第一页为封面页

参赛队员姓名： 王习森

中学： 北京市二十一世纪国际学校

省份： 北京

国家/地区： 中国

指导教师姓名： 刘亦柠

指导教师单位： 北京市二十一世纪国际学校

论文题目：针对双单摆训练的长短期记忆网络的白噪音检测

White Noise Tests on the LSTM Model Trained with Double Pendulum

White Noise Tests on the LSTM Model Trained with Double Pendulum

Xisen Wang

Abstract

This paper describes the intrinsic qualities of a simple double pendulum (DP) with a visual representation, a rigorous deduction of the Lagrangian equation, and a concrete factor analysis. The Long-short-term-memory (LSTM) model was utilized to simulate the double pendulum's periodic and chaotic behaviors and used to evaluate the effectiveness of the model. The auto-correlation coefficients were calculated. Meanwhile, the Box-Pierce tests and Ljung-Box tests for various state-dependent time series were conducted to give various initial conditions to explore the DP system's random characteristics. The research results are as follows: 1) Chaos did not lead to direct randomness; 2) seasonality could coexist with chaos; 3) the highly auto-regressive nature of DP's time series data was found. Therefore, it can be concluded that the chaos in a double pendulum is positively related to the likelihood of being a random white noise series.

Keywords:Chaos; Randomness; Double Pendulum; LSTM Model; White Noise Test;

CONTENT TABLE

1.INTRODUCTION

1.1 Background

The sensitive dependence on initial conditions is the hallmark of chaos. Chaos theory and its phenomena exist in many places: a cantilever beam, fluid-conveying viscoelastic nanotubes, ecological successions, and so forth[15] [18]. It is essential to note that, in particular, a double pendulum exhibits characteristics typical to the chaos when its pendulums are released from relatively large initial angles [16]. Chaotic motion is often unpredictable, which stands true in the double pendulum case as well.

A simple dynamic model in physics, the double pendulum (DP) manifests some of the most intriguing characteristics due to its special double differential equation [5]. The DP system demonstrates a dramatic growth of separations that are intrinsic to chaotic motion**[16].** In physics, the chaotic nature of double pendulums could be examined numerically through Poincare sections and bifurcation diagrams[17].In 2020, Chokri Nouar and Zine El Abidine Guennoun utilized the cryptographic propriety of double pendulum to propose a new mechanism for generating pseudo-random numbers[12]. Despite this useful and successful application, they did not dive into the theoretical connection between chaos and randomness, which this paper aims to accomplish.

Building on the theories proposed and analyses conducted above, this paper proposes the theoretical connections between randomness and chaos by evaluating the intrinsic random nature of the chaotic motion of double pendulums. Given that machine learning predicts the states of the bobs in a way different from the Ordinary Differential Equation (ODE) solution, data from both methods will be evaluated and compared.

In this paper, the following questions will be investigated in different sections: As a tool for our research, what are the fundamental qualities of a simple double pendulum? Secondly, why is the motion equation for the double pendulum chaotic? Thirdly, what is a method to effectively model the DP system and further analyze its data? Fourthly, will the time series data of the double pendulum be pure white noise? Fifthly, what is the correlation between chaos and randomness regarding quantified parameters? Lastly, what are the directions for converting the time-series data to approach “true randomness”? As these questions are answered in the paper, the directions for further explorations will be pointed out and the relationship between chaos and randomness will be discussed. This paper aims to shed light on the theoretical foundation for subsequent interdisciplinary applications.

1.2 Methodology

After presenting a point-mass double pendulum model and illustrating its internal characteristics, the paper will approach the kinematics by running an OED solution through Python. The experiment's approaches include simulations and statistical tests.

By setting parameters such as time and initial angle, the researcher will gather 5000 initial time series data labeled with their state-related data. Once the data set is constructed, the author tests different architectures for the LSTM network and builds a simulated double-pendulum model. The author conducts extensive corresponding tests for zero auto-correlation of the angular velocity and Cartesian positions with various initial conditions to explore the system's randomness characteristics. Random walk will be tested and different components of the time series data will be separated.

1.3 Aim of the paper

The paper aims to illustrate the quantifiable characteristic of the point-mass double pendulum, design a LSTM model to simulate the double pendulum, and apply statistical tests to evaluate the randomness of the data generated from OED sets and the model. Isolating different component and calculating auto-correlation coefficients, it aims to provide working models that could approach "TRUE RANDOM". It explores the connections between chaos and randomness from the angle of double pendulum and aims to provide another image on the theoretical foundation for cryptography, neural network, and other applications that may follow.

2.LITERATURE REVIEW

The difficulty of deciding whether a given (pseudo) random number sequence is "sufficiently random" makes it hard to formulate a rigorous operation definition of randomness in terms of classical mathematical primitives[8]. The existence of chaos in the double pendulum seems to open another door for that.

Checking white noise in time series forecasting is a feasible and appealing way to evaluate randomness[4]. One of the double pendulum's primary characteristics is temporal dependence, as the states of the pendulums are measured sequentially over time. In analyzing the data, inference on the serial condition $\rho_k = corr(x_t, x_{t-k})$ is a common first step in studying the time series data $\{x_t\}$. For a sample $x_1, x_2, ..., x_n$, estimation of ρ_k for various lags k = 1, 2, ... traces its roots back primarily to Yule (1926), who coined the term serial correlation[19]. Bartlett (1946) provided a major step forward in a more general analysis by deriving the Bartlett formula which shows that the asymptotic variance depends only on the auto-correlations themselves [2].

For a given set of data, auto-correlation refers to the correlation between values of the same variable [11]. Its formal definition is built with respect to a set of related instance pairs $(z_i, z_j) \in Z$ and a variable X that is defined on these instances, where the auto-correlation is the correlation between the values of X on these instance pairs. Apart from testing for significant serial correlation at only one lag k, testing the cumulative hypothesis using the portmanteau statistics of Box and Pierce (1970) and Ljung and Box (1978) gives new insights [3][10][6]. In finite samples, the Ljung-Box statistic entails a slightly better performance than the Box-Pierce statistic, which is based on the observation that the matrix W in the asymptotic distribution of ρ_hat reduces to identity matrix under the i.i.d. assumption on x_t.

