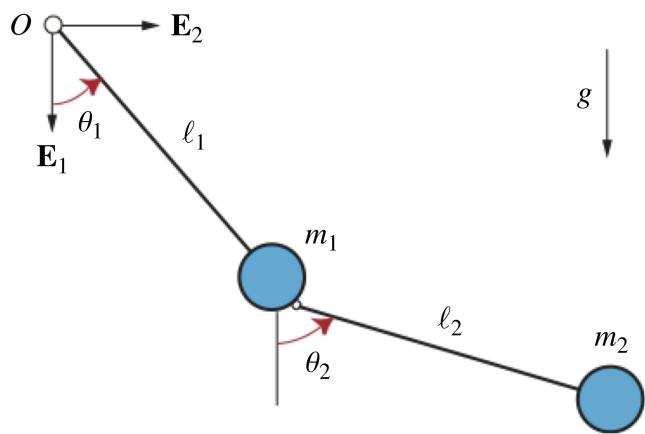


Figure 1.1: Double pendulum

Chapter 2

Derivation of Master Equations

2.1 Notations

Here we label the system using the lengths l_1 and l_2 , masses m_1 and m_2 and angles θ_1 and θ_2 to represent the location of the point masses in 2-D space.

The Cartesian coordinates of the two masses can be derived, and labeled as (x_1, y_1) and (x_2, y_2) respectively:

$$x_1 = l_1 \sin(\theta_1) \quad y_1 = -l_1 \cos(\theta_1) \quad (2.1a)$$

$$x_2 = x_1 + l_2 \sin(\theta_2) \quad y_2 = y_1 - l_2 \cos(\theta_2) \quad (2.1b)$$

2.2 Lagrange's equations

With notations above, Lagrange's equations (2.2) can be used to derive the equations of motion.

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}_i} \right) - \frac{\partial L}{\partial \theta_i} = 0 \quad \text{for } i = 1, 2 \quad (2.2)$$

Here the Lagrangian is $L = T - V$, where T represents the total kinetic energy of the system and V represents the total potential energy.

$$T = \frac{1}{2} m_1 (\dot{x}_1^2 + \dot{y}_1^2) + \frac{1}{2} m_2 (\dot{x}_2^2 + \dot{y}_2^2) \quad (2.3a)$$

$$V = m_1 g y_2 + m_2 g y_2 \quad (2.3b)$$

Given the Cartesian coordinates (2.1), we can find their derivatives as follows:

$$\dot{x}_1 = l_1 \cos(\theta_1) \dot{\theta}_1 \quad \dot{y}_1 = l_1 \sin(\theta_1) \dot{\theta}_1 \quad (2.4a)$$

$$\dot{x}_2 = \dot{x}_1 + l_2 \cos(\theta_2) \dot{\theta}_2 \quad \dot{y}_2 = \dot{y}_1 + l_2 \sin(\theta_2) \dot{\theta}_2 \quad (2.4b)$$

Substituting 2.4 into 2.3, we get

$$T = \frac{1}{2}(m_1 + m_2)l_1^2\dot{\theta}_1^2 + \frac{1}{2}m_2l_2^2\dot{\theta}_2^2 + m_2l_1l_2\dot{\theta}_1\dot{\theta}_2 \cos(\theta_1 - \theta_2) \quad (2.5a)$$

$$V = -(m_1 + m_2)gl_1 \cos(\theta_1) - m_2gl_2 \cos(\theta_2). \quad (2.5b)$$

Consequently, the Lagrangian of the system can be expressed in the following form:

$$\begin{aligned} L = T - V = & \frac{1}{2}(m_1 + m_2)l_1^2\dot{\theta}_1^2 + \frac{1}{2}m_2l_2^2\dot{\theta}_2^2 + m_2l_1l_2\dot{\theta}_1\dot{\theta}_2 \cos(\theta_1 - \theta_2) \\ & + (m_1 + m_2)gl_1 \cos(\theta_1) + m_2gl_2 \cos(\theta_2) \end{aligned} \quad (2.6)$$

Employing Lagrange's equations (2.2), the partial derivatives for θ_1 can be derived as

$$\frac{\partial L}{\partial \dot{\theta}_1} = (m_1 + m_2)l_1^2\dot{\theta}_1 + m_2l_1l_2\dot{\theta}_2 \cos(\theta_1 - \theta_2), \quad (2.7a)$$

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}_1} \right) = (m_1 + m_2)l_1^2\ddot{\theta}_1 + m_2l_1l_2\ddot{\theta}_2 \cos(\theta_1 - \theta_2) - m_2l_1l_2\dot{\theta}_2 \sin(\theta_1 - \theta_2)(\dot{\theta}_1 - \dot{\theta}_2), \quad (2.7b)$$

$$\frac{\partial L}{\partial \theta_1} = -m_2l_1l_2\dot{\theta}_1\dot{\theta}_2 \sin(\theta_1 - \theta_2) - (m_1 + m_2)gl_1 \sin(\theta_1). \quad (2.7c)$$

Similarly, for θ_2 , the derivatives can be written as

$$\frac{\partial L}{\partial \dot{\theta}_2} = m_2l_2^2\dot{\theta}_2 + m_2l_1l_2\dot{\theta}_1 \cos(\theta_1 - \theta_2), \quad (2.8a)$$

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}_2} \right) = m_2l_2^2\ddot{\theta}_2 + m_2l_1l_2\ddot{\theta}_1 \cos(\theta_1 - \theta_2) - m_2l_1l_2\dot{\theta}_1 \sin(\theta_1 - \theta_2)(\dot{\theta}_1 - \dot{\theta}_2), \quad (2.8b)$$

$$\frac{\partial L}{\partial \theta_2} = m_2l_1l_2\dot{\theta}_1\dot{\theta}_2 \sin(\theta_1 - \theta_2) - m_2gl_2 \sin(\theta_2) \quad (2.8c)$$

As a result, the system of Lagrange's equations can be simplified to

$$\begin{cases} (m_1 + m_2)l_1\ddot{\theta}_1 + m_2l_2\ddot{\theta}_2 \cos(\theta_1 - \theta_2) + m_2l_2\dot{\theta}_2^2 \sin(\theta_1 - \theta_2) + (m_1 + m_2)g \sin(\theta_1) = 0 \\ l_2\ddot{\theta}_2 + l_1\ddot{\theta}_1 \cos(\theta_1 - \theta_2) - l_1\dot{\theta}_1^2 \sin(\theta_1 - \theta_2) + g \sin(\theta_2) = 0 \end{cases} \quad (2.9)$$

which are the master equations of the double pendulum.

2.3 First order system

The coupled system as seen in equation (2.9) can be reduced to a system of first order differential equations. Let $\Delta = \theta_1 - \theta_2$. We can then write the system as a matrix:

$$\begin{bmatrix} p_{\theta_1} \\ p_{\theta_2} \end{bmatrix} \begin{bmatrix} m_1 + m_2 & m_2 \cos(\Delta) \\ \cos(\Delta) & 1 \end{bmatrix} = \begin{bmatrix} -m_2 l_2 \sin(\Delta) \dot{\theta}_2^2 - (m_1 + m_2) g \sin(\theta_1) \\ l_1 \sin(\Delta) \dot{\theta}_1^2 - g \sin(\theta_2) \end{bmatrix} \quad (2.10)$$

We divide the top row by m_2 , let $\mu = 1 + \frac{m_1}{m_2}$, and let $\varphi_1 = \dot{\theta}_1$ and $\varphi_2 = \dot{\theta}_2$:

$$\begin{bmatrix} \mu & \cos(\Delta) \\ \cos(\Delta) & 1 \end{bmatrix} \begin{bmatrix} l_1 \dot{\varphi}_1 \\ l_2 \dot{\varphi}_2 \end{bmatrix} = \begin{bmatrix} -\mu g \sin(\theta_1) - l_2 \sin(\Delta) \varphi_2^2 \\ l_1 \sin(\Delta) \varphi_1^2 - g \sin(\theta_2) \end{bmatrix} \quad (2.11)$$

Let \mathbf{A} be the 2x2 matrix on the LHS. $\mu = 1 + \frac{m_1}{m_2} > 1$ as $m_1, m_2 > 0$ and since $\cos^2(\Delta) \leq 1$, we must have that \mathbf{A} is invertible. Then one can obtain

$$\mathbf{A}^{-1} = \frac{1}{\mu - \cos^2(\Delta)} \begin{bmatrix} 1 & -\cos(\Delta) \\ -\cos(\Delta) & \mu \end{bmatrix}. \quad (2.12)$$

Therefore,

$$\begin{bmatrix} l_1 \dot{\varphi}_1 \\ l_2 \dot{\varphi}_2 \end{bmatrix} = \frac{1}{\mu - \cos^2(\Delta)} \begin{bmatrix} 1 & -\cos(\Delta) \\ -\cos(\Delta) & \mu \end{bmatrix} \begin{bmatrix} -\mu g \sin(\theta_1) - l_2 \sin(\Delta) \varphi_2^2 \\ l_1 \sin(\Delta) \varphi_1^2 - g \sin(\theta_2) \end{bmatrix}. \quad (2.13)$$

Remembering that $m_2 \mu = m_2 + m_1$, we can rearrange the system to get:

$$\dot{\varphi}_1 = \frac{m_2 g \sin(\theta_2) \cos(\Delta) - m_2 \sin(\Delta) [l_1 \varphi_1^2 \cos(\Delta) + l_2 \varphi_2^2] - (m_1 + m_2) g \sin(\theta_1)}{l_1 [m_1 + m_2 \sin^2(\Delta)]} \quad (2.14a)$$

$$\dot{\varphi}_2 = \frac{(m_1 + m_2) [l_1 \sin(\Delta) \varphi_1^2 - g \sin(\theta_2) + g \sin(\theta_1) \cos(\Delta)] + m_2 l_2 \varphi_2^2 \sin(\Delta) \cos(\Delta)}{l_2 [m_1 + m_2 \sin^2(\Delta)]} \quad (2.14b)$$

Thus, our double pendulum system can be represented by the following first order equations:

$$\begin{cases} \dot{\theta}_1 = \varphi_1 \\ \dot{\theta}_2 = \varphi_2 \\ \dot{\varphi}_1 = \frac{m_2 g \sin(\theta_2) \cos(\Delta) - m_2 \sin(\Delta) [l_1 \varphi_1^2 \cos(\Delta) + l_2 \varphi_2^2] - (m_1 + m_2) g \sin(\theta_1)}{l_1 [m_1 + m_2 \sin^2(\Delta)]} \\ \dot{\varphi}_2 = \frac{(m_1 + m_2) [l_1 \sin(\Delta) \varphi_1^2 - g \sin(\theta_2) + g \sin(\theta_1) \cos(\Delta)] + m_2 l_2 \varphi_2^2 \sin(\Delta) \cos(\Delta)}{l_2 [m_1 + m_2 \sin^2(\Delta)]} \end{cases} \quad (2.15)$$

Chapter 3

Equilibrium Point Analysis

Using the master equations (2.15) we obtained in the last chapter, interesting properties of double pendulums can be derived using both analytical and numerical methods. In this chapter, we will mainly focus on the existence and stability of equilibrium points by linearisation, then show that the two angles θ_1 and θ_2 of the double pendulum near the stable point can be approximately seen as a coupled oscillator.

3.1 Existence

Definition 3.1.1 (Equilibrium point). The point $\tilde{\mathbf{x}} \in \mathbb{R}^n$ is an *equilibrium point* if it is a constant solution for the differential equation

$$\frac{d\mathbf{x}}{dt} = f(t, \mathbf{x}). \quad (3.1)$$

i.e. $f(t, \tilde{\mathbf{x}}) = 0$ for all t ,

Lemma 3.1.1 (Existence of equilibrium points). *In a double pendulum system without external force, there are exactly 4 equilibrium points.*

Proof. Given the differential equations (2.15) we derived, an equilibrium point is a point such that $\dot{\theta}_1, \dot{\theta}_2, \ddot{\theta}_1$ and $\ddot{\theta}_2$ are 0. Thus we have

$$\frac{m_2 g \sin(\theta_2) \cos(\theta_1 - \theta_2) - (m_1 + m_2) g \sin(\theta_1)}{l_1 [m_1 + m_2 \sin^2(\theta_1 - \theta_2)]} = 0 \quad (3.2a)$$

$$\frac{(m_1 + m_2)(-g \sin(\theta_2) + g \sin(\theta_1) \cos(\theta_1 - \theta_2))}{l_2 [m_1 + m_2 \sin^2(\theta_1 - \theta_2)]} = 0 \quad (3.2b)$$

Since the denominators for both equations are nonzero, we can simply solve the following equations

$$m_2 g \sin(\theta_2) \cos(\theta_1 - \theta_2) = (m_1 + m_2) g \sin(\theta_1) \quad (3.3a)$$

$$\sin(\theta_1) \cos(\theta_1 - \theta_2) = \sin(\theta_2). \quad (3.3b)$$

Rearranging 3.3 and applying the trigonometric identities for addition of angles we get:

$$\cos(\theta_1) \sin(\theta_1 - \theta_2) = 0 \quad (3.4a)$$

$$m_2 \sin(\theta_2) \cos(\theta_1 - \theta_2) = (m_1 + m_2) \sin(\theta_1) \quad (3.4b)$$

From 3.4a, we have $\theta_1 = \pi/2 + k\pi$ or $\theta_1 = \theta_2 + k\pi$ for $k \in \mathbb{Z}$.

Suppose $\theta_1 = \pi/2 + k\pi$ **for** $k \in \mathbb{Z}$:

In this case,

$$\sin(\theta_1) = \pm 1 \quad (3.5a)$$

$$\cos(\theta_1) = 0 \quad (3.5b)$$

Then one can derive that

$$\sin^2(\theta_2) = -\frac{m_1 + m_2}{m_2} < 0, \quad (3.6)$$

which is impossible. Thus this case should be ruled out.

Suppose $\theta_1 = \theta_2 + k\pi$ **for** $k \in \mathbb{Z}$:

For the case $k = 2\mathbb{Z}$,

$$\sin(\theta_1) = \sin(\theta_2) = \sin(\theta) \quad (3.7a)$$

$$\cos(\theta_1) = \cos(\theta_2) = \cos(\theta). \quad (3.7b)$$

By substituting 3.7 into 3.4b,

$$m_2 \sin(\theta)(\cos^2(\theta) + \sin^2(\theta)) = (m_1 + m_2) \sin(\theta). \quad (3.8)$$

Then $m_1 \sin(\theta) = 0$ can be derived for this case. Hence $\theta_1 = k\pi$, $\theta_2 = k\pi + 2j\pi$ for $k, j \in \mathbb{Z}$. If we restrict $\theta_i \in (-\pi, \pi]$, we get $\theta_1 = \theta_2 = 0$ or $\theta_1 = \theta_2 = \pi$.

Similarly, in terms of the case when $k = 2\mathbb{Z} + 1$,

$$\sin(\theta_1) = -\sin(\theta_2) \quad (3.9a)$$

$$\cos(\theta_1) = -\cos(\theta_2). \quad (3.9b)$$

Therefore,

$$m_2(-\sin(\theta_1))(-(cos^2(\theta_1) + sin^2(\theta_1)) = (m_1 + m_2) \sin(\theta_1), \quad (3.10a)$$

leading to solutions $\theta_1 = k\pi$, $\theta_2 = k\pi + (2j + 1)\pi$ for $k, j \in \mathbb{Z}$. Thus $\theta_1 = 0$, $\theta_2 = \pi$ or $\theta_1 = \pi$, $\theta_2 = 0$ are only possible solutions for $\theta_i \in (-\pi, \pi]$. Hence the possible coordinates for equilibrium points of (θ_1, θ_2) in $\theta_1, \theta_2 \in (-\pi, \pi]$ are $(0, 0)$, $(0, \pi)$, $(\pi, 0)$, (π, π) . ■

One can even make a guess by physical intuition that $(0, 0)$ is the only stable point. These equilibrium points will be carefully analysed by the next section.

3.2 Stability by linearisation

3.2.1 Case 1: stable equilibrium point $\theta_1 = \theta_2 = 0$

We first consider the equations of motion for small oscillations around the stable equilibrium. To do this, we work to 2nd order in $\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2$, and assume that these quantities are small, which allows us to use the following Taylor approximations of sine and cosine about 0.

$$\sin(\theta) \approx \theta \quad (3.11a)$$

$$\cos(\theta) \approx 1 - \frac{\theta^2}{2} \quad (3.11b)$$

Applying our approximation, the coordinates (2.1) become

$$x_1 \approx l_1 \theta_1 \quad y_1 \approx -l_1 \left(1 - \frac{\theta_1^2}{2}\right) \quad (3.12a)$$

$$x_2 \approx l_1 \dot{\theta}_1 + l_2 \theta_2 \quad y_2 \approx -l_1 \left(1 - \frac{\theta_1^2}{2}\right) - l_2 \left(1 - \frac{\theta_2^2}{2}\right) \quad (3.12b)$$

As before, we consider the Lagrangian (2.2), which consists of the kinetic energy (T) of the system minus the potential energy (V).

$$T = \frac{1}{2} m_1 l_1^2 \dot{\theta}_1^2 + \frac{1}{2} m_2 (\dot{x}_2^2 + \dot{y}_2^2) \quad (3.13)$$

Using 3.12, we differentiate to obtain $y_2 \approx l_1 \dot{\theta}_1 + l_2 \dot{\theta}_2$. Hence upon squaring this term in 3.13, we obtain fourth order terms. In contrast, $\dot{x}_2 \approx l_1 \dot{\theta}_1 + l_2 \dot{\theta}_2$ only has derivative terms. Thus $\dot{x}_2^2 \gg \dot{y}_2^2$, and the second term can be ignored.

Similarly, for potential energy, we substitute into 2.3 to obtain

$$T = \frac{1}{2} m_1 l_1^2 \dot{\theta}_1^2 + \frac{1}{2} m_2 (l_1^2 \dot{\theta}_1^2 + 2l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 + l_2^2 \dot{\theta}_2^2) \quad (3.14a)$$

$$V = \frac{1}{2} m_1 g l_1 \theta_1^2 + \frac{1}{2} m_2 g (l_1^2 \theta_1^2 + l_2^2 \theta_2^2) + K \quad (3.14b)$$

where the constant K can be omitted in subsequent calculations, as the zero point in calculating potential energy is arbitrary. We now apply Lagrange's equations (2.2) to obtain

$$(1 + \frac{m_2}{m_1}) \ddot{\theta}_1 + \frac{m_2 l_2}{m_1 l_1} \ddot{\theta}_2 + (\frac{g}{l_1} + \frac{m_2 g}{m_1 l_1}) \theta_1 = 0 \quad (3.15a)$$

$$\frac{l_1}{l_2} \ddot{\theta}_1 + \ddot{\theta}_2 + \frac{g}{l_2} \theta_2 = 0 \quad (3.15b)$$

This system can also be expressed in matrix form:

$$\begin{bmatrix} 1 + \frac{m_2}{m_1} & \frac{m_2 l_2}{m_1 l_1} \\ l_1/l_2 & 1 \end{bmatrix} \begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} + \begin{bmatrix} \frac{g}{l_1} + \frac{m_2 g}{m_1 l_1} & 0 \\ 0 & \frac{g}{l_2} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = 0 \quad (3.16)$$

If we consider the simple case that $m_1 = m_2 = m$ and $l_1 = l_2 = l$, we obtain:

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} + \frac{g}{l} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = 0 \quad (3.17)$$

We prove below that the equilibrium point $\theta_1 = \theta_2 = 0$ is indeed stable; in other words, small oscillations around it do not diverge, as demonstrated in Section 3.3.

3.2.2 Case 2: unstable equilibrium point $\theta_1 = 0, \theta_2 = \pi$

We now focus on the double pendulum system under the simplified model $l_1 = l_2 = m_1 = m_2 = 1$. Consider first a small perturbation about this unstable equilibrium and let $\pi - \theta = \phi$. Then for $\theta \approx \pi$:

$$\sin(\theta) \approx \pi - \theta = \phi \quad (3.18a)$$

$$\cos(\theta) \approx -1 + \frac{(\pi - \theta)^2}{2} = -1 + \frac{\phi^2}{2} \quad (3.18b)$$

Then we re-approximate x_2 and y_2

$$x_2 = \sin(\theta_1) + \sin(\theta_2) \approx \theta_1 + \phi_2 \quad (3.19a)$$

$$y_2 = -\cos(\theta_1) - \cos(\theta_2) \approx \frac{\theta_1^2}{2} - \frac{\phi_2^2}{2} \quad (3.19b)$$

We can then easily find the derivatives with respect to time.

$$\dot{x}_2 \approx \dot{\theta}_1 + \dot{\phi}_2 \quad (3.20a)$$

$$\dot{y}_2 \approx \theta_1 \dot{\theta}_1 - \phi_2 \dot{\phi}_2 \quad (3.20b)$$

We refer back to equation (2.3) for the total energies which gives us that:

$$T = \frac{\dot{\theta}_1^2}{2} + \frac{\dot{\theta}_1^2 + 2\dot{\theta}_1\dot{\phi}_2 + \dot{\phi}_2^2}{2} \quad (3.21)$$

$$V = g \left[-\left(1 - \frac{\theta_1^2}{2}\right) - \left(1 - \frac{\theta_1^2}{2}\right) + \left(1 - \frac{\phi_2^2}{2}\right) \right] \quad (3.22)$$

Plugging T and V into the Laplacian, we find that:

$$L = \dot{\theta}_1^2 + \dot{\theta}_1\dot{\phi}_2 + \frac{\dot{\phi}_2^2}{2} - g\theta_1^2 + \frac{g\phi_2^2}{2} - g \quad (3.23)$$

Applying Lagrange's equations (2.2), we find:

$$2\ddot{\theta}_1 + \ddot{\phi}_2 = -2g\theta_1 \quad (3.24a)$$

$$\ddot{\theta}_1 + \ddot{\phi}_2 = g\phi_2 \quad (3.24b)$$

We can put these equations into matrix vector form to give us a system of equations.

$$\underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\phi}_2 \end{bmatrix} = g \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \phi_2 \end{bmatrix} \quad (3.25)$$

We can invert \mathbf{A} to give us a matrix equation for the second derivatives of θ_1 and ϕ_2

$$\begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\phi}_2 \end{bmatrix} = g \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \phi_2 \end{bmatrix} \quad (3.26)$$

Which we find to be:

$$\begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\phi}_2 \end{bmatrix} = g \underbrace{\begin{bmatrix} -2 & -1 \\ 2 & 2 \end{bmatrix}}_{\mathbf{B}} \begin{bmatrix} \theta_1 \\ \phi_2 \end{bmatrix} \quad (3.27)$$

Finding the eigenvalues of the matrix \mathbf{B} will tell us the stability of the equilibrium point $(\theta_1, \theta_2) = (0, \pi)$. We find that there are two eigenvalues $\lambda_1 = \sqrt{2}g > 0$ and $\lambda_2 = -\sqrt{2}g < 0$ with corresponding eigenvectors $E_1 = \begin{bmatrix} -\sqrt{2}+1 \\ \sqrt{2} \\ 1 \end{bmatrix}$ and $E_2 = \begin{bmatrix} -(\sqrt{2}+1) \\ \sqrt{2} \\ 1 \end{bmatrix}$.

The equilibrium point $(\theta_1, \theta_2) = (0, \pi)$ is a saddle point, thus it is unstable.

Due to space constraints, detailed analysis of the other two unstable equilibrium points can be found in Appendix A.

3.3 The stable equilibrium point

The system in 3.17 is reminiscent of the general form for simple harmonic motion (SHM). This similarity motivates us to conduct the next piece of analysis: the normal modes of oscillation under SHM; in other words, a linear combination of simple harmonic oscillations[1].

Under SHM, we typically have solutions of the form:

$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \text{Re}(\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} e^{i\omega_1 t} + \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} e^{i\omega_2 t}) \quad (3.28)$$

where the a_{ij} terms are constants. We can also restate this system in real terms as follows:

$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \cos(\omega t + \phi) \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad (3.29)$$

In this case, a_1 and a_2 are real constants, ω gives the angular velocity, and ϕ gives the phase shift. We now use 3.29 as an ansatz in order to solve 3.17. This yields the following matrix equation:

$$\begin{bmatrix} 2\frac{g}{l} - 2\omega^2 & -\omega^2 \\ -\omega^2 & \frac{g}{l} - \omega^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = 0 \quad (3.30)$$

for a non-trivial solution to this system, the matrix must be non-invertible with determinant zero. This gives the equation $2(\frac{g}{l} - \omega^2)^2 - \omega^4 = 0$, which can be rearranged and solved for the natural frequencies:

$$\omega_+ = \sqrt{2 + \sqrt{2}} \sqrt{\frac{g}{l}} \quad (3.31a)$$

$$\omega_- = \sqrt{2 - \sqrt{2}} \sqrt{\frac{g}{l}} \quad (3.31b)$$

Note that upon taking the second square root, we can keep this positive as cosine is an even function. It is also interesting to consider the 2nd component of 3.30, which gives $-\omega^2 a_1 + (\frac{g}{l} - \omega^2) a_2 = 0 \implies \frac{a_2}{a_1} = \frac{\omega^2}{\frac{g}{l} - \omega^2}$. Upon substituting 3.31, we can show

$$\frac{a_2}{a_1} = \frac{2 \pm \sqrt{2}}{-1 \mp \sqrt{2}} = \frac{-2 + 2 \pm 2\sqrt{2} \mp \sqrt{2}}{1 - 2} = \mp \sqrt{2}. \quad (3.32)$$

Considering 3.31 and 3.32, for $\omega = \omega_+ = \sqrt{2 + \sqrt{2}} \sqrt{\frac{g}{l}}$, we have $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \propto \begin{bmatrix} 1 \\ -\sqrt{2} \end{bmatrix}$. This is the first normal mode, depicted in figure 3.1.

The case for $\omega = \omega_+$ corresponds to the case where both masses are restored to equilibrium by gravity, but the bottom mass also provides a restoring force upon the top mass, increasing the overall normal frequency.

Considering 3.31 and 3.32, for $\omega_- = \sqrt{2 - \sqrt{2}} \sqrt{\frac{g}{l}}$, we have $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \propto \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$. This is the second normal mode, depicted in figure 3.2.

The case for $\omega = \omega_-$ corresponds to the case where both masses are restored to equilibrium by gravity, but this time the restoring forces from each mass cancel, resulting in a lower normal frequency.

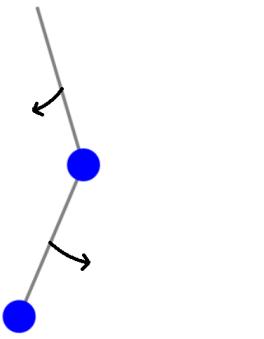


Figure 3.1: Normal Mode 1

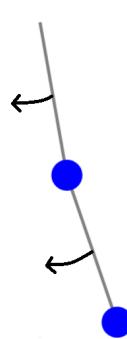


Figure 3.2: Normal Mode 2

In general, we can combine these two normal modes to create the superposition of two cosine functions.

$$\begin{bmatrix} \theta_1(t) \\ \theta_2(t) \end{bmatrix} = A \begin{bmatrix} 1 \\ -\sqrt{2} \end{bmatrix} \cos(\omega_+ t + \phi_1) + B \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix} \cos(\omega_- t + \phi_2) \quad (3.33)$$

This is in fact an example of quasiperiodic behaviour due to the irrational natural frequency ratio [2] under the assumption that $m_1 = m_2 = 1$ and $l_1 = l_2 = l$, as demonstrated below (3.34). Note that this is only the case for these particular parameter values, and it is possible to tweak these to obtain a rational frequency ratio, which would be periodic (please refer to Appendix B).

$$\frac{\omega_+}{\omega_-} = \frac{\sqrt{2 + \sqrt{2}}}{\sqrt{2 - \sqrt{2}}} = \frac{\sqrt{2 + \sqrt{2}}}{\sqrt{2 - \sqrt{2}}} \frac{\sqrt{2 + \sqrt{2}}}{\sqrt{2 + \sqrt{2}}} = \frac{2 + \sqrt{2}}{\sqrt{2}} = 1 + \sqrt{2} \quad (3.34)$$

To prove that $1 + \sqrt{2}$ is irrational, assume that it is rational and equal to $\frac{a}{b}$ where $a, b \in \mathbb{Z}, b \neq 0$. Then, $\sqrt{2} = \frac{a}{b} - 1 = \frac{a-b}{b}$ which is rational. But, we know that $\sqrt{2}$ is irrational, and thus we have a contradiction. ■

We demonstrate this numerically using the FFT and further visualisations in (4.2.1). Furthermore, if we assume that the phase differences are equal to zero, we obtain

$$\begin{bmatrix} \theta_1(t) \\ \theta_2(t) \end{bmatrix} = A \begin{bmatrix} 1 \\ -\sqrt{2} \end{bmatrix} \cos(\omega_+ t) + B \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix} \cos(\omega_- t) \quad (3.35)$$

The general solutions for both masses have concealed within them two distinct simple harmonic oscillations. These combine together in distinct proportions to yield the general solution to the double pendulum. If we consider linear combinations of θ_1 and θ_2 , we can isolate the pure sinusoidal oscillations for each frequency.

$$\sqrt{2}\theta_1 + \theta_2 = 2\sqrt{2}B \cos(\omega_- t) \quad (3.36a)$$

$$\sqrt{2}\theta_1 - \theta_2 = 2\sqrt{2}A \cos(\omega_+ t) \quad (3.36b)$$

Hence the masses themselves are not executing SHM, but the linear combinations of the coordinates of the masses are executing SHM.

3.3.1 Plots against time

The following figures have been plotted using python and the solve_ivp function (from the scipy.integrate sub-package) for the non-linearised system. The initial conditions $(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2) = (\frac{\pi}{10}, \frac{\pi}{100}, 0, 0)$ have been chosen to ensure that the displacement is small, consisting of the superposition of sinusoidal oscillations. This quasiperiodic behaviour for the general solution is clear in Figure 3.3, and will be demonstrated using various tools in the next chapter.

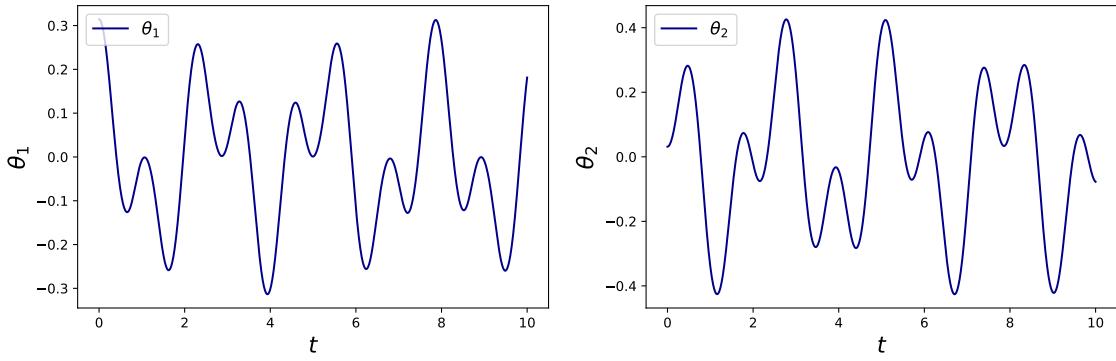


Figure 3.3: Amplitude of θ_1 and θ_2 with 10s time duration

Figure 3.4 highlights the above observation that linear combinations of θ_1 and θ_2 yield pure SHM, as per 3.36. Again, $\sqrt{2}\theta_1 + \theta_2$ oscillates with frequency $\omega_- t = \sqrt{2 - \sqrt{2}}\sqrt{\frac{g}{l}}$, whereas $\sqrt{2}\theta_1 - \theta_2$ oscillates with frequency $\omega_+ t = \sqrt{2 + \sqrt{2}}\sqrt{\frac{g}{l}}$.

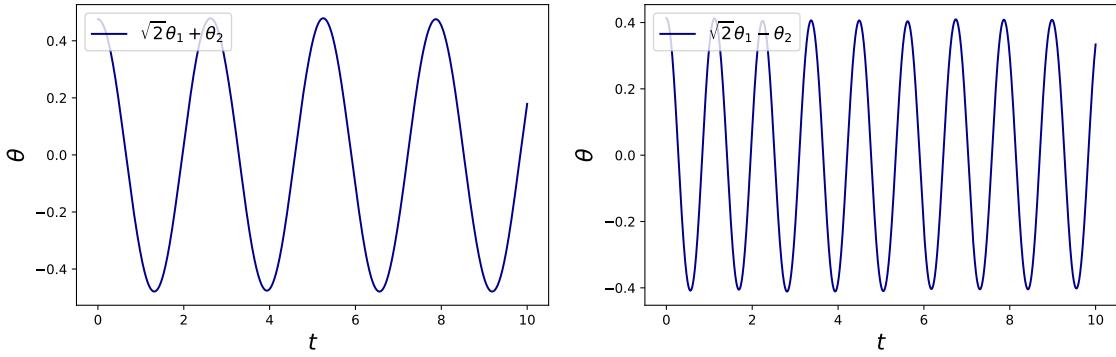


Figure 3.4: Amplitude of $\sqrt{2}\theta_1 + \theta_2$ and $\sqrt{2}\theta_1 - \theta_2$ with 10s time duration

Additionally, by considering initial conditions in the correct ratio we can visualise the normal modes of oscillation of the system.

Firstly, we consider $(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2) = (\frac{\pi}{10}, \sqrt{2}\frac{\pi}{10}, 0, 0)$. This yields the behaviour in normal mode 1, as per Figure 3.1, plotted on the left in Figure 3.5. This corresponds to the case where the each particle is executing SHM under $\omega_+ = \sqrt{2 + \sqrt{2}}\sqrt{\frac{g}{l}}$. However, as the pendulums move in opposite directions, they are π radians out of phase.

Secondly, we can consider $(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2) = (\frac{\pi}{10}, -\sqrt{2}\frac{\pi}{10}, 0, 0)$. This yields the behaviour in normal mode 2, as per Figure 3.2, plotted on the right in Figure 3.5. This corresponds to the case where the entire configuration is executing SHM under $\omega_- = \sqrt{2 - \sqrt{2}}\sqrt{\frac{g}{l}}$ as the pendulums move together.

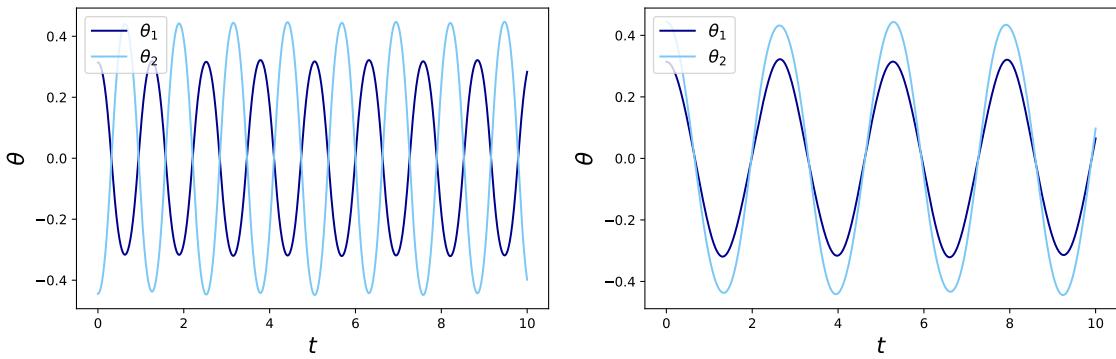


Figure 3.5: Normal Modes of Oscillation

Finally, we can also compare the linearised solutions to the general solution for small displacements. In Figure 3.6, we observe that for small times, the linearisation very closely follows the general solution. However for larger times, there appears to be some divergence; the system shows sensitivity to initial conditions, explored further in the next chapter.

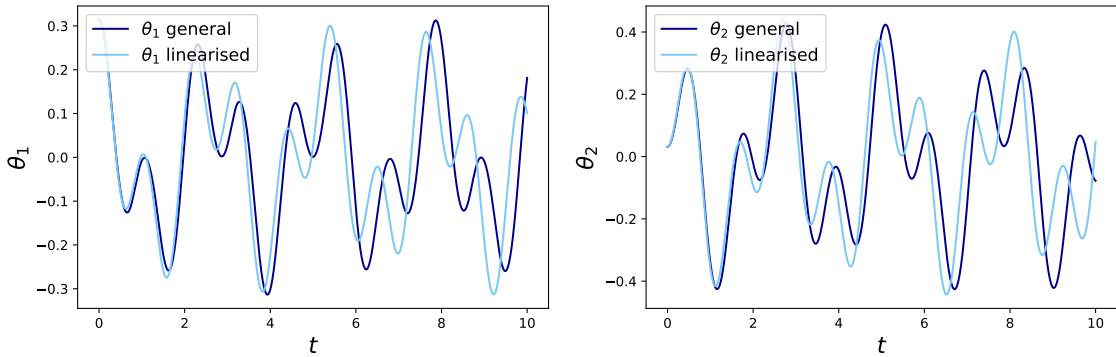


Figure 3.6: Comparison of amplitudes of θ_1 and θ_2 for general, and linearised systems with 10s time duration

3.4 The unstable equilibrium points

Again consider the general solution under simple harmonic motion as before. We use the same notation as in section 3.2.

$$\begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \cos(\omega t + \Phi) \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad (3.37)$$

First we analyse the equilibrium point at $(\theta_1, \theta_2) \approx (0, \pi)$. Under the same logic we have:

$$\begin{bmatrix} -2g + \omega^2 & -g \\ 2g & 2g + \omega^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = 0, \quad (3.38)$$

which gives us the equation

$$(-2g + \omega^2)(2g + \omega^2) + 2g^2 = 0. \quad (3.39a)$$

Solving this we find that $\omega^4 = 2g^2$. Hence ω has two real roots which are $w = \pm\sqrt{\sqrt{2}g}$. As before, the general solution can be written as a function of the eigenvalues and eigenvectors which we calculated earlier in the section 3.2.2.

$$\begin{bmatrix} \theta_1(t) \\ \phi_2(t) \end{bmatrix} = A \begin{bmatrix} -\sqrt{2} + 1 \\ \sqrt{2} \end{bmatrix} \cos(\sqrt{\sqrt{2}gt} + \phi_{11}) + B \begin{bmatrix} -(\sqrt{2} + 1) \\ \sqrt{2} \end{bmatrix} \cos(-\sqrt{\sqrt{2}gt} + \phi_{21}) \quad (3.40)$$

Physically for this equilibrium point, if our initial conditions are in the ratio of the first eigenvalue in the equation we will return to the equilibrium.

Similarly, for $(\theta_1, \theta_2) \approx (\pi, 0)$, we have:

$$\begin{bmatrix} \phi_1(t) \\ \theta_2(t) \end{bmatrix} = A \begin{bmatrix} -(\sqrt{2} + 1) \\ \sqrt{2} \end{bmatrix} \cos(\sqrt{\sqrt{2}gt} + \phi_{12}) + B \begin{bmatrix} -(\sqrt{2} - 1) \\ \sqrt{2} \end{bmatrix} \cos(-\sqrt{\sqrt{2}gt} + \phi_{22}) \quad (3.41)$$

Lastly, for $(\theta_1, \theta_2) \approx (\pi, \pi)$,

$$\begin{bmatrix} 2g + \omega^2 & -g \\ -2g & 2g + w^2 \end{bmatrix} \begin{bmatrix} \phi_{13} \\ \phi_{23} \end{bmatrix} = 0 \quad (3.42)$$

Solutions $\omega = \pm\sqrt{-2g \pm \sqrt{2}g}$ can be derived as eigenvalues. All roots in this case are imaginary, which implies that the general solution is the linear combination of cosh functions. Therefore, one would always derive a divergence from the equilibrium point.

Chapter 4

Dynamical Analysis

4.1 Non-chaotic and chaotic motion

At low initial energy levels, the motion of a double pendulum can be periodic (if the ratio of the two angular velocities is rational) or quasiperiodic (if the ratio is irrational) [2]. For quasiperiodic motion, two double pendulums that are initially close will swing in approximately the same way. This contrasts with chaotic motion, where two double pendulums with arbitrarily close initial conditions will diverge after a sufficiently long period of time, and the motion is said to be ‘memoryless’ [3]. In this chapter, we will mainly focus on numerical methods to support linearisation results in the last chapter, followed by cases where transitions from chaotic to non-chaotic behaviour occur. To simplify the analysis of the transitions, the symmetric initial condition $\dot{\theta}_1(0) = \dot{\theta}_2(0)$ will be carefully illustrated along with other cases. Unless otherwise stated, in this section, we take $\dot{\theta}_1(0) = \dot{\theta}_2(0) = 0$.

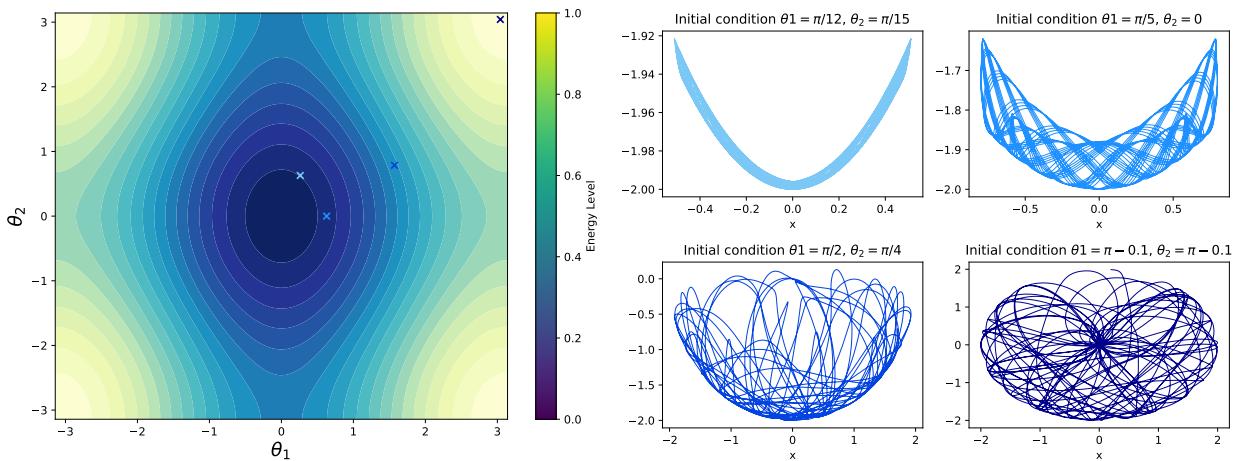


Figure 4.1: Energy level and trajectories of m_2 with zero kinetic energy

4.2 Time series

4.2.1 FFT method

A Fast Fourier transform (FFT) is an efficient algorithm which computes the Discrete Fourier transform (DFT) of a sequence. This allows us to transform a signal in the time/space domain to one in the frequency domain (f). Using $f = \frac{\omega}{2\pi}$, we can then convert the amplitude of θ from the time domain to its angular velocity (ω) domain. Therefore, if the angular velocity domain has two significant peaks with the values found in 3.31, our analytical and numerical results for determining quasiperiodicity match [1].