Time series that show no auto-correlation is called white noise [7]. If the series is white noise, then it is a sequence of random numbers and thus cannot be predicted. White noise contributes to model diagnostics in that the correspondence of forecast errors and white noises represents that all of the signal information is harnessed. By conducting the tests mentioned above and plotting the auto-correlations, the paper would examine whether the average and correlation are zero and whether the variance has changed.

3.Model Illustration (Section1)

3.1 Definition

What determines a chaotic system? A prerequisite for chaotic systems is that the differential equations dictating the dynamics should contain non-linear terms. As Troy Shinbrot et. al (1992) defined, a system is said to be chaotic if the eigenvalue lambda — called the Lyapunov exponent— is positive for typical initial conditions[16].

3.2 Description

As the graph illustrates, the point-mass double pendulum consists of two point-masses, m_1 and m_2, suspended by massless wires of length l1 and l2. O is the pivoting point.

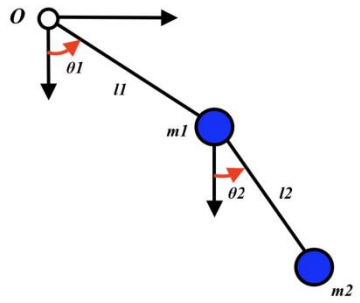


Figure 1: Sketch of a schematic point-mass double pendulum under free oscillation

3.3 Demonstration and deduction

With the origin at the pivotal point of the top pendulum, we could establish a set of equations to shortly investigate the kinematics of double pendulum.

For the positions of the pendulums, we could build on a X-Y coordinate system that would simplify our deduction process, as shown in Equation 1-4.

$$x_1 = L_1 \sin\theta_1 \tag{1}$$

$$y_1 = -L_1 \cos\theta_1 \tag{2}$$

$$x_2 = x_1 + L_2 \sin\theta_2 \tag{3}$$

$$y_2 = y_1 - L_2 \cos\theta_2 \tag{4}$$

When the pendulums are moving, whether periodically or chaotically, they follow these sets of velocity equations, as shown in Equation 5-8.

$$\dot{x_1} = \dot{\theta_1} L_1 \cos\theta_1 \tag{5}$$

$$\dot{y_1} = \dot{\theta_1} L_1 \sin\theta_1 \tag{6}$$

$$\dot{x_2} = \dot{x_1} + \dot{\theta_2} L_2 \cos\theta_2 \tag{7}$$

$$\dot{y_2} = \dot{y_1} + \dot{\theta_2} L_2 \sin\theta_2 \tag{8}$$

Then there are Equation 9-12 for the accelerations:

$$\ddot{x_1} = -\dot{\theta_1}^2 L_1 \sin\theta_1 + \ddot{\theta_1} L_1 \cos\theta_1 \tag{9}$$

$$\ddot{y_1} = \dot{\theta_1}^2 L_1 \cos\theta_1 + \ddot{\theta_1} L_1 \sin\theta_1 \tag{10}$$

$$\ddot{x_2} = \ddot{x_1} - \dot{\theta_2}^2 L_2 \sin\theta_2 + \ddot{\theta_2} L_2 \cos\theta_2 \tag{11}$$

$$\ddot{y_2} = \ddot{y_1} + \dot{\theta_2}^2 L_2 \cos\theta_2 + \ddot{\theta_2} L_2 \sin\theta_2 \tag{12}$$

For energies, let $v_1^2 = \dot{x_1}^2 + \dot{y_1}^2$ and $v_2^2 = \dot{x_2}^2 + \dot{y_2}^2$, then the kinetic energies would be:

$$T_1 = \tfrac{1}{2} m_1 v_1^2 = \tfrac{1}{2} m_1 L_1^2 \dot{\theta_1}^2 \tag{13}$$

$$T_2 = \tfrac{1}{2} m_2 v_2^2 = \tfrac{1}{2} m_2 \left(L_1^2 \dot{\theta_1}^2 + L_2^2 \dot{\theta_2}^2 + 2 L_1 L_2 \cos(\theta_1 - \theta_2) \dot{\theta_1} \dot{\theta_2} \right) \tag{14}$$

Equation 15-16 are for the potential energies of the two pendulum.

$$V_1 = m_1 g y_1 = -m_1 g L_1 \cos\theta_1 \tag{15}$$

$$V_2 = m_2 g y_2 = -m_2 g (L_1 \cos\theta_1 + L_2 \cos\theta_2) \tag{16}$$

Now we form the Lagrangian L= T - V = T_1 + T_2 - V_1 - V_2.

$$\frac{\partial L}{\partial \theta_1} = \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta_1}} \tag{17}$$

$$\frac{\partial L}{\partial \theta_2} = \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta_2}} \tag{18}$$

Applying the Euler Lagrangian Equations (Equation 17 and Equation 18) leads to the Lagrangian form of equations for the motion that is essential in the paper (Equation 19 and 20).

$$-(m_1 + m_2) g L_1 \sin\theta_1 = (m_1 + m_2) L_1^2 \ddot{\theta_1} + m_2 L_1 L_2 \sin(\theta_1 - \theta_2) \dot{\theta_2}^2 + m_2 L_1 L_2 \cos(\theta_1 - \theta_2) \ddot{\theta_2} \tag{19}$$

$$-m_2 g L_2 \sin\theta_2 = m_2 L_2 \ddot{\theta_2} + m_2 L_1 L_2 \cos(\theta_1 - \theta_2) \ddot{\theta_1} + m_2 L_1 L_2 \sin(\theta_1 - \theta_2) \dot{\theta_1}^2 \tag{20}$$

The non-linear terms in the equations reveal that the system is chaotic.