Quasiperiodic cases

Example 4.2.1 (FFT for small oscillation). Consider the initial condition $\theta_1 = \frac{\pi}{12}$, $\theta_2 = \frac{\pi}{15}$, with the pendulum released from rest.

Figure 4.2 demonstrates there are two peaks of angular velocity. The motion of small oscillations can be seen as the superposition of sinusoids, where the peak values are $\sqrt{\frac{g}{l}}\sqrt{2 - \sqrt{2}}$ and $\sqrt{\frac{g}{l}}\sqrt{2 + \sqrt{2}}$. In this case, the FFT method coincides with linearisation results.

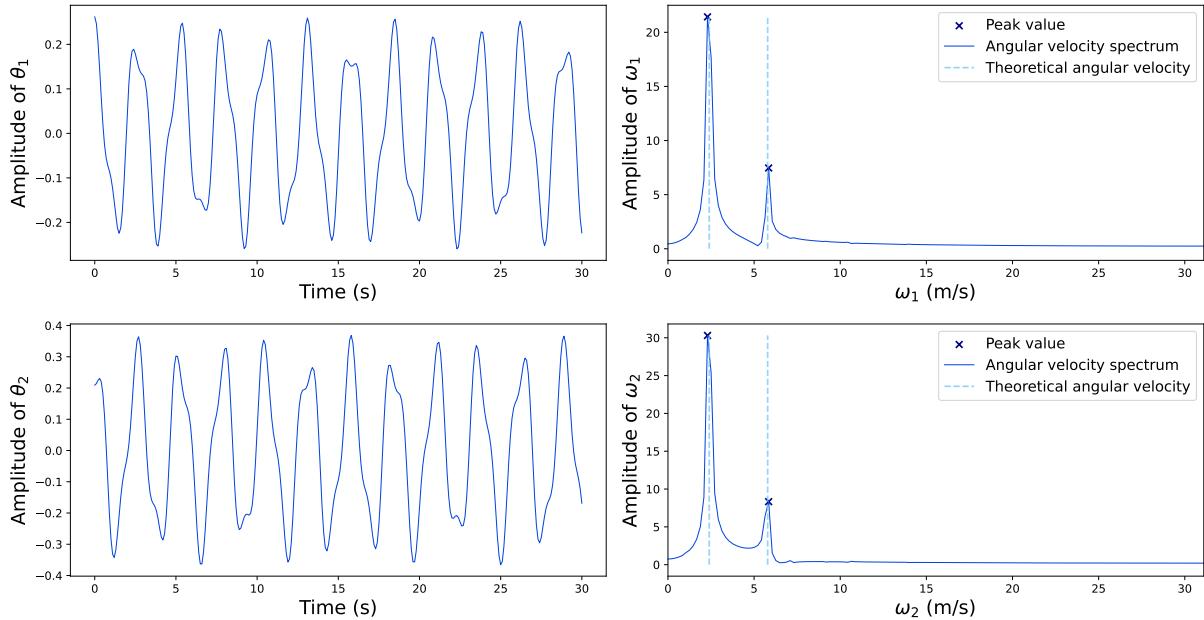


Figure 4.2: Amplitude of θ and ω with 30s time duration

As shown in Figure 4.3, the resolution will be higher in terms of peak angular velocity once we extend the time duration, which further supports the superposition effect.

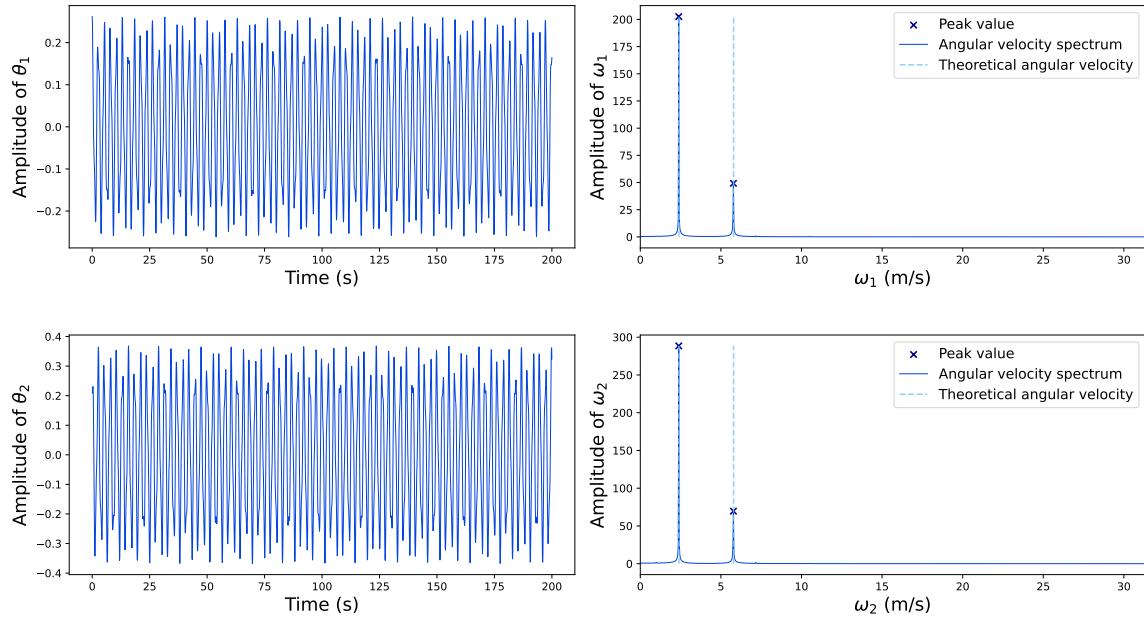


Figure 4.3: Amplitude of θ and ω with 200s time duration

Example 4.2.2 (FFT for an approximating periodic case). Consider the initial condition $\theta_1 = \frac{\pi}{10}$, $\theta_2 = \frac{\sqrt{2}\pi}{10}$, with the pendulum released from rest.

Figure 4.4 indicates that when θ_i is sufficiently small and $\frac{\theta_2}{\theta_1} = \sqrt{2}$, the initial condition would lead to the double pendulum approximately following simple harmonic motion.

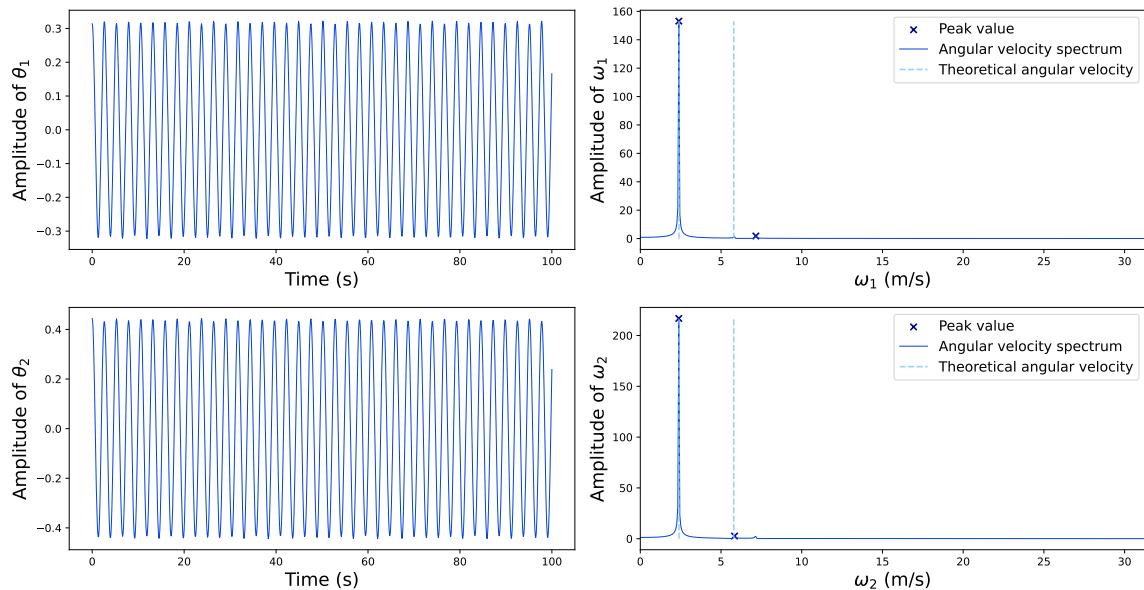


Figure 4.4: Amplitude of θ and ω for an approximating periodic case

Sensitive dependence cases

When θ_i is sufficiently large, linearisation results are not applicable. The motion of the double pendulum will be more sensitive to its initial conditions, which means an arbitrary small difference in initial conditions will cause very different trajectories in the long run, i.e. chaotic motion.

Example 4.2.3. Consider the initial conditions for a first pendulum $y_1 = (\theta_1, \theta_2) = (\frac{\pi}{2}, \frac{\pi}{4})$ and that of a second pendulum $y_2 = (\theta_1, \theta_2) = (\frac{\pi}{2}, \frac{\pi}{4} - 0.01)$, both released from rest.

In Figure 4.5, the two double pendulum systems have roughly the same behaviour in the first 10 seconds, then their trajectories seem to diverge and become uncorrelated, which supports the hypothesis they are more sensitive to their initial conditions compared with the first two small angle cases.

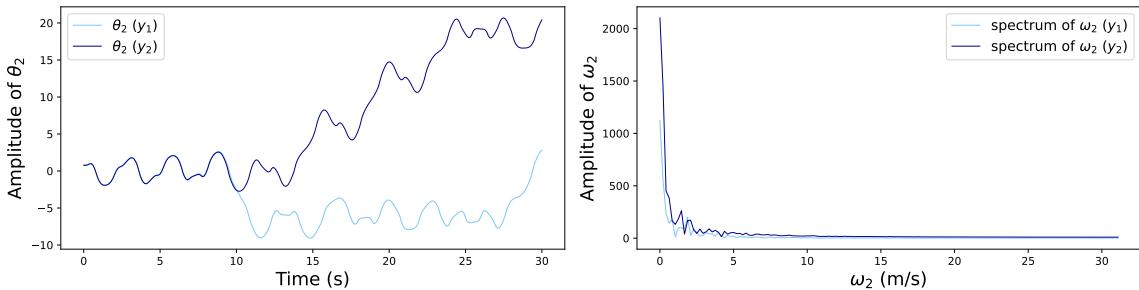
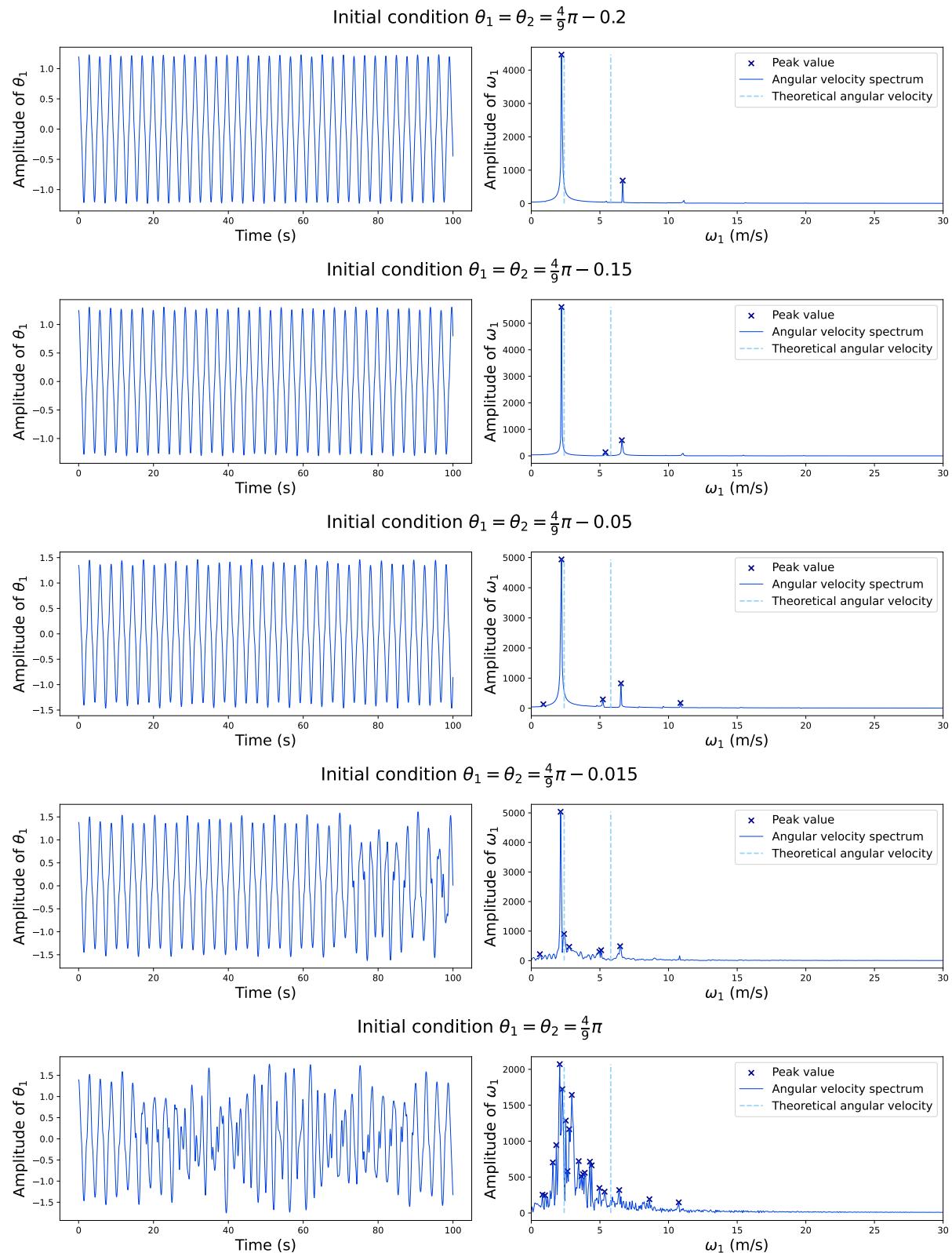


Figure 4.5: Amplitudes of θ_2 and ω_2 given two slightly different initial conditions

Transition cases

One way to examine the transition from non-chaotic to chaotic motion is to examine its angular velocity. The angular velocity spectrum having two peaks is a sufficient condition for us to say that the orbit is quasiperiodic. Chaotic motion will be carried out if a spectrum appears more ‘spiky’. As the point mass in the upper pendulum with angle θ_1 only swings circularly, we will analyse θ_1 and $\dot{\theta}_1$ to see the clear transition.

Example 4.2.4 (Symmetric angle cases). In this example, we will concentrate on symmetric initial conditions, i.e. $\theta_1 = \theta_2$. From Figure 4.6, it has been found that a change of motion behaviour occurs when θ is $\frac{4\pi}{9}$ in radians (80° in degree), which unsurprisingly coincides with Figure 4.13 in Section 4.5 (Lyapunov exponents) and Figure 4.11 in Section 4.6 (bifurcation diagram). We notice that rather than a sudden change in the angular velocity spectrum, the number of peaks increases gradually as the initial angle increases. If we look at the first column of the figure, we can spot a clear transition to chaos occurring in a smaller and smaller time interval.

Figure 4.6: A transition to chaos around $\frac{4\pi}{9}$

Circular cases

If the initial energy is large enough, we may expect a double pendulum to swing circularly indefinitely. This will occur if the pendulum is able to attain a large centripetal force that restricts it to circular motion.

Example 4.2.5. Given the initial conditions $(\theta_1, \theta_2, \omega_1, \omega_2) = (\pi, \pi, 8, 8)$ for a double pendulum, there is a series of zigzag lines for ‘wrapped-around’ θ_1 in Figure 4.7, illustrating jumps from π to $-\pi$ at uniform time intervals. By ‘wrapped-around’, we mean that we map θ_1 to its principal value. This is also supported by the bottom left graph, which shows θ_1 increasing linearly. The corresponding angular velocity domain suggests a peak at zero, which can be interpreted as the Fourier transformation producing only a constant term. Therefore, we can conclude that a double pendulum can undertake circular motion.

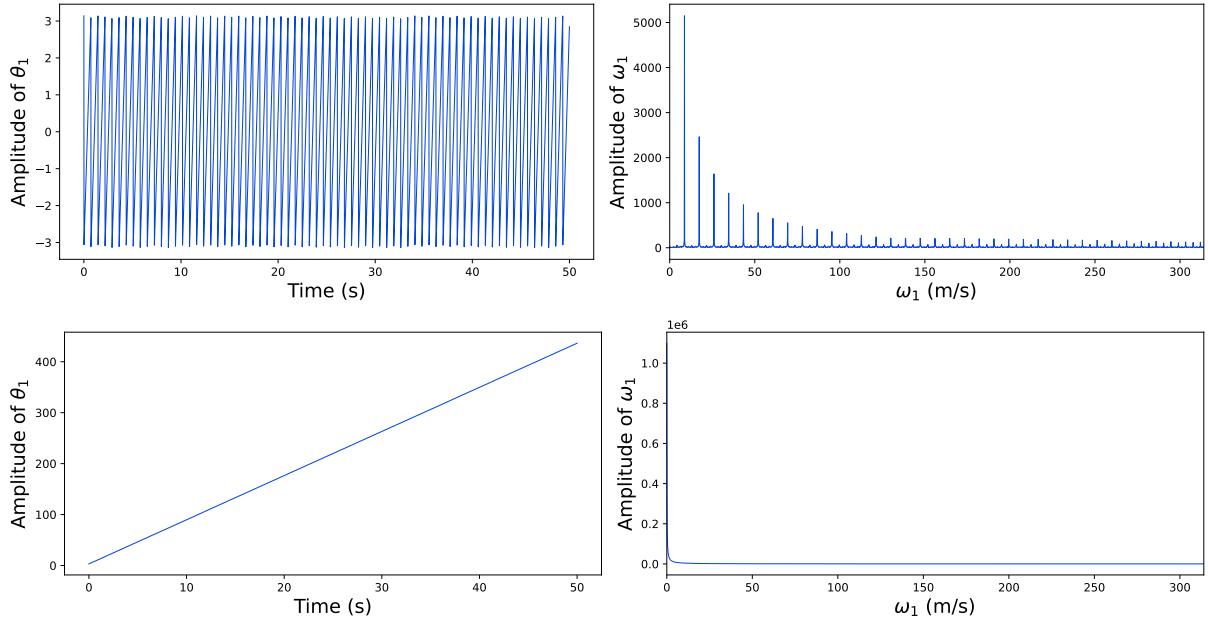


Figure 4.7: θ_1 with and without wrapped round and the corresponding ω_1 during 50s

4.2.2 Recurrence plots

Definition 4.2.1 (Recurrence plot (RP)[4]). Let $x_i \in \mathbb{R}^4$ be a trajectory in phase space $(\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2)$. A RP can be used to efficiently visualise recurrences. It can be formally expressed by the matrix

$$\mathbf{R}_{i,j}(\epsilon) = \Theta(\epsilon - \|x_i - x_j\|), \quad i, j = 1, \dots, N, \quad (4.1)$$

where N is the number of measured points x_i , ϵ is a threshold distance, $\Theta(\cdot)$ is the Heaviside function (i.e. $\Theta(x) = 0$ if $x < 0$, and $\Theta(x) = 1$ otherwise), and $\|\cdot\|$ is a norm.

In this case, the binary matrix \mathbf{R} has entry value 1 if and only if $\|x_i - x_j\| \leq \epsilon$ [5]. Therefore, the recurrence plot will tell us the pattern of similar states of the system[4]. A symmetrical \mathbf{R} with entry value 1 along the diagonal is expected since $d(x_i, x_j) = \|x_i - x_j\|$ is a metric with the properties of symmetry and indiscernibility of identicals. Thus, the recurrence plot is symmetric along $y = x$. Here we define $\|\cdot\|$ as the Euclidean norm and ϵ as 40% of the standard derivation of the signal to get an approximate optimal threshold based on [6].

Example 4.2.6 (Transition state of symmetric angle cases). *From Figure 4.8, we can clearly see that there is a regular pattern for the quasiperiodic case. By increasing the angle of inclination in initial conditions, we get more erratic and less predictable patterns, indicating the transition from quasiperiodic motion to chaotic motion, which can also be interpreted by Figure 4.6 where we use the FFT method.*

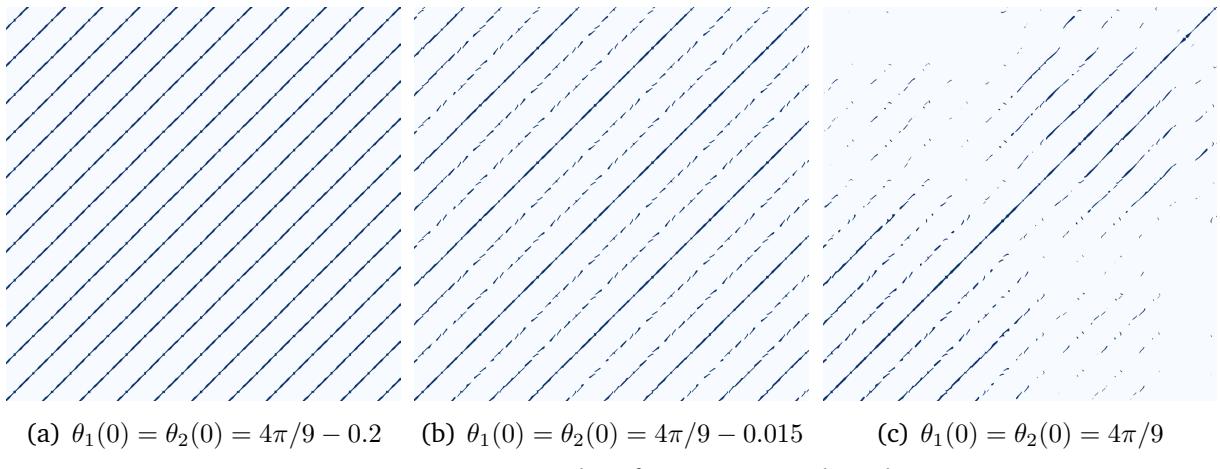


Figure 4.8: Recurrence plots for symmetrical angle cases

Recurrence quantification analysis

Recurrence quantification analysis (RQA) is a tool for quantifying RP[7]. RQA can be done using several different measures, all of which are based on recurrence density or the diagonal/vertical pattern of the RP[4]. Since methods using patterns of diagonal and vertical lines are based on the same idea, only the first two types will be introduced by giving an example for each.

1. Recurrence density

Definition 4.2.2 (Recurrence rate(RR)[4]). In the RP, the *recurrence rate* is a measure of the density of recurrence points, which is defined as

$$RR(\epsilon) = \frac{1}{N^2} \sum_{i,j=1}^N \mathbf{R}_{i,j}(\epsilon) \quad (4.2)$$

Therefore, the RR will illustrate the probability of a state lying in an ϵ -neighbourhood of one of the original states[4]. In other words, RR refers to a recurring probability. Therefore, we would expect the RR to be larger for quasiperiodic motion than for chaotic motion. Taking symmetric initial conditions $\theta_1(0) = \theta_2(0) = \theta$ as an example, a sudden drop can be found for $\theta \approx \frac{4\pi}{9}$ in Figure E.1, which agrees with the former analysis.

2. Diagonal lines

We leave this method to Appendix D as we obtain a less clear result due to the computational complexity.

4.3 Poincaré sections

4.3.1 Theory

Since it is difficult to visualise higher dimensional phase spaces for systems like the double pendulum, the idea of Poincaré sections is to reduce the dimension and concentrate on a specific 3-D section (or its 2-D projection). In other words, instead of studying a continuous flow, one could focus on a discrete map, which is useful for reducing dimension, analysing global dynamics and clarifying concepts[8]. For the double pendulum system, with a 4-D phase space $(\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2)$, Poincaré suggested using the cross section $\theta_1 = 0$ and $\dot{\theta}_1 > 0$ and storing θ_2 and $\dot{\theta}_2$, which means the motion of the lower pendulum will be recorded and analysed[3].

The method for determining Poincaré sections can be found in Algorithm 1.

4.3.2 Numerical results

In this section, we investigate some Poincaré sections to provide more qualitative ‘evidence’ for supporting some of the previous analysis and results. For a periodic orbit, we expect the Poincaré sections to resemble a finite number of points. We expect quasiperiodic orbits to only intersect with very small part of the section, and the sequence of points usually resembles a curve[9].

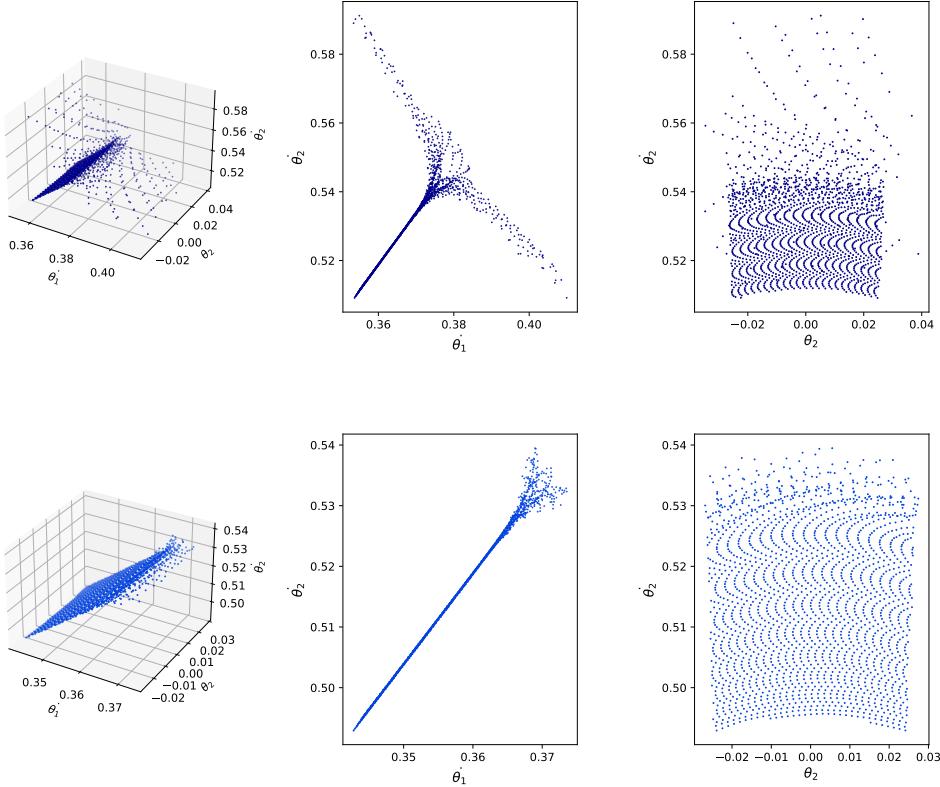
Example 4.3.1. Figure 4.9 uses Poincaré sections to check the linearisation result for approximating periodic behaviour. We consider the initial conditions $\theta_1 = \frac{\pi}{20}, \theta_2 = \sqrt{2} \times \frac{\pi}{20}, \dot{\theta}_1 = \dot{\theta}_2 = 0$. The second row shows Poincaré sections and projections with these initial angles, which seem like reasonable plots for approximating a single point. When the angles start in an initial ratio of 1.5 (first row) or 1.3 (third row) instead, the points start to span a larger region of the graph than we might expect for periodic behaviour.

Algorithm 1 Algorithm for determining Poincaré section

```

procedure POINCARÉ SECTION(initial conditions)
     $T \leftarrow 5000$ 
     $dt \leftarrow 0.01$ 
     $N \leftarrow T/dt$ 
     $x \leftarrow$  initial conditions
     $\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2 \leftarrow$  solution given  $x$ 
     $l_{12}, l_{21}, l_{22} \leftarrow$  empty lists            $\triangleright$  for storing qualified  $\dot{\theta}_1, \theta_2, \dot{\theta}_2$  respectively
    for  $i = 1$  to  $N - 1$  do
        if  $\theta_{1i} < 0$  and  $\theta_{1(i+1)} > 0$  then
            Append  $\dot{\theta}_{1i}$  to  $l_{12}$                    $\triangleright$  these are weighted averages
            Append  $\dot{\theta}_{2i}$  to  $l_{21}$                    $\triangleright$  using  $\theta_{1i}, \theta_{1(i+1)}$  as weights.
            Append  $\dot{\theta}_{2i}$  to  $l_{22}$ 
        end if
    end for
    plot graphs, such as  $l_{22}$  against  $l_{21}$ 
end procedure

```



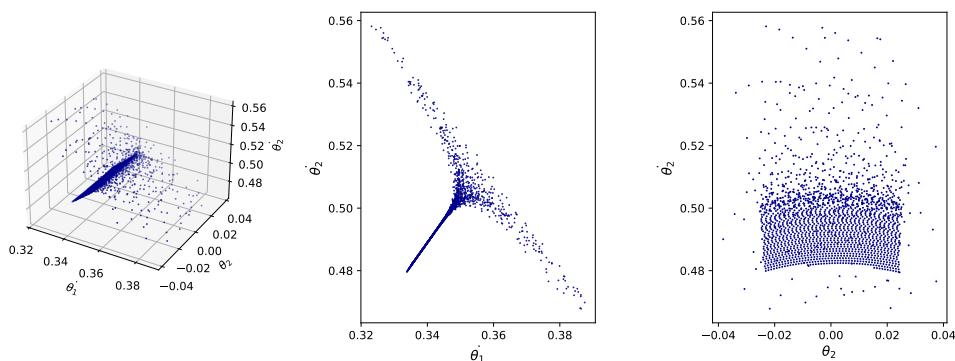
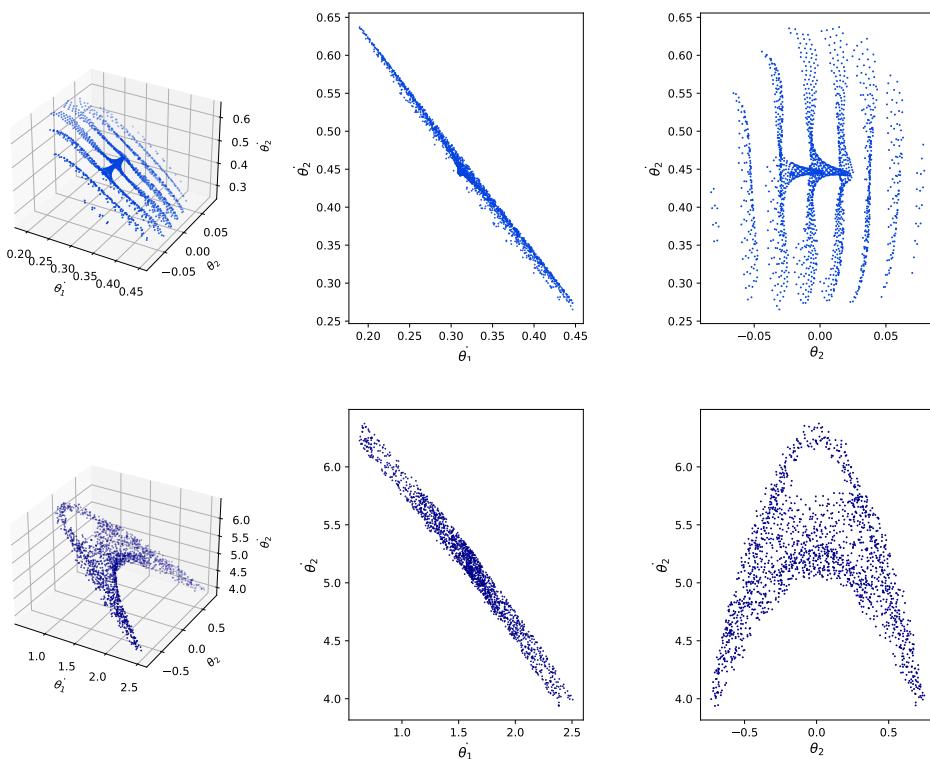


Figure 4.9: The initial conditions for these plots are of the form $\theta_1 = \frac{\pi}{20}, \theta_2 = \frac{c\pi}{20}$, where $c = 1.5, \sqrt{2}$ and 1.3 in the 1st to 3rd rows respectively.

Example 4.3.2. Figure 4.10 depicts Poincaré sections for equal initial angles, zero initial angular velocity. The second and third plots support the transition to chaos found at $\theta_1 = \theta_2 = \frac{4}{9}\pi$. We will later see that this is also supported by Figures 4.11 (bifurcation diagrams) and 4.12 (maximal Lyapunov exponents).



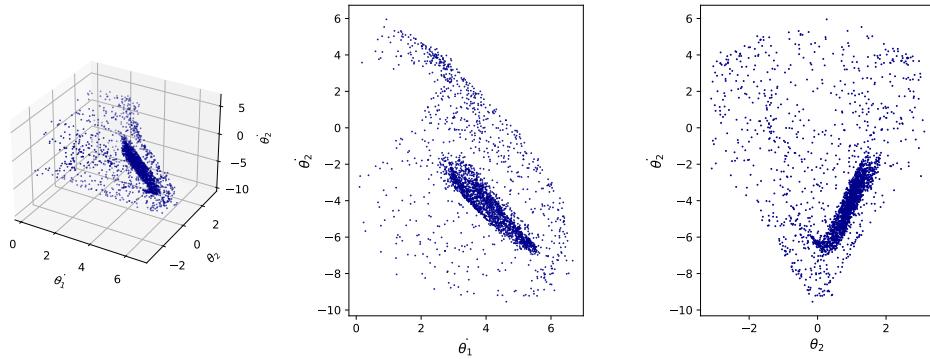


Figure 4.10: The initial conditions for these plots are: first row $\theta_1 = \theta_2 = \frac{\pi}{20}$, second row $\theta_1 = \theta_2 = \frac{4}{9}\pi \times 0.98$, and third row $\theta_1 = \theta_2 = \frac{4}{9}\pi$.

4.4 Bifurcation diagrams

4.4.1 Theory

Thanks to Poincaré sections, we can now focus on the distribution between two variables and develop further analysis of the transition between non-chaotic and chaotic motion using bifurcation diagrams.

Definition 4.4.1 (bifurcation[10]). The appearance of a topologically nonequivalent phase portrait under variation of parameters is called a *bifurcation*.

Definition 4.4.2 (bifurcation diagram[10]). A *bifurcation diagram* of the dynamical system is a stratification of its parameter space induced by the topological equivalence, together with representative phase portraits for each stratum.

Bifurcation diagrams visualise behaviour of a system by showing all possible values of a quantity given a certain parameter[10]. Thus the bifurcation theorem is a study of how topological structure varies as bifurcation parameters are varied[11]. By carefully choosing a parameter against the quantity that can distinguish quasiperiodic and chaotic motion, we will then depict the route to chaos[12]. It is reasonable to expect that trajectories in quasiperiodic motion will cross a line with same angular velocities, while trajectories in chaotic motion will never coincide with former trajectories.

4.4.2 Numerical results

From the theory above, if we plot all values of $\dot{\theta}_1$ with initial conditions having a small increment of angle each time, we would expect solid line region at the small angle part, as well as dense and discrete points in a large area at the large angle part[12].

Example 4.4.1 (Bifurcation diagrams with different initial conditions). *Figure 4.11 coincides with the theory and shows a clear transition around 80° for the first case, which is also supported by Figure 4.6 in Section 4.2 FFT method. Interestingly, in all figures, angles where a transition occurs also correspond to different points of the sharp spiky boundary in Figure 4.13 and Figure 4.16, illustrating a great transition in motion with relatively small angle change. It is worth noting that the first part of bifurcation diagram is roughly a straight line when initial conditions are $\frac{\theta_2}{\theta_1} = \pm\sqrt{2}$, which is reasonable since the trajectory is approximately periodic (shown in 3.5).*

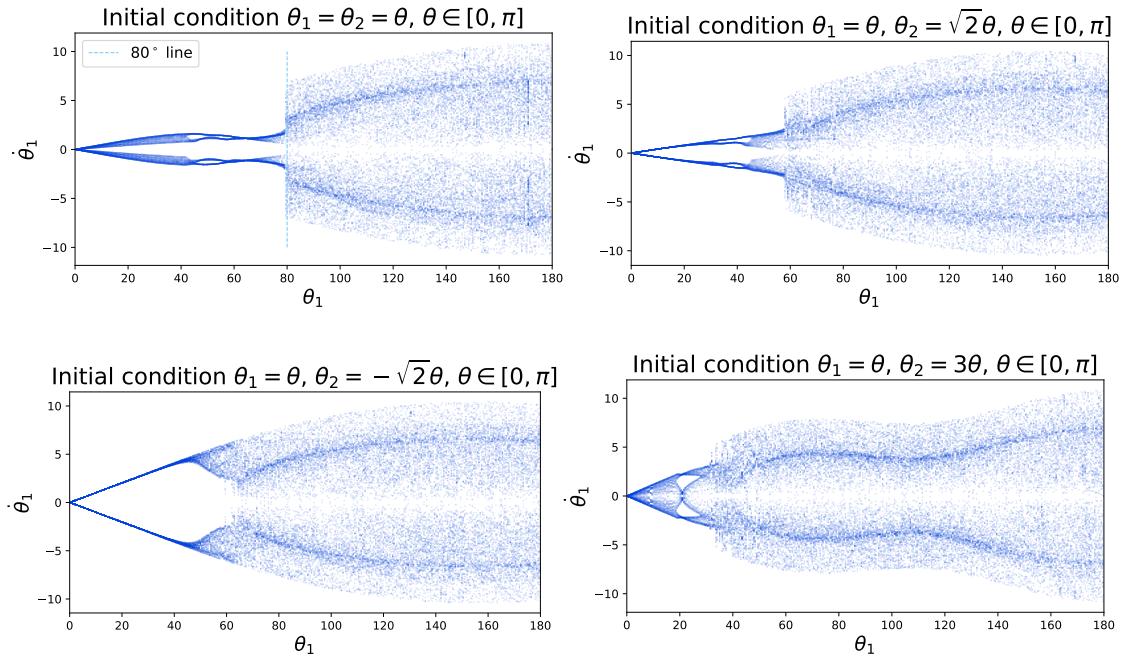


Figure 4.11: Bifurcation diagrams with different initial conditions

4.5 Lyapunov exponents

4.5.1 Theory

Another way of conducting sensitivity analysis is through considering Lyapunov exponents, as positive Lyapunov exponents usually indicate chaos [8], since it tells us the exponential rate of divergence between two infinitesimally close trajectories. The definition of a Lyapunov exponent can be found in Appendix C.

For the double pendulum problem, each initial condition has four local Lyapunov exponents, corresponding to the eigenvalues of the dimension 4 Jacobian matrix. For categorising non-chaotic and chaotic regimes, we consider the maximal Lyapunov exponent (MLE), i.e. the largest exponent in the Lyapunov spectrum.

Due to the exponential growth rate, for any orbit beginning from a perturbation with some component in the direction associated with the direction of the MLE, the impact of the MLE will dominate over the other exponents. Hence, one way to determine the MLE [13] is by considering the average:

$$\text{MLE} = \lim_{t \rightarrow \infty} \left(\lim_{\|\epsilon\| \rightarrow 0} \left(\frac{1}{t} \ln \frac{\|\phi(t, \mathbf{x}_0 + \epsilon) - \phi(t, \mathbf{x}_0)\|}{\|\epsilon\|} \right) \right) \quad (4.3)$$

where ϵ is the perturbation, \mathbf{x}_0 is the initial condition, and $\phi(t, x)$ denotes the flow after the system has evolved for length of time t . The algorithm for determining the MLE numerically can be found in Algorithm 2.

Since the (non-trivial) trajectories are continuous, do not terminate at a fixed point, and are bounded (by the conservation of energy), initial conditions leading to nonchaotic behaviour must have a zero Lyapunov exponent [14]. Here is a sketch proof: consider a perturbation in the direction of the flow at the point, say $\phi(t)$, for small t . By quasiperiodicity, any flow would need to return to be close to some point it has previously visited, i.e. there exists some τ such that $\phi(0)$ is close to $\phi(\tau)$ and $\phi(t)$ is close to $\phi(t + \tau)$. This means that $\|\phi(t + \tau) - \phi(\tau)\| \approx \|\phi(t) - \phi(0)\|$. Hence, on average, the separation between these two trajectories does not grow or shrink. Using this perturbation would lead to computing a 0 Lyapunov exponent. Therefore, Algorithm 2 will always return (roughly) at least zero, which can be seen in Figure 4.12. To find chaotic regimes, we therefore look for initial conditions with (strongly) positive maximal Lyapunov exponents. Moreover, the MLE gives us a quantitative measurement of the sensitivity [15].

Algorithm 2 Algorithm for determining maximal Lyapunov exponents

```

procedure MLE(initial conditions, perturbation)
     $T \leftarrow 1$ 
     $N \leftarrow 500$ 
     $x \leftarrow$  initial conditions
     $\epsilon \leftarrow$  perturbation
    for  $i = 1$  to  $N$  do
         $\epsilon \leftarrow \epsilon \times \frac{0.01}{\|\epsilon\|}$ 
         $x \leftarrow \phi(T, x)$ 
         $x' \leftarrow \phi(T, x + \epsilon)$ 
         $\epsilon \leftarrow \|x' - x\|$ 
         $\lambda_i \leftarrow \ln(\frac{\epsilon}{0.01})$ 
    end for
    return  $\frac{1}{N} \sum_{i=1}^N \lambda_i$ 
end procedure

```

4.5.2 Numerical results

In order to get an overall view of the maximal Lyapunov exponents for various initial conditions, we can create a colour plot. Due to the time complexity of the algorithm, we only ran the algorithm for $N=500$ to approximate the maximal Lyapunov exponent. There are some dangers of trying to determine Lyapunov exponents precisely using this method [15]. However by running the algorithm for selected initial conditions at varying values of N (please see Figure E.2 in the appendix), we find that the results are roughly consistent for distinguishing between non-chaotic (maximal Lyapunov exponent equals 0) and chaotic cases (MLE is positive).