3.4 Factor analysis

A double pendulum is stable and predictable with a small initial angle yet becomes unstable and chaotic with a large angle initial angle. Given that the time duration for predictable behavior is positively correlated with accuracy, the inability to reach infinite accuracy dictates the presence of chaotic motion[16]. Linear stability analysis of double pendulum indicates that its two natural frequencies depend on the mass and length of each pendulum in that system[13].

To further examine the characteristics of the double pendulum and determine how the factors are involved, the author created functions that returned the right-hand side of the ordinary differential equations that describe the system and transformed theta and omega to Cartesian coordinates. A time series of x_1, y_1, x_2, y_2, vx_1, vy_1, vx_2, vy_2 was created, consisted of 30,000 pieces of data from

t=0s to t =50s with an interval of 0.01s to capture the precise movement of the system.

For the sake of simplicity and convenience, the initial condition was set as the following: $L_1 = L_2$ =1.0m. m_1=3kg and m_2=1.0kg.(m_1 was set as 3kg to make the observations more obvious) The gravitational acceleration was set as 9.8kgm/s^2 as is the standard to similar problems. Both initial angles θ_1 and θ_2 were 0.25π.

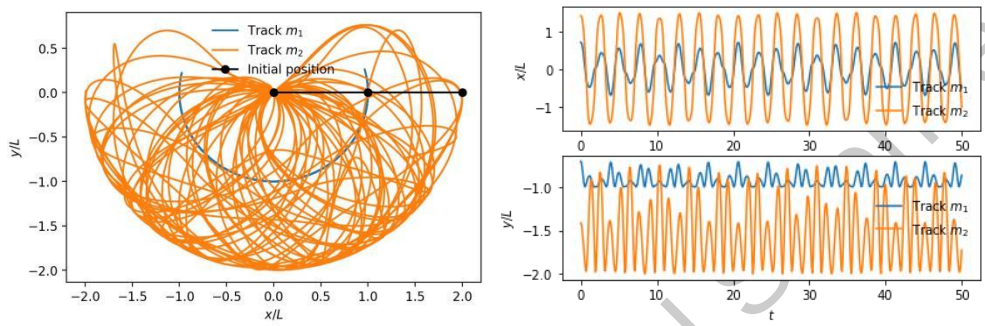


Figure 2: A trajectory simulated through OED

As shown in figure 2, the trajectory of the position of the second bob shows a complex path. Yet if the system is separated by x-axis and y-axis, it would show quasi-periodic behaviors.

4. LSTM MODELING (Section 2)

4.1Input Gathering

Although linear regression with a polynomial feature map could only model a double pendulum with small initial angles, a Long Short-Term Memory Network can accurately give predictions on the kinematics of the system[9].

In this section, the author gathered input data from the ODE solution that it acquired and built upon Keras to design a Recurrent Neural Network (LSTM in particular) to model the position of the two pendulums in x-y axes. With the future defined as $t = t_0 + 20\delta t$, the array of inputs is split into overlapping windows. Arrays from a_i to a_{i+50} are input and a_{i+50} to $a_{i+50+20}$ are labels. (a is an element from the time series of the parameters)

Our dataset contains multiple time series of the states of the double pendulum as described by 4 parameters: the position in x and y coordinates of the two bob [x_1, y_1, x_2, y_2]. These are obtained by the ODE solution from the previous section.

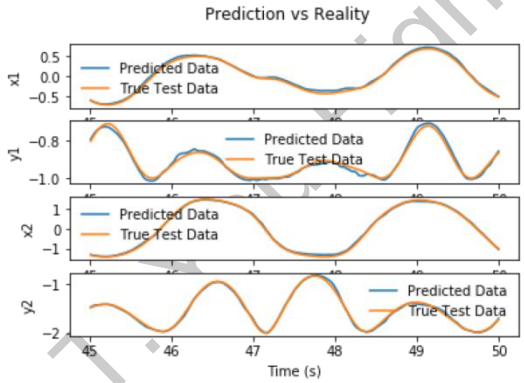
Each time series ranges from 0ms to 5000ms, containing 5001sequences. To assess the performance of the trained LSTM network, the author partitioned the training and validation datasets into different segments and kept the last 500 sequences for validation.

4.2Model Design

The design used the Long-term-short-memory network consisting of 4 hidden layers and an output layer with 4 nodes, namely [x_1, y_1, x_2, y_2]. The researcher used a batch size of 8 and ran the model for 25 epochs while experimenting with different activation functions before deciding upon using a ReLU activation function: f(x) = max (ax+b, 0) for the hidden layers and a linear activation function for the output layer for the structure of our network.

4.3Results

Using the Cartesian coordinates of the bobs, the LSTM model on time series data across various initial conditions was trained. The model worked well on predicting the trajectory of the pendulum (from t=0ms to t=5000ms). The average MSE was $2.35*10^{-3}$in the training set and $1.9*10^{-3}$ in the validation set.



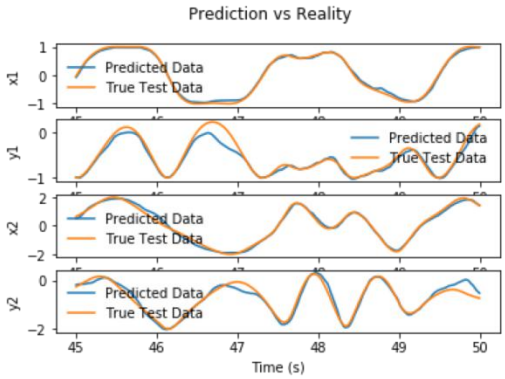
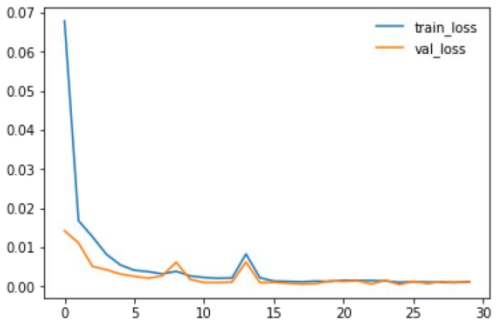


Figure3:The prediction comparison (chaotic) Figure4: The prediction comparison(periodic)

*The chaotic case had initial conditions of $\theta_1=\pi/2$, $\theta_2=\pi/2$; the periodic case had initial conditions of $\theta_1=\pi/4$, $\theta_2=0$.