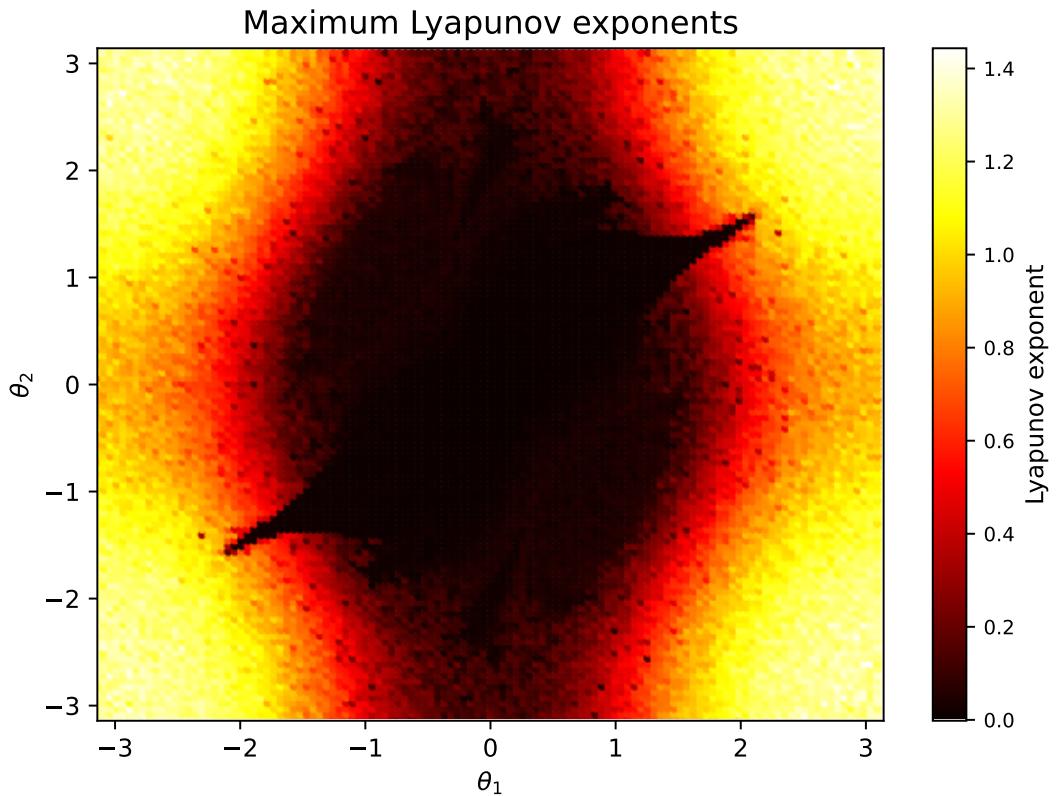


Figure 4.12: Colour plot depicting the maximal Lyapunov exponent for initial conditions with zero angular velocity.

The black shape (where the maximal Lyapunov exponent should be nonpositive) of this plot corresponds to where we expect non-chaotic regimes. We find that this shape is also present when we plot colour plots under different metrics used to evaluate chaoticity, such as in the ‘time to flip’ heatmap Figure 4.16 and the auto-correlation heatmap Figure 4.20.

Example 4.5.1. In Figure 4.13, we see that for different initial angle ratios, the MLE be-

comes (strongly) positive for different values of θ_1 . We see these values are in agreement with the bifurcation diagrams in Figure 4.11.

MLEs may also suggest how abruptly the transition from quasiperiodic to chaotic behaviour occurs. Figure E.3 in Appendix E depicts a relatively quick transition from quasiperiodic to chaotic orbits around the initial conditions corresponding to the steep region of purple line in Figure 4.13. We can compare this to systems starting from initial conditions lying on the light blue line, shown in Figure E.4 in Appendix E, where the transition to chaos seems more gradual.

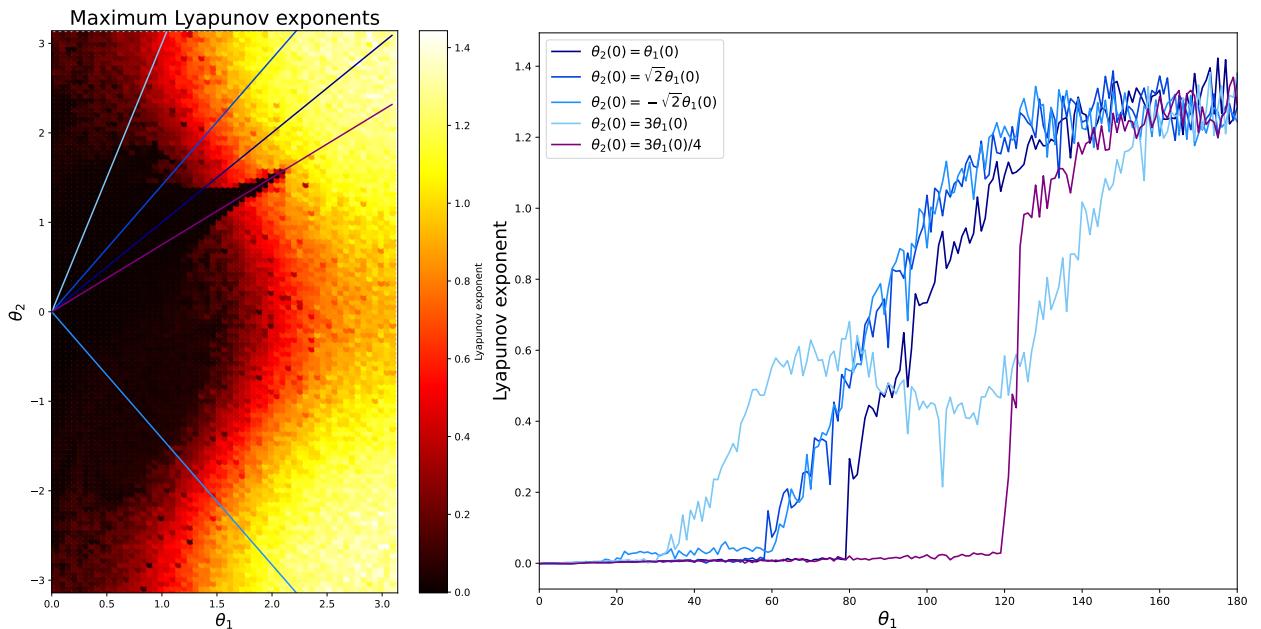


Figure 4.13: Maximal Lyapunov exponents for certain initial angle ratios

4.6 Sensitive dependence globally

While the previous section predominantly focused on results that allow us to look at specific initial conditions, this section will focus on the bigger picture of the double pendulum. We will assume throughout this section that the lengths of both pendulums and the mass of both bobs are 1. Unless stated otherwise, we also set the initial angular velocities to 0.

4.6.1 Time for error to double

The first metric we will use to look at sensitivity is the 'time it take for the error to double'. Mathematically, we define this as follows:

1. We define two pendulums p_1 and p_2 represented by $(\theta_1(t), \theta_2(t), \dot{\theta}_1(t), \dot{\theta}_2(t))$ and $(\varphi_1(t), \varphi_2(t), \dot{\varphi}_1(t), \dot{\varphi}_2(t))$ with initial conditions $\varphi_1(0) = \theta_1(0) + \delta$, $\varphi_2(0) = \theta_2(0) + \delta$, and $\dot{\theta}_1(0) = \dot{\theta}_2(0) = \dot{\varphi}_1(0) = \dot{\varphi}_2(0) = 0$ where δ is chosen beforehand (we use $\frac{\pi}{60}$ in this paper).
2. We plot the first time unit, $T \in \mathbb{R}_{>0}$, such that $d_\infty \left(\begin{bmatrix} \theta_1(T) \\ \theta_2(T) \end{bmatrix}, \begin{bmatrix} \varphi_1(T) \\ \varphi_2(T) \end{bmatrix} \right) \geq 2\delta = \epsilon$.
Note: $d_\infty \left(\begin{bmatrix} \theta_1(T) \\ \theta_2(T) \end{bmatrix}, \begin{bmatrix} \varphi_1(T) \\ \varphi_2(T) \end{bmatrix} \right) = \max(|\theta_1 - \varphi_1|, |\theta_2 - \varphi_2|)$, and one time unit is defined as $\sqrt{\frac{L}{g}}$ seconds where L is the common length of l_1 and l_2 . In this case $L = 1$.

To produce the colour-plot below, we investigate the system for 100 time units, and sample at equally spaced points separated by a distance of $\frac{\pi}{60}$. Any initial condition that did not diverge by ϵ within the 100 time units is plotted as white.

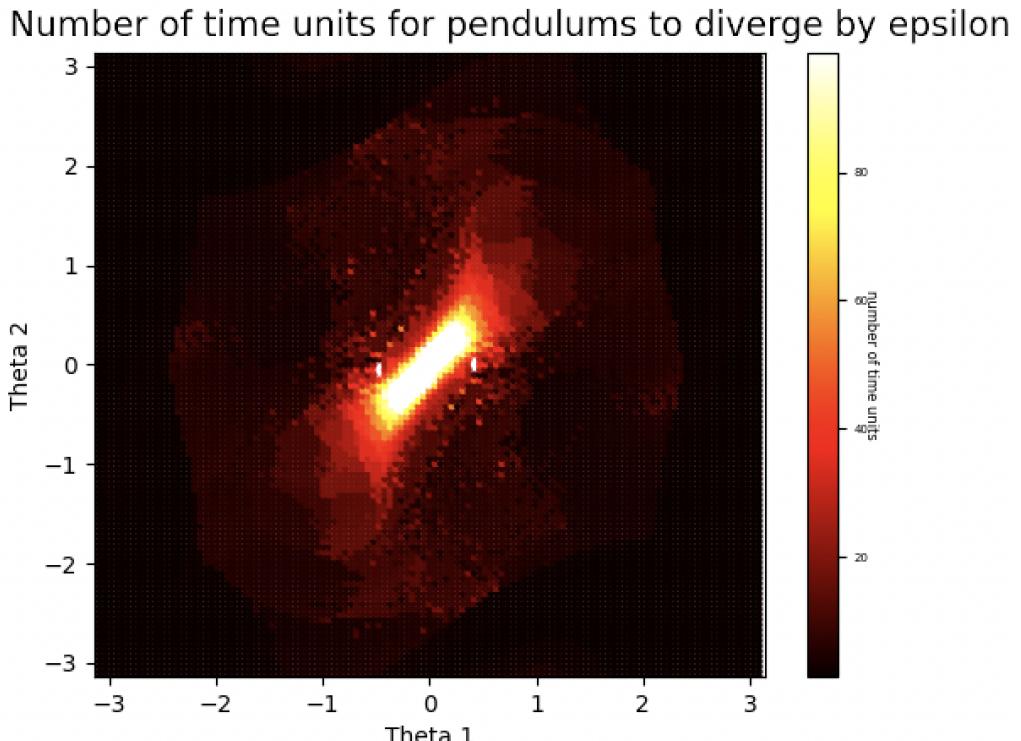


Figure 4.14: Colour plot depicting number of time units it takes for the initial difference to double.

Interestingly, the plot depicts a white rectangle in the centre, where the difference between the pendulums isn't doubled. Outside the rectangle, the number of time units taken to diverge decreases dramatically. If we create an animation of the values that θ_1 and θ_2 take for a given small initial condition, the shape resembles this rectangle, which

suggests that there may be a positively invariant set centred around $(0, 0)$ (see footer for the animation for initial conditions $\dot{\theta}_1 = \dot{\theta}_2 = \frac{\pi}{20}$)¹. From the animation, we can see that the rectangle is filled in a ‘regular pattern’, which is to be expected as we are not under chaotic motion, but also it seems that the whole rectangle gets filled over time. For example, this is what the graph looks like after 500 seconds:

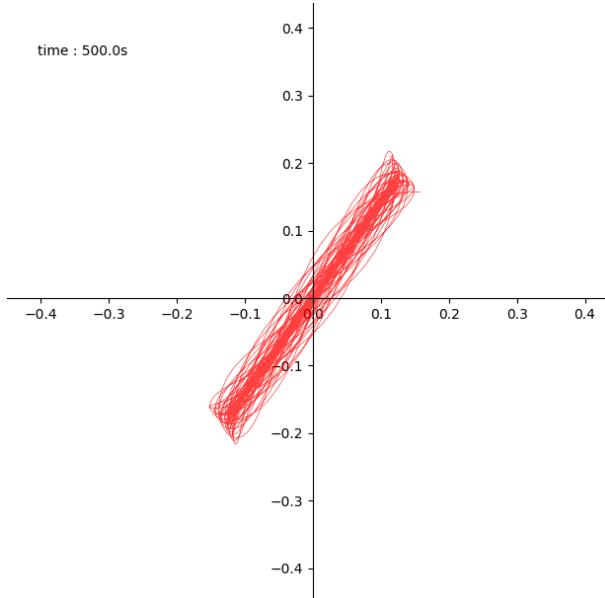


Figure 4.15: Plot of θ_1 against θ_2 after 500 seconds

This is a consequence of the Poincaré recurrence theorem, which states that for certain dynamical systems, the system will return arbitrarily close to its initial conditions in finite time [16]. See Appendix F, where we write a full description of the Poincaré recurrence theorem and why it is applicable to our double pendulum system. It can also be seen that if we use a photo-editing program to draw the lines of symmetry of the rectangle, the lines are approximately $y = x$ and $y = -x$. The line of symmetry across $y = -x$ is particularly interesting as it is the shorter line of symmetry for the rectangle, suggesting that sensitivity to initial conditions varies the most along this line. 4.12 corroborates this, as $y = -x$ is also approximately the shortest line joining the two parallel sides of the central black parallelogram-like shape.

4.6.2 Time for pendulum to flip

Another interesting lens through which to look at sensitive dependence is to investigate how long it takes for a pendulum to flip given various initial conditions. A flip is defined as when either $|\theta_1|$ or $|\theta_2|$ increases past π . The colour plot represents the number of time units it takes for a flip to take place, and white regions represent initial conditions where neither mass flipped in the allotted 150 time units. $\dot{\theta}_1 = \dot{\theta}_2 = 0$ [17]:

¹<https://drive.google.com/file/d/1d2IkIaEaM3JhqoQCh39W8HYhlxmXtHVx/view>

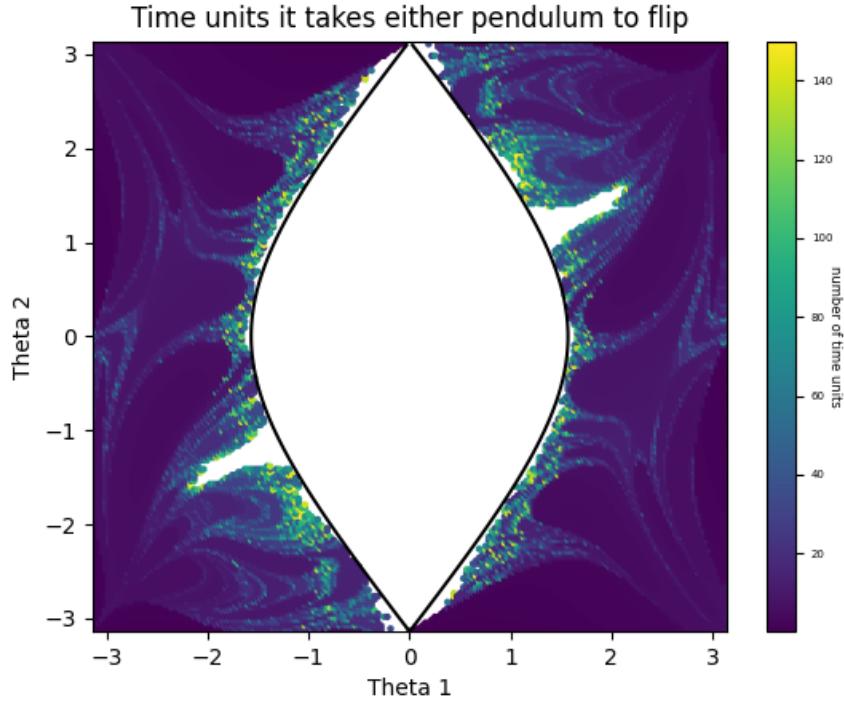


Figure 4.16: Colour plot depicting number of time units it takes for either pendulum to flip. White represents initial conditions where neither pendulum flipped.

The shape enclosed by the black line are initial conditions where it is energetically impossible for the pendulum to flip. The total energy in the system is

$$\dot{\theta}_1^2 + \cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 + \frac{1}{2}\dot{\theta}_2^2 - 2g \cos(\theta_1) - g \cos(\theta_2).$$

From physical intuition, it is clear that it takes less energy for the bottom mass to flip than it takes for the top mass to flip. This is due to the fact that in order for the top mass to flip, the bottom mass has to be raised to a certain level as well. Thus, for the pendulum to ‘barely flip’, we have $\theta_2 = 0$ and $\dot{\theta}_1 = \dot{\theta}_2 = 0$. So, the total energy is then $-2g \cos(\theta_1) + g$ which is at least $-g$. So, the system won’t flip if the initial total energy, which in our case is $-g \cos(\theta_1) - g \cos(\theta_2)$, is less than $-g$. Thus the region where it is impossible to flip is $2 \cos(\theta_1) + \cos(\theta_2) > 1$.

Interestingly, the boundary of the white region is fractal in nature, due to the two branches either side of the middle ellipsoid-esque shape. We can find the fractal dimension of one of the branches by using the box counting algorithm.

Definition 4.6.1 (Box-Counting Dimension[18]). Given a set S in the Euclidean space \mathbb{R}^n , we define the *Box-Counting Dimension* as $D_b = \lim_{r \rightarrow 0} \frac{\log N(r)}{\log \frac{1}{r}}$ where $N(r)$ is the number of boxes of side length r needed to cover the set S .

First, we need to get a relatively high resolution of one of the branches. Without loss of generality, we chose the right one, and create a $2^9 \times 2^9$ pixel representation of it:

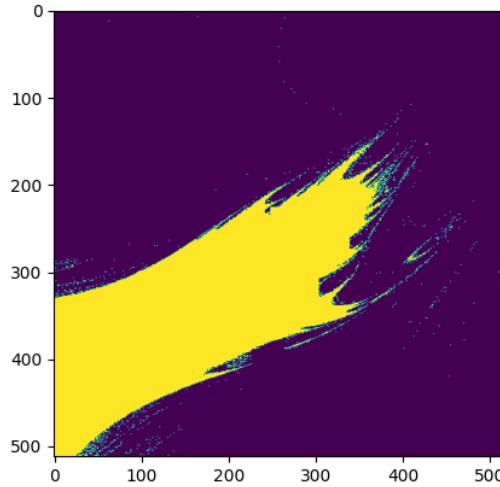


Figure 4.17: Right branch of the time to flip graph.

Then, to estimate the box-counting dimension of the boundary, we split the image into boxes and then take $N(r)$ to be the number of boxes that contain the boundary of the shape. We took the box side lengths (in pixels) to be the powers of 2 less than 2^9 . Then, we plotted a graph of $\log \frac{1}{r}$ against $\log N(r)$:

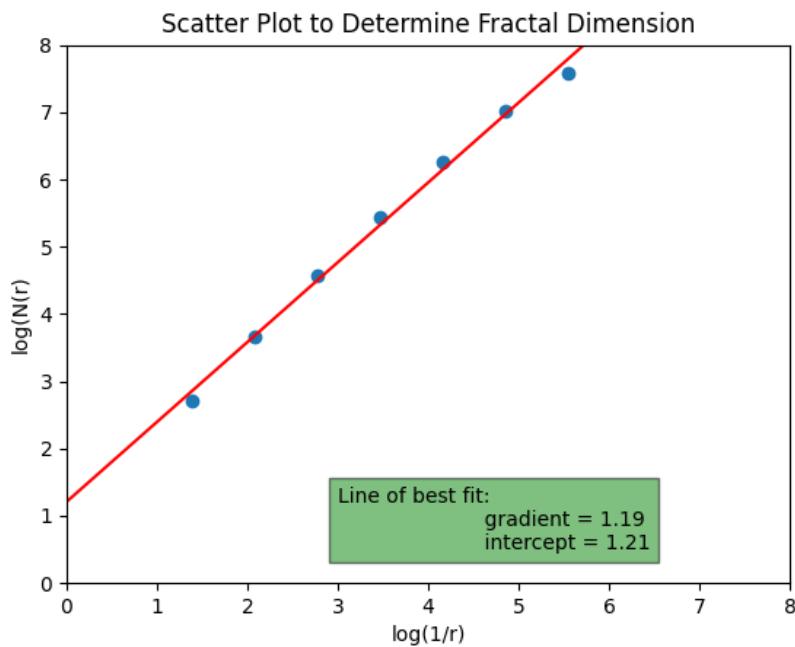


Figure 4.18: Plot of $\log(1/r)$ against $\log(N(r))$

We can see that the plots very closely follow a straight line. So, we can estimate the box-counting dimension by finding the gradient of the line of best fit. This works out to be 1.19. Since 1.19 is not close to an integer, there is evidence to suggest that the boundary of our graph is indeed fractal.

4.6.3 Auto-Correlation Coefficients

The Auto-Correlation coefficient provides us with an alternative to study the periodicity and also the chaoticness of our system.

Definition 4.6.2 (Auto-Correlation Coefficient). The Auto-Correlation Coefficient is the correlation between a time series and a delayed version of itself, where the amount of delay is known as the *lag*.

Mathematically, The Auto-Correlation Function (ACF) is defined as

$$Q_k = \frac{\text{Cov}(e_t, e_{t+k})}{\sqrt{\text{Var}(e_t)} \sqrt{\text{Var}(e_{t+k})}}$$

where e is our time series, t is the time step, and k is the lag. However, since our time series is stationary (i.e. there is no dependence on time, which is clear as we only require $\theta_1, \theta_2, \dot{\theta}_1$ and $\dot{\theta}_2$ to be equal for two pendulums to have the same motion), the formula can be reduced to

$$Q_k = \frac{\text{Cov}(e_t, e_{t+k})}{\sigma^2}$$

where σ^2 is the variance of the time series.

Q_k takes values between $[-1, 1]$ because of the Cauchy–Schwarz inequality. In practice, the actual covariance and variance will be unknown, so we have to estimate it. We can use the estimator

$$\hat{R}_k = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2}$$

where t is the time step, k is the lag, \bar{x} is the sample mean.

Lemma 4.6.1 (Consistency). \hat{R}_k is a consistent estimator for Q_k

Proof. We can write the estimator as

$$\frac{n-k}{n-1} \cdot \frac{\frac{1}{n-k} \sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\frac{1}{n-1} \sum_{t=1}^n (x_t - \bar{x})^2}$$

Now, the numerator of our estimator is known to be consistent for $\text{Cov}(e_t, e_{t+k})$ [19]. The denominator is the sample variance, which is also known to be a consistent estimator for σ^2 [20]. Furthermore, $\lim_{n \rightarrow \infty} \frac{n-k}{n-1} = 1$. Thus by Slutsky's theorem [21], we must have that $\hat{R}_k \xrightarrow{p} Q_k$ and hence by definition the estimator is consistent. ■

An important property of the ACF is that it maps periodic functions to periodic functions.

Lemma 4.6.2 (Periodic Invariance). *The ACF of a periodic function is also periodic, with the same period.*

Proof. Lets let our time series e repeat itself every m time steps. Thus, $e_t = e_{t+km}$ for $k \in \mathbb{Z}$. Take an arbitrary r . Then

$$\begin{aligned} Q_r &= \frac{\text{Cov}(e_t, e_{t+r})}{\sqrt{\text{Var}(e_t)} \sqrt{\text{Var}(e_{t+r})}} \\ Q_{r+km} &= \frac{\text{Cov}(e_{t+km}, e_{t+r+km})}{\sqrt{\text{Var}(e_{t+km})} \sqrt{\text{Var}(e_{t+r+km})}} \\ &= \frac{\text{Cov}(e_t, e_{t+r})}{\sqrt{\text{Var}(e_t)} \sqrt{\text{Var}(e_{t+r})}} \\ &= Q_r \end{aligned}$$

■

Because of this property, we can use the ACF as another tool to inspect the periodicity and chaoticness of our system. If the system possesses periodicity or non-chaoticness based on the initial conditions we choose, we would expect the ACF to also be periodic and for the coefficients to not tend to 0, as neighbouring points will stay close. The maximal Lyapunov exponent for these cases should be close to 0. On the other hand, if our system is chaotic we would expect nearby points in the phase space to eventually move far apart, and hence the ACF should tend to 0. The maximal Lyapunov exponent for these cases should be relatively large.

Before looking at plots of the ACF, it is worthwhile to define white-noise and the Bartlett test.

Definition 4.6.3 (White noise). [22] A time series is deemed to be white noise if the variables at each time step are independent and identically distributed, and have zero mean. In this case $Q_k = 0$ as the covariance between the time series and a lagged version of itself will be 0.

Definition 4.6.4 (Bartlett Test). [23] Bartlett (1946) demonstrated that when we have a series generated by a white noise process, the estimators \hat{R}_k will approximately follow a $N(0, \frac{1}{M})$ distribution where M is the number of samples in our time series.

Below we create a stem plot for the ACF for a variety of initial conditions with respect to the angle θ_2 for the bottom mass. Before computing the ACF, we assign each angle to its principal value as to reduce the range to $[-\pi, \pi)$. We look at lags from 1 to 100 (as we are using a time step of 0.05 when calculating the solution this corresponds to lags from 0.05 to 5 seconds), and the first 100 seconds of the solution.

We use the Bartlett test to determine if the auto-correlation coefficients are statistically different from 0. We use a significance level of 0.01. In our case, we have 2001 samples so our coefficients follow a $N(0, \frac{1}{2001})$ distribution. Thus, the coefficient is statistically significant if its magnitude is greater than 0.0576. We represent this threshold on the graphs with a black dashed line.

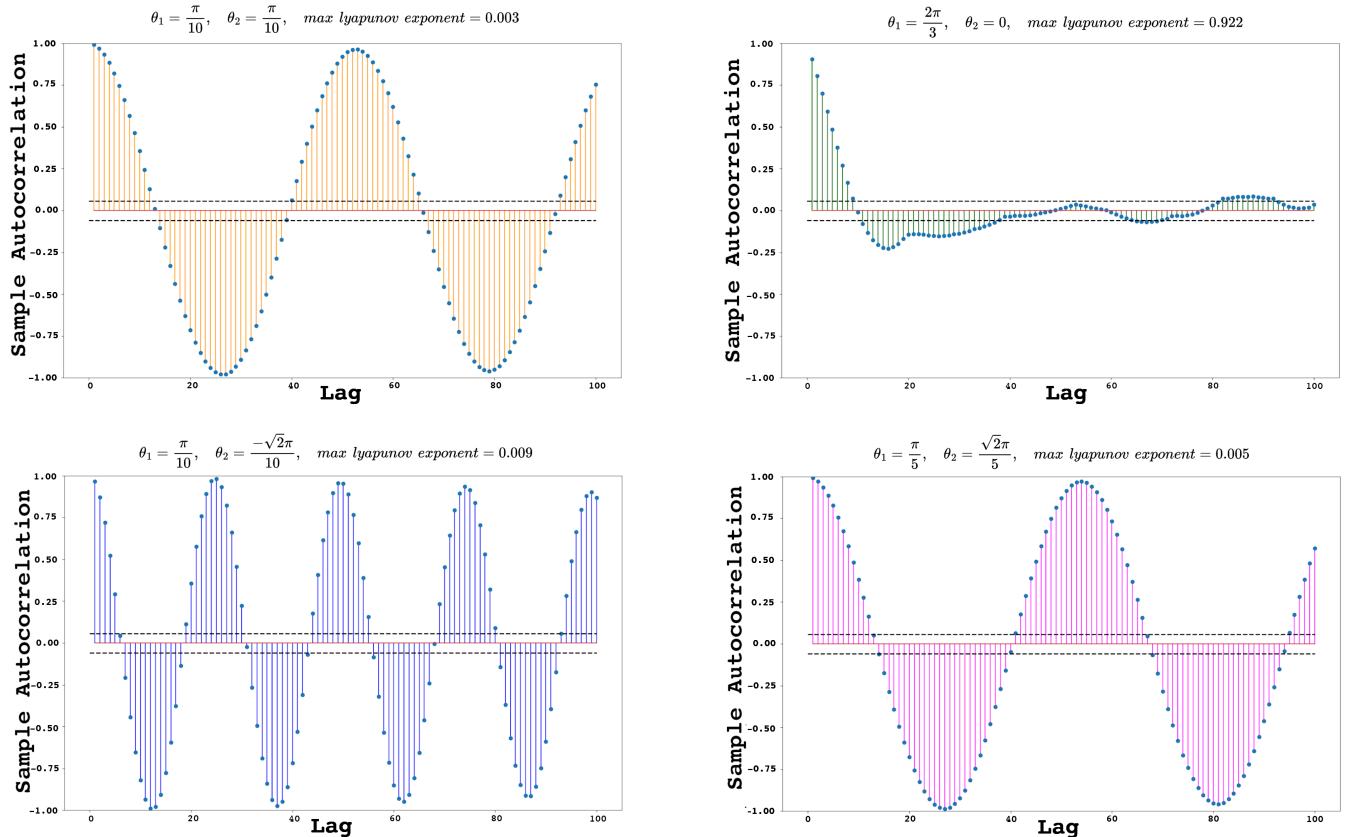


Figure 4.19: Stem Plots for the ACF for a variety of initial conditions.

As we can see, the initial conditions close to the stable equilibrium are indeed periodic and do not tend to 0. The bottom two plots represent the ACF for the normal modes of motion as discussed in 3.3.1. The top right plot tends to 0, which is expected for initial conditions far from $(0, 0)$.

Below we also make a plot of the fraction of coefficients within the 0.0576 threshold. We look at the first 200 lags, and because small lags generally have a higher degree of auto-correlation, we start from a lag of 50.

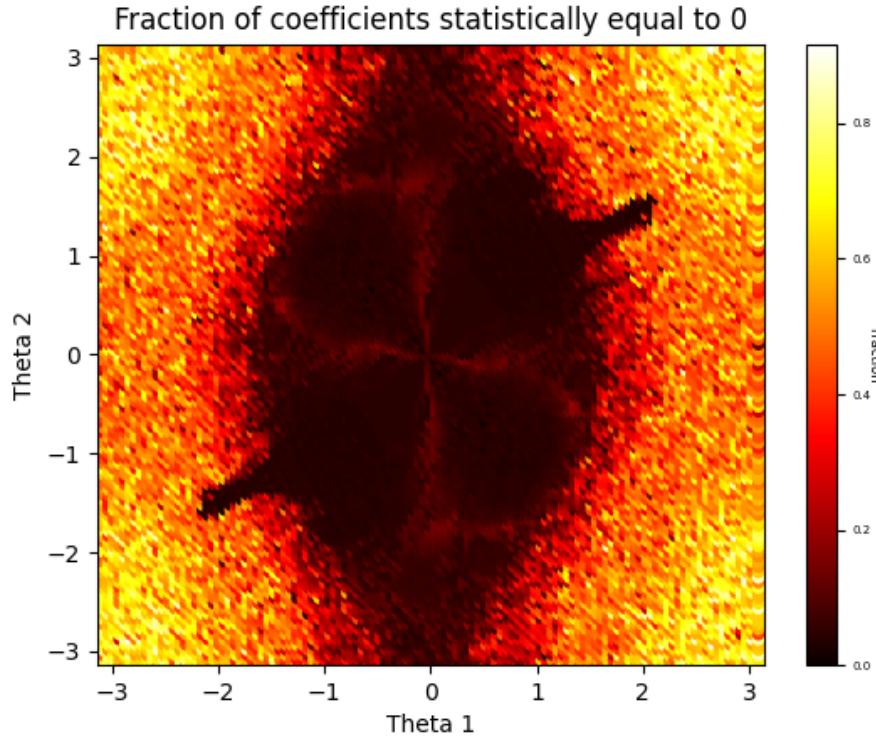


Figure 4.20: Fraction of lags within the 99% confidence interval threshold.

From the graph, we can see a clear central region where the fraction is near to 0 and the double pendulum is likely to not be chaotic. Outside the region, the fraction quickly increases to be around 0.8, indicating a sudden transition to chaos.

Surprisingly this graph, which was generated from statistical results, is extremely similar to the graph of how long it takes for either mass to flip (4.16), a purely mechanical result. Thus, the ‘branches’ of this graph are likely to be fractal too. This connection could perhaps be investigated further if more time was available.

While the Bartlett test is useful to determine whether individual coefficients are statistically equal to 0, we need a new test if we wish to look at multiple coefficients simultaneously. Hence, we introduce the Ljung-Box test.

Definition 4.6.5 (Ljung-Box test). [24] The Ljung-Box test asserts that

$$L_m = n(n+2) \sum_{k=1}^m \frac{\hat{R}_k^2}{n-k} \sim \chi_m^2$$

where n is the number of samples in our time series and $m \in \mathbb{Z}_{>0}$, under the null hypothesis that the samples in our time series are independent.

We carry out a hypothesis test with a significance level of 0.01 and use Bonferroni correction [25] to create a joint test for the auto-correlation coefficients of θ_1 and θ_2 . Although Bonferroni correction is conservative and thus the test will not have a relatively high power. However, other methods, such as those used in [26] where it is proved that for $X \sim \chi_m^2$ and $Y \sim \chi_m^2$ with correlation ρ , $X + Y$ approximately follows $\Gamma(\frac{m}{1+\rho}, 2(1+\rho))$, are more computationally intensive.

So, we reject H_0 if either of the test statistics for θ_1 or θ_2 are greater than c where $P(\chi_m^2 > c) = 0.005$. In our case, we use the first 100 lags so $m = 100$ and $c = 140.17$.

Below is a log plot for the smallest of the two test statistics for a variety of initial conditions, and the red plane represents the value of c .

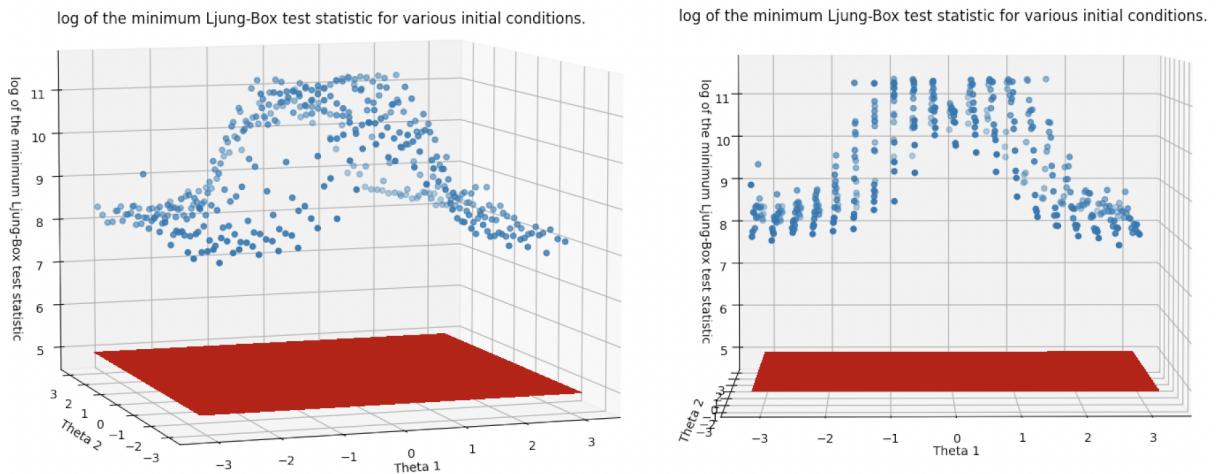


Figure 4.21: Ljung-Box test statistic for a variety of initial conditions

As we can see from the graph, every point we sampled from was quite above the c -threshold. Therefore, there is a high likelihood that the values at each time step are not independent. This demonstrates that even though the double pendulum may be chaotic at various initial conditions, its motion does not resemble white noise and is not random. Furthermore, we can see that the correlations are roughly normally distributed, which is what we would expect because of the Bartlett Test and the Delta method [27].

Chapter 5

Conclusion

In this report, we have found periodic, quasiperiodic, chaotic and circular regimes of a double pendulum system with unit masses and rod lengths. Linearisation around the stable fixed points suggested that for small angles, the two masses behave according to the sum of two cosine functions. This result is then substantiated by a variety of numerical tools commonly used to analyse dynamical systems. Taking a more global view, we adopted various ways of characterising quasiperiodic and chaotic behaviour, and found that the various methods all reinforce each other. For example, the same shape corresponding to nonchaotic behaviour is present in the heatmaps generated by different ways of quantifying sensitivity. Further analysis on one of the ‘branches’ of this shape shows that it has fractal dimension 1.19, and so is likely to be fractal. We also find that some transitions from non-chaos to chaos are more abrupt than others.

There are still many aspects of the problem that could be explored, such as varying the parameters, which have been taken to be unit in our study and doing further investigations into varying initial angular velocities.

Bibliography

- [1] Jyotirmoy Roy, A. Mallik, and Jayanta Bhattacharjee. Role of initial conditions in the dynamics of a double pendulum at low energies. *Nonlinear Dynamics*, 73, 2013. DOI: [10.1007/s11071-013-0848-1](https://doi.org/10.1007/s11071-013-0848-1). pages 9, 16
- [2] Ivana Kovacic, Miodrag Zukovic, and Dragi Radomirovic. Normal modes of a double pendulum at low energy levels. *Nonlinear dynamics*, 99(3), 2020. DOI: [10.1007/s11071-019-05424-5](https://doi.org/10.1007/s11071-019-05424-5). pages 11, 15
- [3] Louis N. Hand. *Analytical mechanics*. Cambridge University Press, Cambridge, 1998. pages 15, 22
- [4] Norbert Marwan, M Carmen Romano, Marco Thiel, and Jürgen Kurths. Recurrence plots for the analysis of complex systems. *Physics reports*, 438(5):237–329, 2007. DOI: [10.1016/j.physrep.2006.11.001](https://doi.org/10.1016/j.physrep.2006.11.001). pages 20, 21, 22, 49
- [5] Norbert Marwan, Stefan Schinkel, and Jürgen Kurths. Recurrence plots 25 years later -gaining confidence in dynamical transitions. *Europhysics letters*, 101(2):20007–, 2013. DOI: [10.1209/0295-5075/101/20007](https://doi.org/10.1209/0295-5075/101/20007). pages 21
- [6] S Schinkel, O Dimigen, and N Marwan. Selection of recurrence threshold for signal detection. *The European physical journal. ST, Special topics*, 164(1):45–53, 2008. DOI: [10.1140/epjst/e2008-00833-5](https://doi.org/10.1140/epjst/e2008-00833-5). pages 21
- [7] Norbert Marwan, Niels Wessel, Udo Meyerfeldt, Alexander Schirdewan, and Jürgen Kurths. Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Phys. Rev. E*, 66:026702, Aug 2002. DOI: [10.1103/PhysRevE.66.026702](https://doi.org/10.1103/PhysRevE.66.026702). pages 21
- [8] Stephen Wiggins. *Introduction to applied nonlinear dynamical systems and chaos*. Texts in applied mathematics ; 2. Springer, New York, 2nd edition, 2003. pages 22, 26, 48
- [9] F.C. Moon. Nonlinear dynamics. In Robert A. Meyers, editor, *Encyclopedia of Physical Science and Technology (Third Edition)*, pages 523–535. Academic Press, New York, 3rd edition, 2003. DOI: [10.1016/B0-12-227410-5/00484-1](https://doi.org/10.1016/B0-12-227410-5/00484-1). pages 22

- [10] IU. A. (IUrii Aleksandrovich) Kuznetsov. *Elements of applied bifurcation theory*. Applied Mathematical Sciences ; 112. Springer, New York, 3rd edition, 2004. pages 25
- [11] Paul. Glendinning. *Stability, instability, and chaos : an introduction to the theory of nonlinear differential equations*. Cambridge texts in applied mathematics. Cambridge University Press, Cambridge, 1994. pages 25
- [12] A M Calvão and T J P Penna. The double pendulum: a numerical study. *European Journal of Physics*, 36(4):045018, May 2015. DOI: [10.1088/0143-0807/36/4/045018](https://doi.org/10.1088/0143-0807/36/4/045018). pages 25
- [13] Keying Guan. Important notes on lyapunov exponents. 2014. DOI: [10.48550/ARXIV.1401.3315](https://arxiv.org/abs/1401.3315). pages 27
- [14] H. Haken. At least one lyapunov exponent vanishes if the trajectory of an attractor does not contain a fixed point. *Physics Letters A*, 94(2):71–72, 1983. DOI: [10.1016/0375-9601\(83\)90209-8](https://doi.org/10.1016/0375-9601(83)90209-8). pages 27
- [15] Joe Chen. Chaos from simplicity: An introduction to the double pendulum. https://ir.canterbury.ac.nz/bitstream/handle/10092/12659/chen_2008_report.pdf, 2008. Student assignment. pages 27, 28
- [16] Tomoo Yokoyama. Generalization of poincaré recurrence theorem for flows on surfaces and characterization of minimal flows on compact surfaces - arxiv.org. <https://arxiv.org/pdf/2110.05705.pdf>, Feb 2022. pages 31, 55
- [17] Jeremy S Heyl. The double pendulum fractal. <https://www.famaf.unc.edu.ar/~vmarconi/fiscomp/Double.pdf>, Aug 2008. pages 31
- [18] Jiaxin Wu, Xin Jin, Shuo Mi, and Jinbo Tang. An effective method to compute the box-counting dimension based on the mathematical definition and intervals. *Results in Engineering*, 6:100106, 2020. DOI: [10.1016/j.rineng.2020.100106](https://doi.org/10.1016/j.rineng.2020.100106). pages 32
- [19] Lie Wang. Estimation of covariance matrix. <https://math.mit.edu/~liewang/ECM.pdf>. pages 34
- [20] Tom Kennedy. Properties of point estimators and finding them. https://www.math.arizona.edu/~tgk/466_s18/chap9.pdf, 2018. pages 34
- [21] Jun Shao. Lecture 21: Convergence of transformations and generating a random variable. <https://pages.cs.wisc.edu/~shao/stat609/stat609-21.pdf>, 2015. pages 34

- [22] Ruey S. Tsay. *Analysis of financial time series*. Wiley series in probability and statistics. Wiley, Hoboken, N.J, 3rd edition, 2010. pages 35
- [23] Richard M. Levich and Rosario C. Rizzo. Alternative tests for time series dependence based on autocorrelation coefficients. <https://pages.stern.nyu.edu/~rlevich/wp/LR1.pdf>, Dec 1998. pages 35
- [24] Xisen Wang. White noise tests on the lstm model trained with double pendulum. <https://iopscience.iop.org/article/10.1088/1742-6596/2137/1/012032/pdf>, 2021. pages 37, 57
- [25] Megan Goldman. Why is multiple testing a problem? <https://www.stat.berkeley.edu/~mgoldman/Section0402.pdf>, 2008. pages 38
- [26] Alberto Ferrari. A note on sum and difference of correlated chi-squared variables. <https://arxiv.org/pdf/1906.09982.pdf>, Jun 2019. pages 38
- [27] Reg Kulperger. Delta method. <http://fisher.stats.uwo.ca/faculty/kulperger/SS3858/Handouts/DeltaMethod.pdf>, 2017. pages 38
- [28] Grant Gustafson. Chapter 11: Systems of Differential Equations. <http://www.math.utah.edu/~gustafso/2250systems-de.pdf>. Lecture notes from the University of Utah 2250 course. pages 45, 46
- [29] Martin Rasmussen. MATH50004 Differential Equations Lecture Notes Spring Term 2021/22. pages 46
- [30] Diego Assencio. The double pendulum: Hamiltonian formulation. <https://diego.assencio.com/?index=e5ac36fcb129ce95a61f8e8ce0572dbf>, May 2014. pages 54
- [31] Marc Dvorak and Tim Ohno. Liouville's theorem - inside mines. https://inside.mines.edu/~tohno/teaching/PH505_2011/liouville_dvorak.pdf, Dec 2011. pages 55
- [32] James Meiss. Hamiltonian systems. http://www.scholarpedia.org/article/Hamiltonian_systems#Eq-2, 2007. pages 55
- [33] Marc R. Roussel. Hamiltonian systems. <https://people.uleth.ca/~roussel/nld/Hamiltonian.pdf>, Oct 2005. pages 55
- [34] Aurélien Géron. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow : concepts, tools, and techniques to build intelligent systems*. O'Reilly, Beijing, 2nd edition, 2019. pages 57

Appendix A

Further stability by linearisation

A.0.1 Case 3: unstable equilibrium point $\theta_1 = \pi, \theta_2 = 0$

Lets now consider the equilibrium point $(\theta_1, \theta_2) \approx (\pi, 0)$ then using the same conventions as before with $\pi - \theta_1 = \phi_1$. We have

$$x_1 = \sin(\pi - \phi_1) = \sin(\phi_1) \approx \phi_1 \quad (\text{A.1a})$$

$$y_1 = -\cos(\pi - \phi_1) = \cos(\phi_1) \approx 1 - \frac{\phi_1^2}{2} \quad (\text{A.1b})$$

$$x_2 = \sin(\phi_1) + \sin(\theta_2) \approx \phi_1 + \theta_2 \quad (\text{A.1c})$$

$$y_2 = \cos(\phi_1) - \cos(\theta_2) \approx -\frac{\phi_1^2}{2} + \frac{\theta_2^2}{2} \quad (\text{A.1d})$$

The derivatives with respect to time are then:

$$\dot{x}_2 \approx \dot{\phi}_1 + \dot{\theta}_2 \quad (\text{A.2a})$$

$$\dot{y}_2 \approx -\phi_1 \dot{\phi}_1 + \theta_2 \dot{\theta}_2 \quad (\text{A.2b})$$

Since we are only considering up to second order terms we don't need to calculate \dot{y}_2 as once we square this term we will be left with 4th order terms.