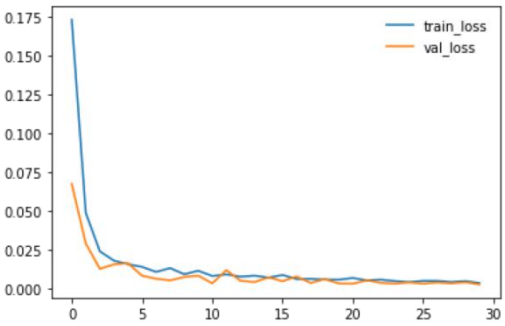


Figure5: The MSE (Pi/4,Pi/4) **Figure6:The MSE (Pi/2,Pi/2)**

As presented in figure5 and figure6, the validation loss as described by MSE keeps an overall decreasing trend and becomes stable since epoch 15.

4.4The Effect of Changing Offset

Setting the initial conditions as stationary pendulums with $\theta_1=\frac{\pi}{4}$ and $\theta_2=\frac{\pi}{4}$, the researcher measured the effect of different offsets on the validation losses of the model. Subsequently, the author gathered 1000 sequences for each time series from 0.0s to 100.0s with a window size of 50.

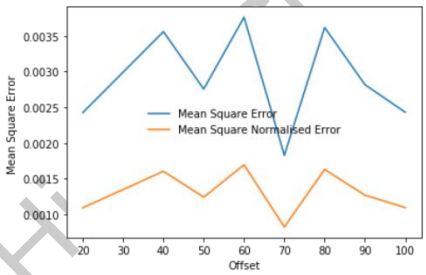


Figure7: the effect of different offsets (from 20 to 100)

As figure 7 shows, the MSE is at minimum when the offset is 70 and at maximum when the offset is 40, 60, and 80. There is no general trend for the effectiveness of offset size. It oscillates around a certain value, so is the MSNE.

Until now, the author has successfully modeled the double pendulum system, analyzed the effect of offsets, and would use the model for further tests.

5.Randomness Evaluation

The author would generate time series with 64 pairs of initial angles and evaluate their randomness by checking white noise series, respectively. The initial conditions consist of the permutation of both angles that have 8 values equally divided by the range Pi/8 to Pi.

The dataset of the angular velocities of each pair in the initial conditions in time series from t=0ms to t=5000ms would be gathered using the ODE method. Their auto-correlation coefficient on lag k = 1, 2, ... would then be evaluated and the Box-Pierce Test and Ljung-Box Test would be conducted to find out if the time series that reflect the states of the double pendulums are merely white noises.

In addition, the author would use the LSTM model to do Box-Pierce Tests and Ljung-Box Test on $[x_1]$, $[y_1]$, $[x_2]$, $[y_2]$ for all the 64 pairs of initial angles. The chaos factor, α, or how chaotic would the system be, is measured in an intuitive sum of the initial angles. Q-α graphs were used to visualize the results.

5.1Test for zero auto-correlation

5.1.1 Method

In theory, if the time series is white noise, then its current value T_i should not be correlated at all with the past values $T_{(i-1)}$, $T_{(i-2)}$, and etc. Thus, the corresponding auto-correlation coefficients ρ_1, ρ_2, and ...etc will be zero or close to zero. The first step to define a white noise is to calculate its auto-correlation coefficient: p_k at any lag k.

For a series $\{x_t\}$ of uncorrelated random variables, the condition of the auto-correlation function (ACF) $\rho_k = \text{corr}(x_t, x_{t-k}) = 0$ at lag k = 1,2,... is well defined for all t with or without the assumption of stationarity. The formula for empirical use, based on observed data $x_1,...x_n$, for the sample auto-correlation ρ_k at lag k !=0 is

$$\widehat{\rho}_k = \frac{\sum_{t=k+1}^{n}(x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}, \quad \bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t \qquad (1)$$

At k=0, the autocorrelation coefficient would simply be zero as the value is always perfectly correlated with itself.

As ρ_k is normally distributed, for a white noise time series, it should have a mean close to zero and some variance σ_k^2. For σ_k, Anderson(1964), Quenouille(1949), and Bartlett(1946) have developed the standard deviation to be as $\sigma_k = 1/\sqrt{n}$, where n is the sample size, in this case, 5000 [1][2][14].

For all 64 time series, the author would calculate and plot the auto-correlation coefficients for 100 lags by the statsmodels library. The paper set the confidence level to be 95% (alpha = 0.05) and concentrated on the time series for angular velocity and for Cartesian positions.

The lag number, step, method, and type of time series were then altered to reshape and reform these data to gain new insights.

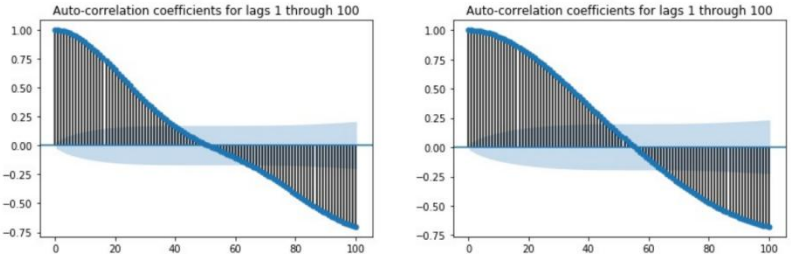


Figure 8: Auto-correlation coefficient plots

*In figure 8, (a) is the auto-correlation coefficient plot for $\theta_1=\frac{\pi}{2}$ and $\theta_2=\frac{\pi}{2}$ & (b) is the auto-correlation coefficient plot for $\theta_1=\frac{\pi}{4}$ and $\theta_2=\frac{\pi}{4}$ generated through the ODE solution. The plots are for angular velocity w_1 but the same pattern is true for w_2. The blue region shows a 95% confidence interval.