We use the energy formulas once more and find:

$$T = \frac{\dot{\phi}_1^2}{2} + \frac{\dot{\phi}_1^2 + 2\dot{\phi}_1\dot{\theta}_2 + \dot{\theta}_2^2}{2} \quad (\text{A.3})$$

$$V = g \left[\left(1 - \frac{\phi_1^2}{2} \right) + \left(1 - \frac{\phi_1^2}{2} \right) - \left(1 - \frac{\theta_2^2}{2} \right) \right] \quad (\text{A.4})$$

Plugging T and V into the Laplacian we find that:

$$L = \dot{\phi}_1^2 + \dot{\phi}_1\dot{\theta}_2 + \frac{\dot{\theta}_2^2}{2} + g\phi_1^2 - \frac{g\theta_2^2}{2} - g \quad (\text{A.5})$$

Applying Lagrange's equations from (2.2) we get the result below.

$$2\ddot{\phi}_1 + \ddot{\theta}_2 = 2g\phi_1 \quad (\text{A.6a})$$

$$\ddot{\phi}_1 + \ddot{\theta}_2 = -g\theta_2 \quad (\text{A.6b})$$

We can put these equations into matrix vector form to give us the system:

$$\underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} \ddot{\phi}_1 \\ \ddot{\theta}_2 \end{bmatrix} = g \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \theta_2 \end{bmatrix} \quad (\text{A.7})$$

We can rewrite this as

$$\begin{bmatrix} \ddot{\phi}_1 \\ \ddot{\theta}_2 \end{bmatrix} = g \underbrace{\begin{bmatrix} 2 & 1 \\ -2 & -2 \end{bmatrix}}_{\mathbf{B}} \begin{bmatrix} \phi_1 \\ \theta_2 \end{bmatrix} \quad (\text{A.8})$$

Finding the eigenvalues of the matrix \mathbf{B} will tell us the stability of the equilibrium point $(\theta_1, \theta_2) = (0, \pi)$. We find that there are two eigenvalues $\lambda_1 = \sqrt{2}g > 0$ and $\lambda_2 = -\sqrt{2}g < 0$ with corresponding eigenvectors $E_1 = \begin{bmatrix} -(\sqrt{2}+1) \\ \sqrt{2} \\ 1 \end{bmatrix}$ and $E_2 = \begin{bmatrix} -(\sqrt{2}-1) \\ \sqrt{2} \\ 1 \end{bmatrix}$.

The equilibrium point $(\theta_1, \theta_2) = (\pi, 0)$ is a saddle point, thus it is unstable.

A.0.2 Case 4: unstable equilibrium point $\theta_1 = \theta_2 = \pi$

The last equilibrium point to consider is when $(\theta_1, \theta_2) \approx (\pi, \pi)$. We notice from the previous two equilibrium points that the differences in the derivation to the first equilibrium point arises due to the gravitational potential energy and having a positive contribution instead of a negative contribution. We have the same effect here and so the system can be derived as:

$$\underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} \ddot{\phi}_1 \\ \ddot{\phi}_2 \end{bmatrix} = g \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \quad (\text{A.9})$$

$$\begin{bmatrix} \ddot{\phi}_1 \\ \ddot{\phi}_2 \end{bmatrix} = g \underbrace{\begin{bmatrix} 2 & -1 \\ -2 & 2 \end{bmatrix}}_{\mathbf{B}} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \quad (\text{A.10})$$

The eigenvalues of this matrix are $\lambda_1 = (2 + \sqrt{2})g > 0$ and $\lambda_2 = (2 - \sqrt{2})g > 0$ with corresponding eigenvectors $E_1 = \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ 1 \end{bmatrix}$ and $E_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 1 \end{bmatrix}$. Therefore, the equilibrium point $(\theta_1, \theta_2) = (\pi, \pi)$ is unstable.

Appendix B

Varying parameters for a periodic linearisation result

One way of varying parameters to try to achieve a rational frequency ratio is by considering the mass ratio $M = \frac{m_2}{m_1}$. We will still assume that the rods have equal lengths $l_1 = l_2 = l$. Under this assumption, equation 3.16 simplifies to

$$\begin{bmatrix} 1+M & M \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} + \begin{bmatrix} \frac{g}{l} + M\frac{g}{l} & 0 \\ 0 & \frac{g}{l} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = 0 \quad (\text{B.1})$$

This implies that

$$\begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} = - \begin{bmatrix} 1+M & M \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \frac{g}{l} + M\frac{g}{l} & 0 \\ 0 & \frac{g}{l} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \frac{-g}{l} \begin{bmatrix} 1+M & -M \\ -(1+M) & 1+M \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \quad (\text{B.2})$$

Here, we will take a slightly different approach to section 3.3 to show a different method. We can rewrite equation B.2 as a first order differential equation by letting $\varphi_i = \dot{\theta}_i$, \mathbf{A} be the matrix in equation B.2, \mathbf{I}_2 and $\mathbf{0}$ be the 2×2 identity and zero matrices, and \mathbf{y} be $(\theta_1, \theta_2, \varphi_1, \varphi_2)^T$. Thus we have

$$\dot{\mathbf{y}} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_2 \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \mathbf{y} \quad (\text{B.3})$$

Let \mathbf{B} denote the block matrix in equation B.3. To find the eigenvalues of \mathbf{B} , we note that:

Lemma B.0.1. \mathbf{B} has characteristic polynomial $\det(\mathbf{A} - \lambda^2 \mathbf{I}_2)$ [28].

Therefore, λ is an eigenvalue of \mathbf{B} iff

$$\lambda^2 = -\frac{g}{l}(1+M \pm \sqrt{M(1+M)}) \quad (\text{B.4})$$

It also follows immediately from Lemma B.0.1 that,

Lemma B.0.2. λ is an eigenvalue of \mathbf{B} iff λ^2 is an eigenvalue of \mathbf{A} .

By B.4, \mathbf{B} has 4 distinct purely imaginary eigenvalues, which must be in conjugate pairs.

Lemma B.0.3. If $\mathbf{u} + i\mathbf{v}$ is an eigenvector of \mathbf{B} corresponding to eigenvalue λ , where \mathbf{u} and \mathbf{v} are real vectors, then $\mathbf{u} - i\mathbf{v}$ is an eigenvector of $\bar{\lambda}$ [29].

Lemma B.0.4. $(\mathbf{B} - \lambda \mathbf{I}_4)\mathbf{x} = 0$ iff $\mathbf{A}\mathbf{x}_1 = \lambda^2 \mathbf{x}_1$ and $\mathbf{x}_2 = \lambda \mathbf{x}_1$, where $\mathbf{x}_1, \mathbf{x}_2$ are two dimensional vectors and $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)^T$ [28].

Using the above lemmas and general results for solving systems of first order ODEs with constant coefficients, we can proceed as follows:

$$\text{Let } \omega_+ = \sqrt{\frac{g}{l}} \sqrt{1 + M + \sqrt{M(1 + M)}} \text{ and } \omega_- = \sqrt{\frac{g}{l}} \sqrt{1 + M - \sqrt{M(1 + M)}}.$$

We find that the eigenvectors of \mathbf{A} corresponding to its eigenvalues $-\omega_+^2$ and $-\omega_-^2$ are $\begin{bmatrix} 1 \\ -\sqrt{\frac{1+M}{M}} \end{bmatrix}$ and $\begin{bmatrix} 1 \\ \sqrt{\frac{1+M}{M}} \end{bmatrix}$, so we obtain the more general form of equation 3.33:

$$\begin{bmatrix} \theta_1(t) \\ \theta_2(t) \end{bmatrix} = a \begin{bmatrix} 1 \\ -\sqrt{\frac{1+M}{M}} \end{bmatrix} \cos(\omega_+ t + \phi_1) + b \begin{bmatrix} 1 \\ \sqrt{\frac{1+M}{M}} \end{bmatrix} \cos(\omega_- t + \phi_2) \quad (\text{B.5})$$

where a, b, ϕ_1, ϕ_2 are determined by the initial conditions.

For periodic behaviour, we require a rational frequency ratio:

$$\frac{\omega_+}{\omega_-} = \frac{\sqrt{1 + M + \sqrt{M(1 + M)}}}{\sqrt{1 + M - \sqrt{M(1 + M)}}} = x, x \in \mathbb{Q} \quad (\text{B.6})$$

Squaring both sides and rationalising the denominator leads to

$$\frac{(1 + M)^2 + M(1 + M) + 2(1 + M)\sqrt{M(1 + M)}}{1 + M} = 1 + 2M + 2\sqrt{M(1 + M)} = x^2 \quad (\text{B.7})$$

With some rearranging, we find that a mass ratio of $M = \frac{(1 - x^2)^2}{4x^2}, x \in \mathbb{Q}$ leads to periodic linearisation result.

For the following time plots, we will assume that $l = 1$, $g = 9.81$, $m_1 = 1$, and $\dot{\theta}_1 = \dot{\theta}_2 = 0$. The time plots suggest the superposition of two waves is a good approximation for the full non-linear system for small angles and small lengths of time. In Figure B.1, $x = 1.4$ in the first plot, which seems periodic, while $x = \sqrt{2}$ in the second plot, which seems quasiperiodic. In Figure B.2, we consider how the full system deviates from the linearised system by plotting θ_1 against the solution for θ_1 from the linearisation for $x = 2, 5$ and 10 for the initial angles $\theta_1 = 0, \theta_2 = \frac{\pi}{10}$. Systems with lower values of x appear to stay ‘close’ to the linearisation for longer, which is reasonable when we compare the full master equations for the system against the linearisation.

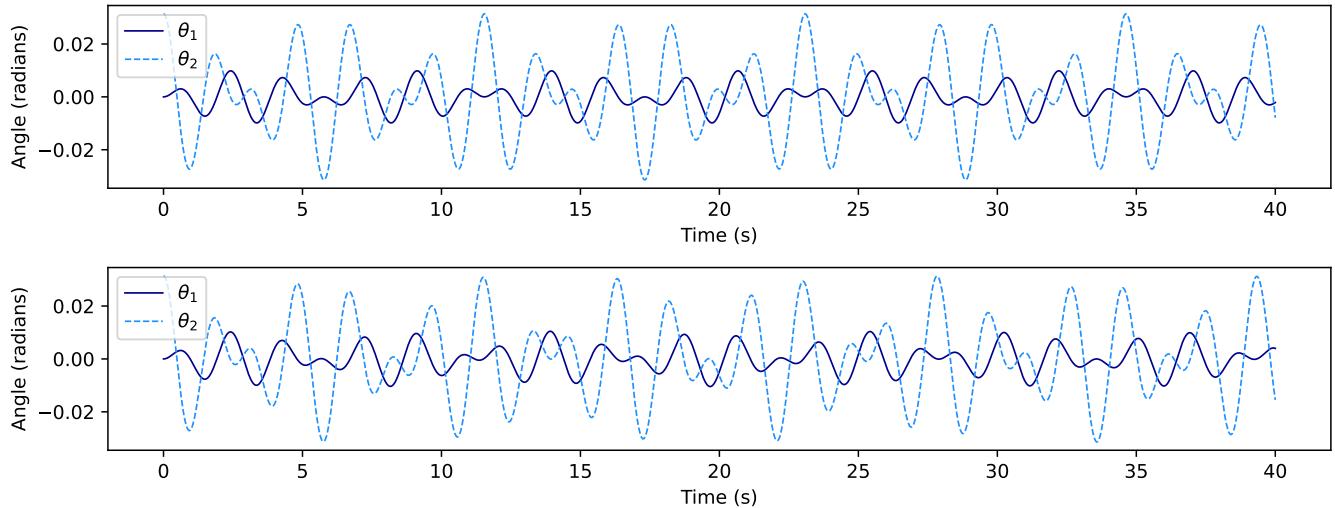


Figure B.1: Time plots for initial angles $\theta_1 = 0$, $\theta_2 = \frac{\pi}{100}$ for comparing a rational and irrational frequency ratio.

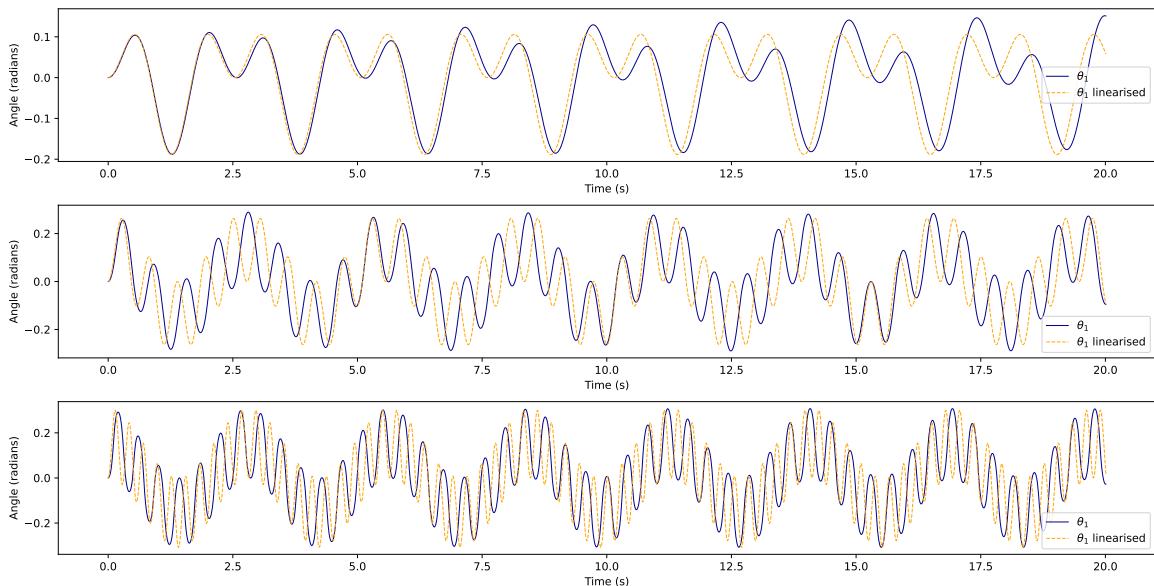


Figure B.2: Comparing deviations from the linearised system for $x = 2, 5, 10$.

Appendix C

Definition of Lyapunov exponent

Definition C.0.1 (Lyapunov exponents[8]). Consider the C^r ($r > 1$) vector field

$$\dot{x} = f(x), \quad x \in \mathbb{R}^n. \quad (\text{C.1})$$

Let $x(t, x_0)$ be a trajectory of [C.1](#) satisfying $x(0, x_0) = x_0$. Consider the orbit structure of the linearisation of [C.1](#) about $x(t, x_0)$ given by

$$\dot{\epsilon} = Df(x(t))\epsilon, \quad \epsilon \in \mathbb{R}^n \quad (\text{C.2})$$

Let $X(t; x(t, x_0))$ be the fundamental solution matrix of [C.2](#), and let $e \neq 0$ be a vector in \mathbb{R}^n . Then the *coefficient of expansion in the direction e along the trajectory through x_0* is defined to be

$$\lambda_t(x_0, e) \equiv \frac{\|X(t; x(t, x_0))e\|}{\|e\|} \quad (\text{C.3})$$

where $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ with $\langle \cdot, \cdot \rangle$ denoting the standard scalar product on \mathbb{R}^n . The *Lyapunov exponent in the direction e along the trajectory through x_0* is defined to be

$$\chi(X(t; x(t, x_0)), x_0, e) \equiv \overline{\lim_{t \rightarrow \infty}} \left(\frac{1}{t} \log \lambda_t(x_0, e) \right) \quad (\text{C.4})$$

Appendix D

More about recurrence quantification analysis

Definition D.0.1 (determinism[4]). *Determinism* of the system is defined as

$$DET = \frac{\sum_{l=l_{min}}^N lP(\epsilon, l)}{\sum_{l=1}^N lP(\epsilon, l)}, \quad (\text{D.1})$$

where $P(\epsilon, l)$ is a histogram of diagonal lines of length l , i.e.

$$P(\epsilon, l) = \sum_{i,j=1}^N (1 - \mathbf{R}_{i-1,j-1}(\epsilon))(1 - \mathbf{R}_{i+l,j+l}(\epsilon)) \prod_{k=0}^{l-1} \mathbf{R}_{i+k,j+k}(\epsilon). \quad (\text{D.2})$$

From the definition, $P(\epsilon, l)$ counts the number of diagonal lines with length exactly l . Thus DET stands for the fraction of the number of diagonal lines with length at least l_{min} . The diagonal lines with length l means the orbit will return to resembling a particular state exactly l times. In this case, the less chaotic the behaviour is, the higher we expect the DET value to be.

However, due to the complexity of the algorithm, we did not get a really clear transition pattern using this method. But it is still worth mentioning since the method provide an intuitive way to quantify recurrence plots.

Appendix E

Figures referred to in Chapter 4

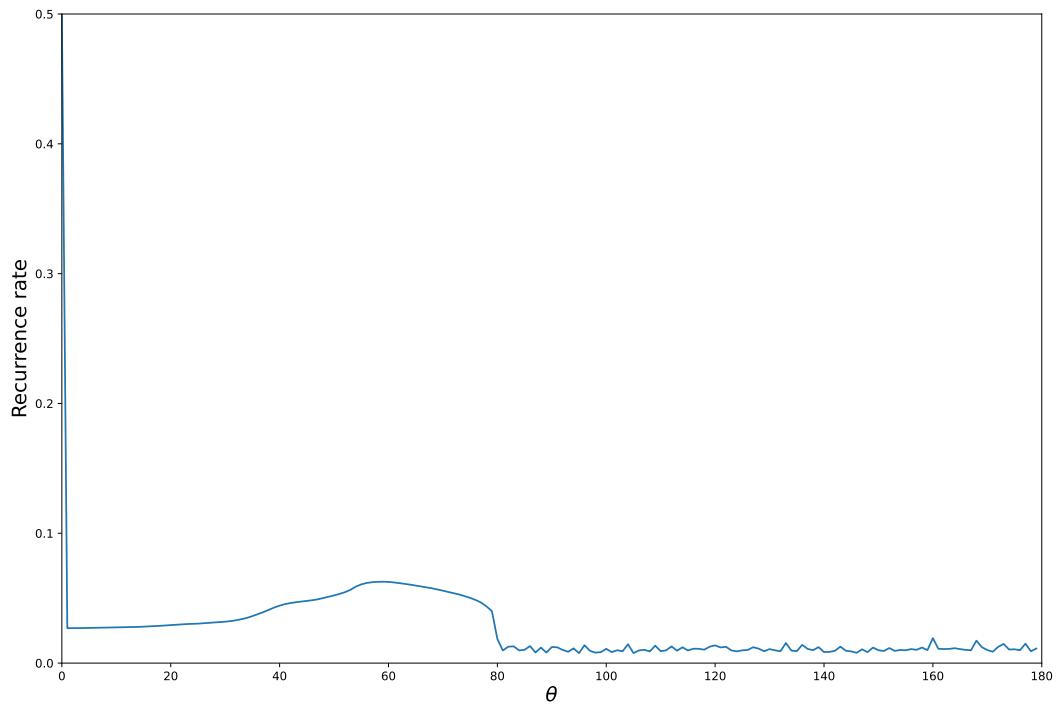


Figure E.1: Recurrence rate for symmetric initial condition $\theta_1(0) = \theta_2(0) = 0$

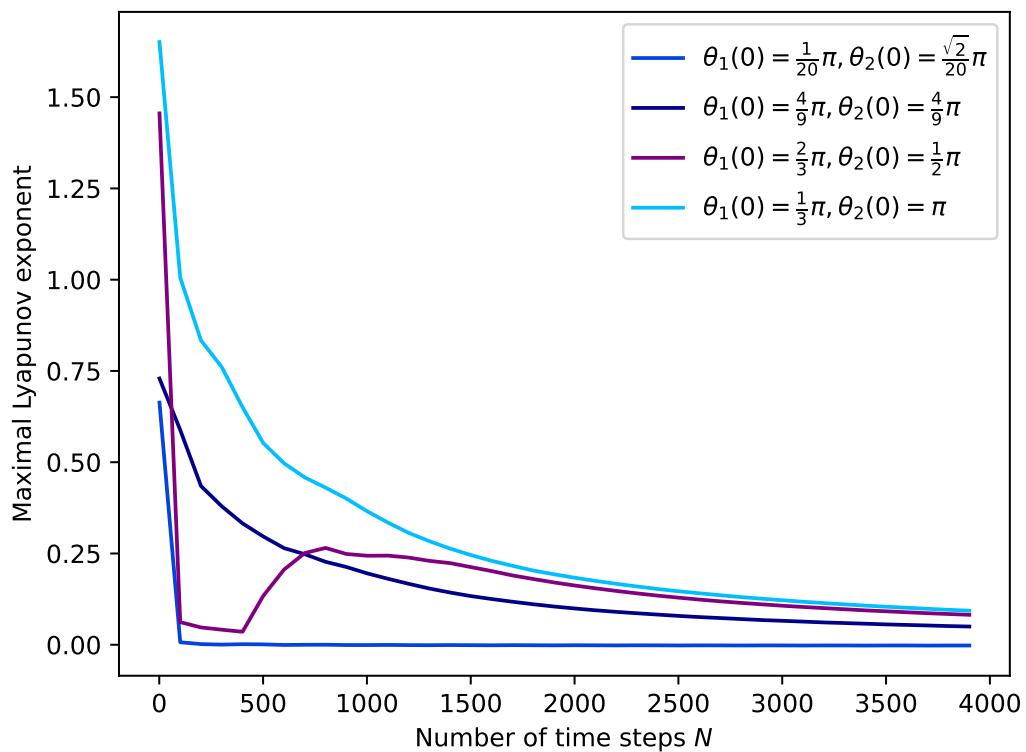


Figure E.2: Maximal Lyapunov exponents determined numerically for different values of N , the number of time steps, for varying initial angles and zero initial velocity.