5.1.2Results

After scrutinizing all 64 plots of auto-correlation, the author found that DP's Cartesian positions and angular velocities, in their current forms, have a rather high degree of auto-correlation and are not white noise. One could make the conclusion that the data come from an underlying auto-regressive model with strong positive auto-correlation. Starting with a high auto-correlation of slightly less than one, they continue decreasing, until the auto-correlation becomes negative and starts showing an increasing negative auto-correlation. The decreasing auto-correlation is generally linear with little noise, indicating strong auto-correlation and high predictability.

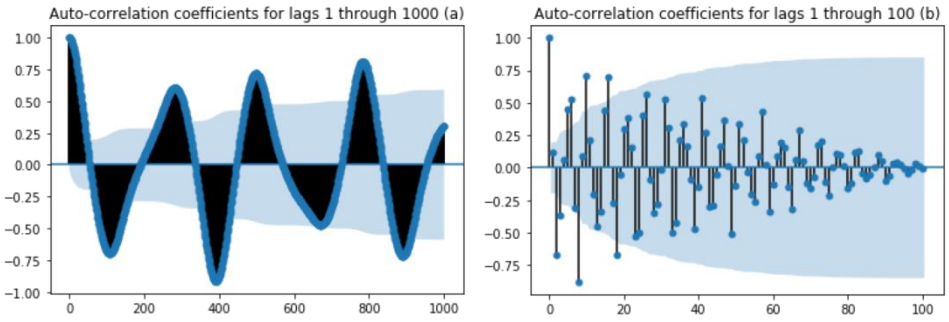


Figure 9, Reformed Auto-correlation I

*In figure 9, the initial condition was set as $\frac{\pi}{4}, \frac{\pi}{4}$. It was generated through the LSTM model and the time series was [w_1], and similar results were shown in [w_2] as well.

In figure 9(a), the author increased the lag k to 1000 to grasp the overview of the data pattern. Strong seasonality was shown in the plots, as the plot exhibited an alternating sequence of positive and negative spikes that are not decaying to zero. Apart from the obvious strong auto-correlation, the plot was actually also quite periodic.

In figure 9(b), the step was increased to 50 to gain a greater representation of the data. Each sequence of 20 now represented 1000 [x_1] values. The resulting graph showed an alternating sequence of positive and negative spikes decaying to zero after about 68 lags. The confidence interval included auto-correlation coefficients for lags is more than 20.

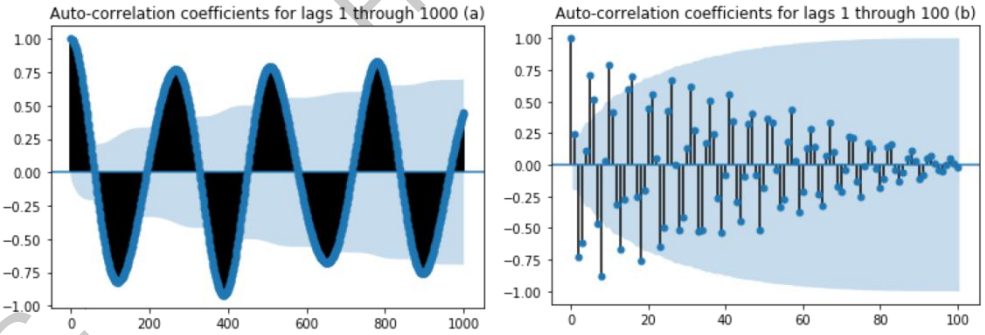


Figure 10: Reformed auto-correlation plots II

*In figure 10, the initial condition was set as $\frac{\pi}{4}, \frac{\pi}{4}$. It was generated through the LSTM model and the time series was [x_1]. Similar results were shown with the OED method.

In figure 10(a), one could see strong seasonality indicated by wave-like patterns that resemble figure 9(a). The only difference is that the x_1 time series has less noise than the angular velocity time series does.

In figure 10(b), the paper found a pattern similar to figure 9(b). The pattern is symmetric and reasonably sketches the landscape of figure 10(a). The absolute value of the auto-correlation coefficient winds down to zero almost linearly.

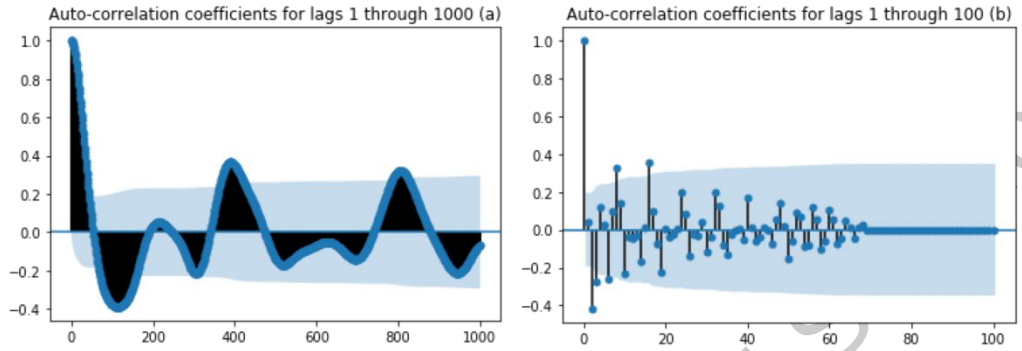


Figure 11: Reformed auto-correlation plots III

*****For figure 11, the initial condition was set as π, π. It was generated through the LSTM model and the time series was [x_1], and similar results were shown with the OED method.

The initial angles for figure 11's case would result in a α of 2π, which is much greater than the α for figure 10, namely $\frac{\pi}{2}$. One could observe that the pattern is less symmetric and periodic, but negative and positive values still alternate. The double pendulum's chaotic nature did add noises to the periodic pattern but the nature of oscillation remained.

In the 64 plots that the researcher scanned, it became obvious that the noise increased with α. In addition, the result indicated that the more chaotic the system is, the less lag k it needs to keep auto-correlation coefficient 0 or stay within the confidence interval, making it easier to be white noise. When the lags were low, DP's parameters that describe its states hardly showed strong auto-correlation. Seasonality was shown through k=0 to k=1000, but the noise correlated with the chaotic behaviors of DP obscured the pattern.

In addition, the similar patterns and statistical conclusions revealed by the OED solution and the LSTM model indicated that the dataset was not white noise either and confirmed the effectiveness of the model in this new perspective.

5.2 Box-Pierce Test

5.2.1Method

$$Q=n*\sum_{i=1}^{k} r_i^2 \qquad (2)$$

For the Box-Pierce Test, if the underlying data set is merely white noise, the expected value of Q statistic is zero. For any given time series, looking up the p-value of the test statistic in the Chi-square distribution according to k degree of freedom would enable us to check if the Q statistics deviates from zero in a statistically significant manner.

In the paper, it is assumed that a p-value of less than 0.05 indicates a significant auto-correlation that cannot be attributed to chance and is not considered random.

5.2.2Results

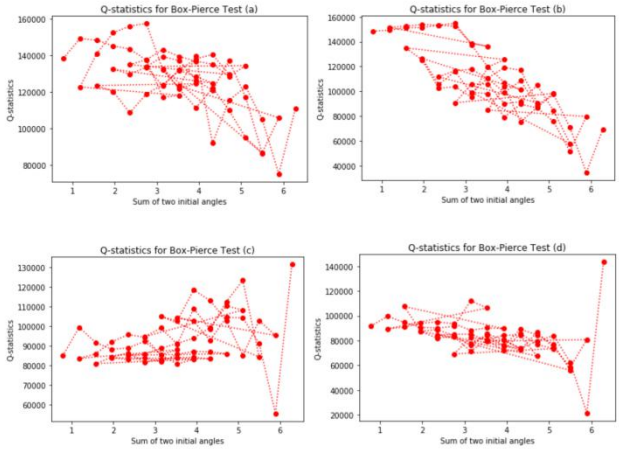


Figure 12: Box-Pierce Statistics Overview

*Figure 12 is the scatter plot of various Q-statistics according to different α. 12(a) is based on the time series of x_1, 12(b) is based on the time series of x_2, 12(c) is based on the time series of y_1, 12(d) is based on the time series of y_2.

After calculating the Box-Pierce Q-statistics for various initial conditions as described above, the paper came up with the following statistics: the mean for X_1 is 126665.29125594, for X_2 is 108893.17834141, for Y_1 is 93139.7486203, for Y_2 is 83181.82788792.

Except for the undistinguished pattern in figure 12(c), the rest revealed that higher α negatively correlates with Q-statistics. Because α is assumed to be positively correlated with the degree of chaos in a double pendulum, the more chaotic a DP system is, the fewer its Q-statistics would be.

Note that in fact these auto-correlations are strong, and the p-value is near to 0.00 for all the points above, indicating that the test-statistics are significantly greater than the Chi-square table value with corresponding degrees of freedom. Although one could still see the trend of decreasing Box-Pierce statistics, the p-value remains statistically small for all cases studied. A sound conclusion should incorporate such limitations.

5.3 Ljung-Box Test

5.3.1Method

The Ljung-Box test improves upon the Box-Pierce test and makes the distribution look closer to the Chi-square distribution. It is a diagnostic tool to test the lack of fit of a time series model.

$$Q=n(n+2)*\sum_{i=1}^{k}\frac{r_i^2}{n-k} \quad (3)$$

The paper conducted the test through the Statsmodel library and attained the test statistics and the p-value.

Just with steps described in Section 5.3, the Q-statistics for the Ljung-Box Test was calculated under the same set of initial conditions.

5.3.2Results

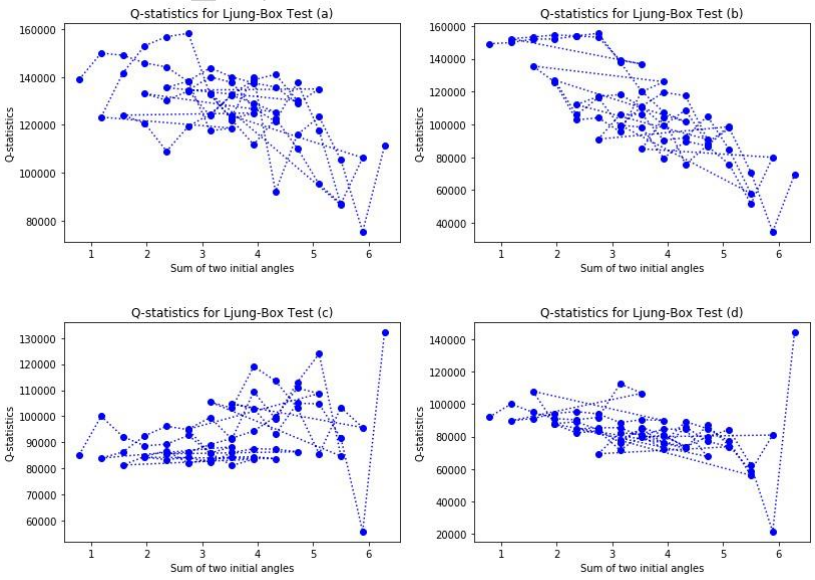


Figure 13: Ljung-Box Statistics Overview

*Figure 13 is the scatter plot of various Q-statistics according to different α. 13(a) is based on the time series of x_1, 13(b) is based on the time series of x_2, 13(c) is based on the time series of y_1, 13(d) is based on the time series of y_2.