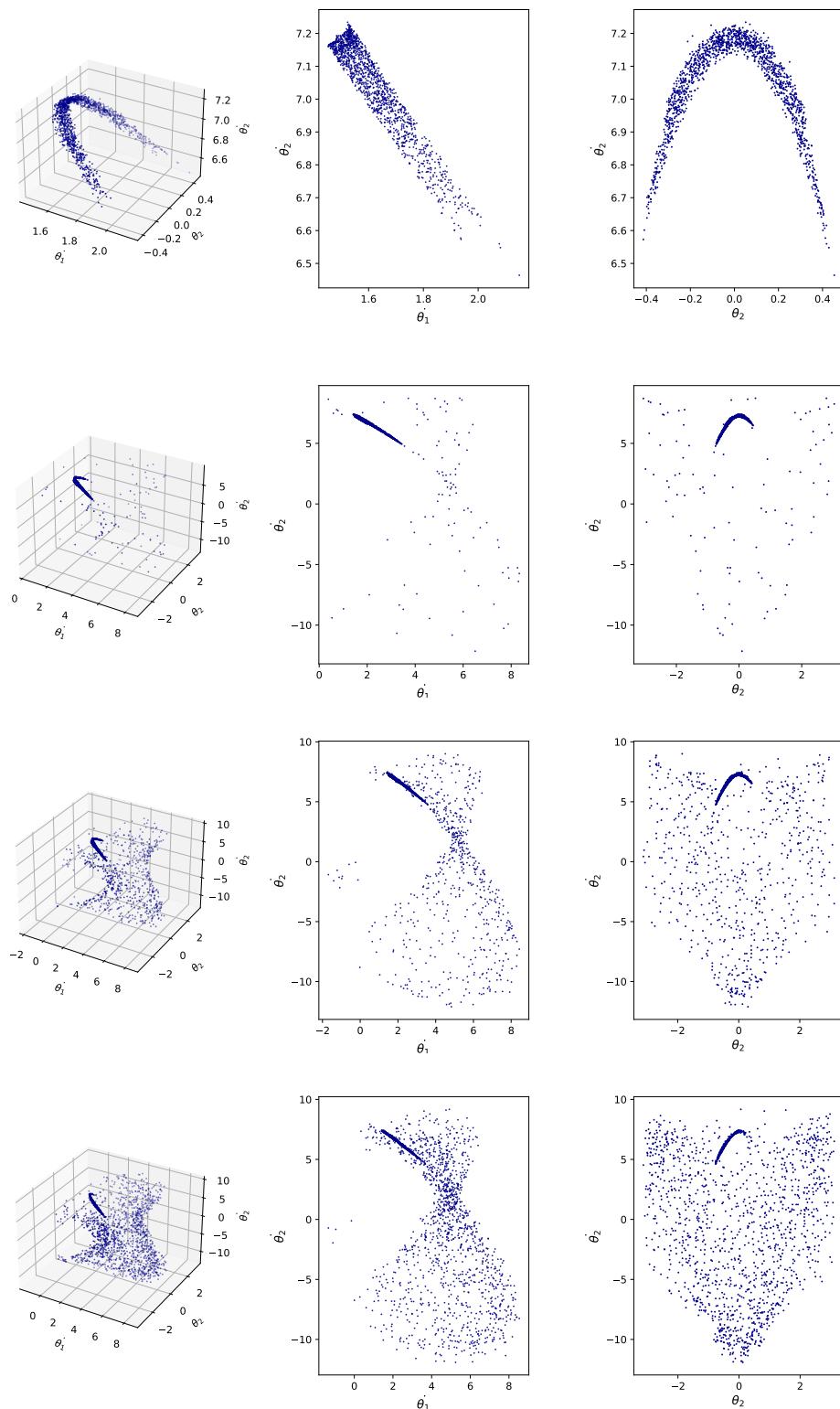


Figure E.3: The initial conditions for these plots are of the form $\theta_1 = c \times \frac{2}{3}\pi$, $\theta_2 = c \times \frac{1}{2}\pi$, $\dot{\theta}_1 = \dot{\theta}_2 = 0$, where $c = 0.98, 0.99, 1, 1.1$ for the 1st to 4th plots respectively. This corresponds to the purple line of Figure 4.13

Chapter E. Figures referred to in Chapter 4

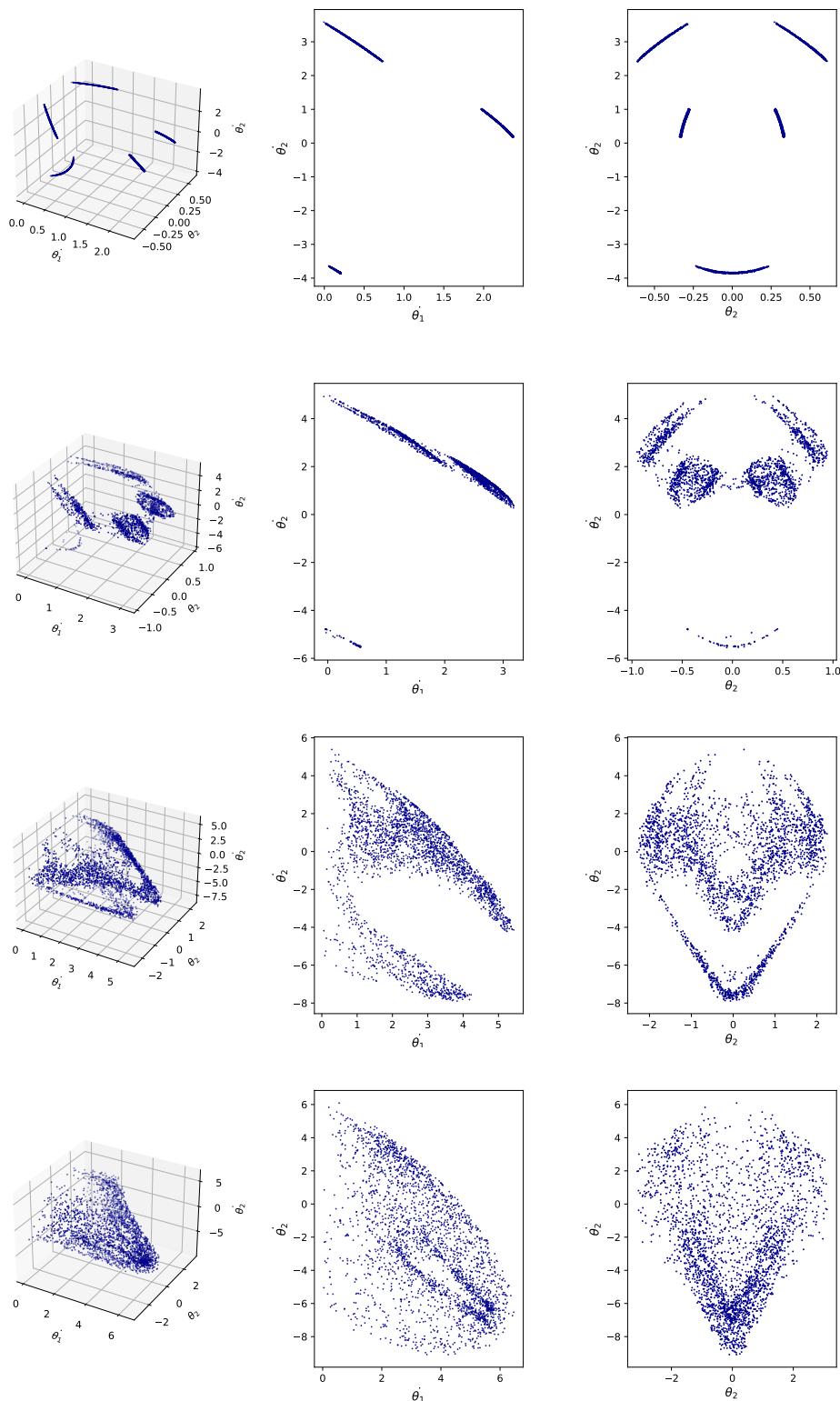


Figure E.4: The initial conditions for these plots are of the form $\theta_1 = c \times \frac{1}{6}\pi$, $\theta_2 = c \times \frac{1}{2}\pi$, $\dot{\theta}_1 = \dot{\theta}_2 = 0$, where $c = 0.7, 1, 1.2, 1.5$ for the 1st to 4th plots respectively. This corresponds to the lightest blue line of Figure 4.13

Appendix F

Explanation of the applicability of the Poincaré recurrence theorem

First, we show that our double pendulum is a Hamiltonian system. To do that, we show that the equations of motion for the double pendulum can be derived through the Hamiltonian. In this way, the method covered here can be seen as an alternative to the 'Lagrange's equations subsection'. However, here we do not derive the full equations of motion. What is important to show here is that they *can* be derived. The full equations of motion can be found at [30], which we have used to help write the formulation below.

The Hamiltonian of the system is defined as $H = \sum_{i=1}^2 \dot{\theta}_i p_{\theta_i} - L$ where L is the Lagrangian as defined in 2.6

and p_{θ_i} are the canonical momenta of the system. They are defined as

$$\begin{cases} p_{\theta_1} = \frac{\partial L}{\partial \dot{\theta}_1} = (m_1 + m_2)l_1^2 \dot{\theta}_1 + m_2 l_1 l_2 \dot{\theta}_2 \cos(\theta_1 - \theta_2) \\ p_{\theta_2} = \frac{\partial L}{\partial \dot{\theta}_2} = m_2 l_2^2 \dot{\theta}_2 + m_2 l_1 l_2 \dot{\theta}_1 \cos(\theta_1 - \theta_2) \end{cases} \quad (\text{F.1})$$

The set of equations for motion will then be defined as:

$$\dot{\theta}_i = \frac{\partial H}{\partial p_{\theta_i}}, \dot{p}_{\theta_i} = -\frac{\partial H}{\partial \theta_i}$$

However, our Hamiltonian H contains $\dot{\theta}_1$ and $\dot{\theta}_2$. These need to be removed before we can start deriving the equations of motion.

We can represent equations F.1 as the matrix system below. Let $\Delta = \theta_1 - \theta_2$:

$$\begin{bmatrix} p_{\theta_1} \\ p_{\theta_2} \end{bmatrix} = \begin{bmatrix} (m_1 + m_2)l_1^2 & m_2 l_1 l_2 \cos(\Delta) \\ m_2 l_1 l_2 \cos(\Delta) & m_2 l_2^2 \end{bmatrix} \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} \quad (\text{F.2})$$

Now, the determinant of the matrix on the RHS is $m_1 m_2 l_1^2 l_2^2 + m_2^2 l_1^2 l_2^2 [1 - \cos^2(\Delta)] = m_1 m_2 l_1 l_2^2 + m_2^2 l_1^2 l_2^2 \sin^2(\Delta) \geq m_1 m_2 l_1^2 l_2^2 > 0$ as $m_1, m_2, l_1, l_2 > 0$.

Thus, we can obtain expressions for θ_1 and θ_2 by inverting this system, and then plug them into our expression for H , before then deriving the four equations of motion. Therefore, our double pendulum can indeed be represented as a Hamiltonian system. We can now look at Liouville's theorem for dynamical systems.

Theorem F.0.1 (Liouville's theorem [31][32]). *Liouville's theorem states that the phase space density of a certain element as it moves in phase space is fixed. This holds for any divergence free vector-field.*

A restatement of the above theorem is that if we take an arbitrary space of initial conditions in our phase space (which in our case will be a 4 dimensional object) and the measure its volume, if Liouville's theorem holds, that volume will stay constant throughout the evolution of the system. Mathematically, if we have a set of initial conditions U and let the flow of our system be represented as $\phi(x, t)$ for initial condition x , $\int_U dz = \int_{\phi(U, t)} dz$ where z represents the phase space variables.

We now show that every Hamiltonian system obeys Liouville's theorem.

Theorem F.0.2 (Every Hamiltonian system obeys Liouville's theorem [32]).

Proof. We need to show that the divergence of every Hamiltonian system is 0. According to [32], every Hamiltonian system with $2n$ degrees of freedom (where a degree of freedom is defined as a (θ_i, p_i) pair [33]) can be written as

$$\dot{z} = \mathbf{J} \nabla H(z, t)$$

where $\mathbf{J} = \begin{bmatrix} 0 & \mathbf{I}_n \\ -\mathbf{I}_n & 0 \end{bmatrix}$ and \mathbf{I}_n is the $n \times n$ identity matrix, H is a smooth scalar function, z consists of the phase space variables (which in our case will consist of θ_1, θ_2 and the two canonical momenta), and t is time.

$\nabla \cdot \dot{z} = \nabla \cdot \mathbf{J} \nabla H = \sum_{i,j} \mathbf{J}_{ij} \frac{\partial H}{\partial z_i \partial z_j}$. Now, \mathbf{J} is skew-symmetric as $\mathbf{J}^T = -\mathbf{J}$ and the Hessian of H is symmetric due to Schwartz' Theorem. Thus, this sum is 0 and hence the divergence of our system is 0. Therefore, the system obeys Liouville's theorem. ■

Now that we have shown that every Hamiltonian system obeys Liouville's theorem and is thus volume-preserving, we can prove that the Poincaré recurrence theorem holds for our double pendulum.

Theorem F.0.3 (Poincaré Recurrence Theorem [16]). *The Poincaré recurrence theorem implies the density of recurrent points for volume-preserving dynamical systems on compact domains.*

A perhaps simpler restatement of the above theorem is that if a dynamical system is volume-preserving and compact, then after a finite amount of time the system will return arbitrarily close to its initial conditions (and equivalently every other point in its orbit that can be reached in finite time).

We know our double pendulum is volume-preserving. It is also clear to see that for any initial condition, our system will move within a compact domain of the state space - θ_1 and θ_2 will be restricted to the range $[-\pi, \pi)$, and $\dot{\theta}_1$ and $\dot{\theta}_2$ will have an upper bound determined by the initial total energy of the system. Thus, the Poincaré recurrence theorem applies to our double pendulum.

Appendix G

Proof of concept of obtaining the numerical solution via a neural network

While throughout this paper we used the Sympy Python module to obtain a numerical solution to our double pendulum system, according to [24] it is also possible to use Machine Learning methods to obtain the solution. Here, we give a rough proof of concept that such a method may be viable, and is something that could be further investigated given more time.

As in the paper, we attempt to train a Long Short-Term Memory model (LSTM) to predict the future x and y co-ordinates of the masses given the previous ones. An LSTM is a recurrent neural network, which means it is suited to work with temporal/sequences of data.

We show that an LSTM can quite accurately predict the last second of motion having been trained on the first 4 seconds of motion for a double pendulum with initial conditions ($\theta_1 = \frac{\pi}{20}$, $\theta_2 = \frac{\pi}{20}$, $\dot{\theta}_1 = 0$, $\dot{\theta}_2 = 0$), which is a non-chaotic initial condition.

Here is the code that trains the LSTM (we used [34] to help with the code):

Listing G.1: Code to train the LSTM

```
y0 = [np.pi/20, 0, np.pi/20, 0]
p1 = Pendulum(np.pi/3, 0, np.pi/3, 0, y0, tmax=5, dt=0.001)
# Get the (x,y) co-ordinates of the bobs in the numerical
full_sol = np.array([p1.x1, p1.y1, p1.x2, p1.y2])
full_sol = full_sol.T
ax_0_len = full_sol.shape[0]
train_len = int(0.8 * ax_0_len)
```

```
train_dataset = tf.data.Dataset.from_tensor_slices(full_sol[:train_len,:])
n_steps = 100
window_length = n_steps + 1
dataset = train_dataset.window(window_length, shift=1, drop_remainder=True)
dataset = dataset.flat_map(lambda window: window.batch(window_length))
from tensorflow.python.ops.gen_dataset_ops import
    dataset_to_single_element_eager_fallback
batch_size = 8
dataset = dataset.batch(batch_size)
dataset = dataset.map(lambda windows: (windows[:, :-1], windows[:, 1:]))
dataset = dataset.prefetch(1)

model = keras.models.Sequential([
    keras.layers.LSTM(128, return_sequences=True, input_shape=[None, 4],
                      dropout=0.2, recurrent_dropout=0.2, activation='relu'),
    keras.layers.LSTM(128, return_sequences=True, input_shape=[None, 4],
                      dropout=0.2, recurrent_dropout=0.2, activation='relu'),
    keras.layers.TimeDistributed(keras.layers.Dense(4))
])

model.compile(loss=euc_dist_keras, optimizer=tf.keras.optimizers.Adam(),
              metrics="accuracy")
history = model.fit(dataset, epochs=10)

#Calculate Euclidean distance between predictions and actual result
from keras import backend as K
def euc_dist_keras(y_true, y_pred):
    return K.sqrt(K.sum(K.square(y_true - y_pred), axis=-1, keepdims=True))
```

Note:

1. We train for 10 epochs (which means that we go through the entire dataset 10 times while training).
2. Just like the paper, we use a batch size of 8 (which means that the weights of the neural network are updated every time after 8 inputs have been processed.)
3. We used a window size of length 100. This means that the network will use the last 100 elements in the sequence to help make its next prediction.
4. We use two LSTM layers, each with 128 units.
5. Just like the paper, we use the ReLU activation function and let our last layer consist of a layer of four units.

6. After experimenting, we decided to use a Euclidean distance function as our loss function.
7. We decided to use the Adam optimiser rather than Gradient Descent to train our neural network.

At the end of training, we tested the model on the second of motion of the double pendulum, and plotted the euclidean distance between the vector of the actual x-y coordinates of the two masses, and the predicted value.

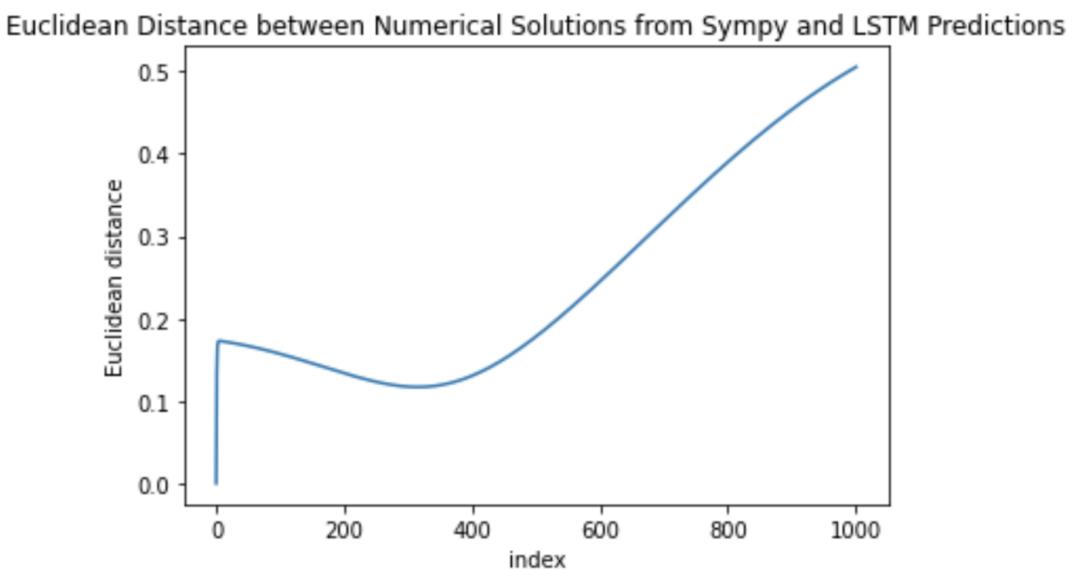


Figure G.1: Plot showing the error in the predictions of the motion for the last second of the double pendulum

As we can see, the error is quite small for the first 500ms, however, after that it starts to increase quite rapidly. So, this method of obtaining predictions seems that it could be promising, however, more time and model tuning would be needed to determine if this method is viable, especially for initial conditions that lead to chaotic motion.