After calculating the Ljung-Box Q-statistics for various initial conditions as described above, the research came up with the following statistics: the mean Q-stats for X_1 is 127130.95881355, for X_2 is 109263.12955357, for Y_1 is 93421.88835164, for Y_2 is 83424.46384691. The p-values are so small that the computer algorithm returned an approximate value of 0.00, regardless of the initial conditions. Since 0.00 is less than 0.05, the paper could reject the null hypothesis that the residuals were independently distributed.

For each time series, the percentage difference between the two tests generally stays within $3.7*10^{-3}$, which explains the similarity in the patterns of the two test plots and confirms the analysis in Section 5.2.

5.4 Random Walk Check

5.4.1Method

Although the time series analyzed above may seem to not be white noise, there is a possibility that they were random walks. Random walks are in fact highly auto-correlated, and their auto-correlated detected by the above tests may make us conclude that the random walks are not white noise.

To avoid making the mistake, the author would create a new list of data that represents the differences of angular velocity w_k and w_{k+1}. White noise tests would run on the differenced time series.

5.4.2Results

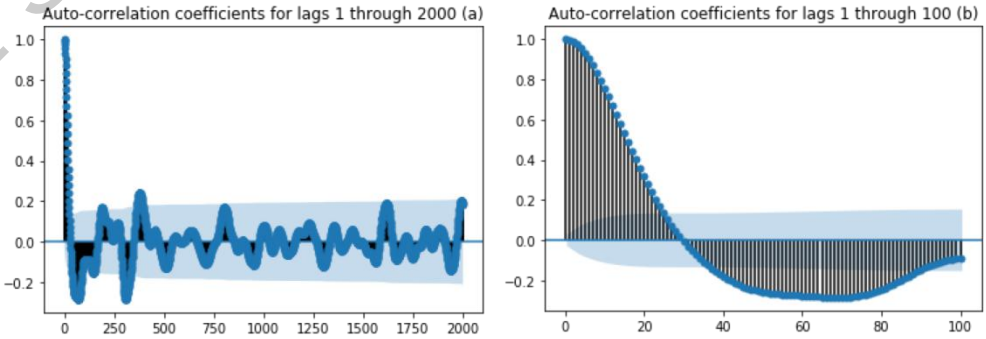


Figure 14: Auto-correlation Coefficients for differenced series

***Figure 14 is created under initial angles (Pi/4, Pi/4), based on the time series of angular velocity w_1.**

As one could perceive, the differenced series also showed patterns of alternating positive and negative spikes and a strong auto-correlation, which confirmed the findings in Section 5.1, 5.2 and 5.3. However, looking at more lags (figure 14(a)), most of the data was included within the blue range, the 95% confidence interval, demonstrating that the differenced series approaches more to being white noise. Hence, the differenced series can be maneuvered for simulating random digits or other further applications.

5.5Time Series Decomposition

In this section, the time-series data generated previously will be decomposed into different components to better capture their patterns.

5.5.1Method

The time-series data could be considered as a result of additive or multiplicative combinations of the noise, seasonal, cyclical, and trend components[n]. To model the time-series, let's define Y_n as the value of the time series, T_n as the trend component, S_n as the seasonal component, N_n as the noise component, all at the n^{th} step.

In an additive model, the time-series data could be represented as:

$$Y_n = T_n + S_n + N_n$$

In a multiplicative model, the time-series data could be represented as:

$$Y_n = T_n S_n N_n$$

The additive model is chosen, because negative values exist in the time-series data.
The seasonal component is assumed to have a length of 400ms and thus frequency of 400. 400ms centered moving averages will be calculated. The moving average transformation will highlight the trend component, which will then be subtracted and divided (according to the two models) to retrieve the seasonal and noise components. By calculating the average value of seasonal components for different 400ms duration, the noisy seasonal component could be acquired. Since the noisy seasonal component includes both the noise component and the averaged out seasonal component, the noisy seasonal component will be divided or subtracted by the averaged out seasonal component to gain the noise component.

5.5.2Results

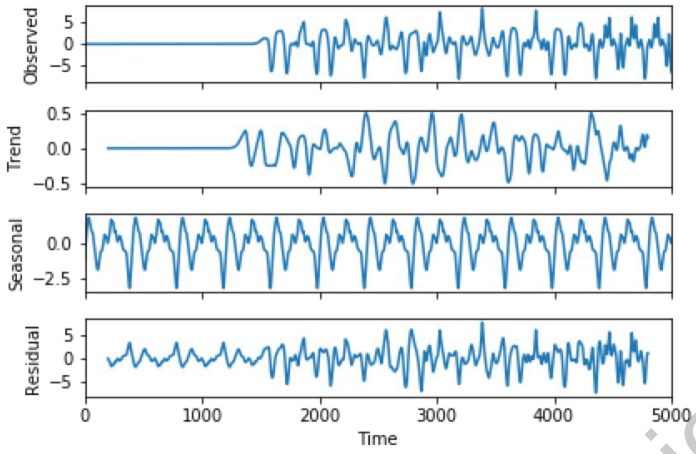


Figure 15: Separation of Components

The seasonal component accounts for the periodic ups and downs, and different seasonal periods often superimpose to form a time series. As one could observe, the seasonal component is periodic and fluctuates from a range of -3 to 2. It implies that normalizing the data to -1 to 1scale may help design a novel probability algorithm.

The trend component stands for the pattern that ranges all information across seasonal periods. Unlike other time-series data, the chaotic data had a relatively unpredictable range. If such characteristics is applied in cryptography, it will be much hard to decipher a code.

The cyclical component represent patterns that do not have fixed period but happen across seasonal periods, which will has included in the trend component due to the difficulty of separation. Hence, it is reasonable to infer that the variations in the trend comes from the cyclical component.

The noise, or random component, is what remains after the other components are all separated. In the chaotic simulation, the noise includes the unknown effects of the known and unknown factors in the double pendulum.

6. DISCUSSION

The research has been mostly successful and conclusive analysis has been given, yet several improvements could still be made. First, normalizing the data on a MIN-MAX scale may help LSTM model to converge faster. Second, comparing the effect of LSTM and OED on same

parameters and initial conditions could better quantify the effectiveness of the LSTM model designed in the paper. Third, the findings were limited to the fact that the p-value remained 0.0 throughout the experiment.

There are several directions to further the research as well. Maneuvering the time-series data according to their chaotic characteristics could generate random data that approaches to true randomness. In addition, utilizing the noise components to encrypt important codes is also expected to reach good results. If one seeks to analyze the problem beyond the initial angle, exploring the effect of the initial velocity, which has been set as 0 throughout the paper, will suffice.

7. CONCLUSION

Contradictory to what many may believe, chaos does not directly lead to randomness. The paper designed an effective LSTM model to generate data for statistical analysis of the random nature in the time series of the Double Pendulum. With rigorous approaches of analyzing the input, this paper provided different cases, simulated extensive scenarios, conducted valid tests, and discovered important and general trends, separated different components, and concluded the results as follows: 1) (the higher the α) the more chaotic the system is, the greater noise it has. 2) Periodic traits do not contradict chaotic traits, because the chaotic behaviors still show characteristic seasonality. 3) The time-series data that describe the state of a typical double pendulum are strongly auto-correlated. 4) The more chaotic a DP system is, the more likely it is to be converted to a white noise series. These are important first-hand discoveries that shed light on the use of chaotic theory that could be applied to design novel learning algorithms, detect cyber poisoning attacks, or create new cryptographic ciphers.

8. ACKNOWLEDGMENTS

It is a pleasure to thank professor Mario from UCL, whose lectures on mechanics and machine learning supported me to grow a solid understanding of the field of physics. In addition, I would also like to present my gratitude to Qiang Zheng, who gave valuable suggestions that enhanced the paper's academic rigor and depth. Finally, I would thank my instructor, Yining Liu. As the computer science and statistics teacher of my school, she guided me in brainstorming and provided me with useful research strategies. In the whole process of researching half-a-year, I proposed the topic, conducted research, run experiments, and analyzed the results entirely on my own. However, without the gracious professionalism and support from these honorable people, my paper could not be accomplished with such degree of completion.

9.REFERENCES

[1]Anderson, R. L. (1942). Distribution of the serial correlation coefficient. *The Annals of Mathematical Statistics*, *13*(1), 1–13. https://doi.org/10.1214/aoms/1177731638

[2]Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time-series. *Supplement to the Journal of the Royal Statistical Society*, *8*(1), 27. https://doi.org/10.2307/2983611

[3]Box, G. E., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, *65*(332), 1509–1526. https://doi.org/10.1080/01621459.1970.10481180

[4]Brownlee, J. (2020, August 14). *White noise time series with python*. Machine Learning Mastery. https://machinelearningmastery.com/white-noise-time-series-python/.

[5]Carl W. Akerlof (2012, September 26) The Chaotic Motion of a Double Pendulum. *UMICHEDU* http://instructor.physics.lsa.umich.edu/adv-labs/Chaotic%20Double%20Pendulum/Pendulum_2012_09_26.pdf

[6]Dalla, V., Giraitis, L., & Phillips, P. C. B. (2019, April). *Robust tests for white noise and cross-correlation*. COWLES FOUNDATION DISCUSSION PAPER NO. 2194R. https://cowles.yale.edu/sites/default/files/files/pub/d21/d2194-r.pdf.

[7]Hyndman , R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice (2nd ed)*. 2.9 White noise. https://otexts.com/fpp2/wn.html.

[8]James, F. (1995). Chaos and randomness. *Chaos, Solitons & Fractals*, *6*, 221–226. https://doi.org/10.1016/0960-0779(95)80028-f

[9]Klinkachorn , S., & Parmar , J. (n.d.). *Evaluating Current Machine Learning Techniques On Predicting Chaotic Systems*. STANFORD EDU. http://cs229.stanford.edu/proj2019spr/report/38.pdf.

[10]LJUNG, G. M., & BOX, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, *65*(2), 297–303. https://doi.org/10.1093/biomet/65.2.297

[11]Nevillle, J., S， im， sek O¨ zgu¨ r, & Jensen, D. (2004). *Autocorrelation and Relational Learning: Challenges and Opportunities*. CS.PERDU.EDU. https://www.cs.purdue.edu/homes/neville/papers/neville-et-al-srl2004.pdf.

[12]Nouar, C., & Guennoun, Z. E. A. (2020). A pseudo-random number generator using double pendulum. *Applied Mathematics & Information Sciences*, *14*(6), 977–984. https://doi.org/10.18576/amis/140604

[13]Ohlhoff, A., & Richter, P. H. (2000). Forces in the double pendulum. *ZAMM*, *80*(8), 517–534. https://doi.org/10.1002/1521-4001(200008)80:8<517::aid-zamm517>3.0.co;2-1

[14]Quenouille, M. H. (1949). The joint distribution of serial correlation coefficients. *The Annals of Mathematical Statistics*, *20*(4), 561–571. https://doi.org/10.1214/aoms/1177729948

[15]Ruelle, D. (1980). Strange attractors. *The Mathematical Intelligencer*, *2*(3), 126–137. https://doi.org/10.1007/bf03023053

[16]Shinbrot, T., Grebogi, C., Wisdom, J., & Yorke, J. (1992). Chaos in a double pendulum. *American Journal of Physics*, *60*(6), 491–499. https://doi.org/10.1119/1.16860

[17]Stachowiak, T., & Okada, T. (2006). A numerical analysis of chaos in the double pendulum. *Chaos, Solitons & Fractals*, *29*(2), 417–422. https://doi.org/10.1016/j.chaos.2005.08.032

[18]Strogatz, S. (2010, March 7). *Finding your roots*. The New York Times. https://opinionator.blogs.nytimes.com/2010/03/07/finding-your-roots.

[19]Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series?--a study in sampling and the nature of time-series. *Journal of the Royal Statistical Society*, *89*(1), 1. https://doi.org/10.2307/2341482