University of Exeter
Department of Mathematics

Time series analysis and modelling of the freezing of gait phenomenon

Ai Wang

Submitted by Ai Wang, to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Mathematics, October, 2023.

I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for the award of a degree by this or any other University has been acknowledged.

Signed: Ai Wang. . . . .

Abstract

Freezing of Gait (FOG) is one of the most debilitating symptoms of Parkinson's Disease and is associated with falls and loss of independence. The patho-physiological mechanisms underpinning FOG are currently poorly understood. In this thesis we combine time series analysis and mathematical modelling to study the FOG phenomenon's dynamics. We focus on the transition from stepping in place into freezing and treat this phenomenon in the context of an escape from an oscillatory attractor into an equilibrium attractor state. We analyze the experimental data by two different approaches. In the first approach we use a stochastic Hopf bifurcation normal form model to study the escape time from oscillatory behavior to small-amplitude fluctuations. For the other approach we extract a discrete-time discrete-space Markov chain from experimental data and divide its state space into communicating classes to identify the transition into freezing. This allows us to develop a methodology for computationally estimating the time to freezing as well as the phase along the oscillatory (stepping) cycle of a patient experiencing Freezing Episodes (FE). The developed methodology is general and could be applied to any time series featuring transitions between different dynamic regimes including time series data from forward walking in people with FOG.

Acknowledgements

As a student on the path of learning mathematics, what I have been thinking about is neither the meaning of mathematics for human beings, nor how to delve into more advanced mathematical theories. How to make a small contribution to the humanities and life sciences together with my good friend Mathematics is the original intention of me to study for a PhD. Like most people who aren't taking math as their major, combing math and life science is unimaginable at the first place, I would like to send my most sincere thanks to my primary supervisor, Prof Krasimira Tsaneva-Atanasova. She was the one who opened the doors to this fascinating, vivid, and logical world of mathematics for me. Her unrelenting dedication to scientific research, her rigorous professional ethos, and her commitment to weekly meetings in the face of a demanding schedule have left a deep impression on me.

I am equally indebted to my second supervisor, Prof Jan Sieber, my freshman mathematics programming teacher. Despite the immense challenges of your class due to my lack of prior exposure to computer-related courses, your abundant guidance laid the foundation for my programming knowledge. Your endorsement encouraged me to pursue a PhD and accept me as your student. Over the years, you have patiently explored mathematics with me, tutored me in programming, assisted me in my research, extensively edited my initial attempts at writing, and imparted invaluable life lessons. You are not merely an academic mentor but a guide for life.

In recognition of their support during my PhD journey, I wish to express my appreciation for Jen Creaser, Kyle Wedgwood, Vadim Biktashev, Jennifer Catto, Peter Ashwin, Andrew Gilbert, Harry Green, Julian Newman, Ana Rodrigues, Congping Lin, Zhongkai Tao, Young Will, Mark Holland and Zhi Zhang.

The role of my family has been instrumental in my accomplishments. I want to express my deepest gratitude to my parents Yueming Wang and Yufang Miao, my sister Ke Miao, and my grandparents. Without your unwavering support, tireless sacrifice, and loving companionship, I would not be where I am today. Thank you, my boyfriend Chengjie Tang, for appearing in my life. I wish you all the best in your PhD journey. Lastly, I wish to acknowledge my beloved pets Xiaohu, ZhuYou, Sixteen, Huahua, and Bayi. Your presence has filled my life with joy and love.

Contents

List of Figures

List of Tables

1. Introduction

People with Parkinson's Disease (PD) will often walk with reduced gait speed and shorter stride length. Spatial and temporal characteristics of Parkinsonian gait are also typically more variable compared to age-matched controls [3, 4]. Furthermore, approximately 50% of people with advanced Parkinson's will experience freezing of gait (FOG) [5]. Patients describe FOG as the sensation that their feet are glued to the floor, preventing them from initiating a new step. Indeed, a greater variability in walking patterns has been observed in Parkinson's patients with FOG. This is characterised by increased step coefficient of variation [6], asymmetry, rhythmicity [7] and difficulty coordinating [8], compared to patients without freezing. These studies focus on descriptive statistics of the walking and/or stepping time series. In this study, in addition to performing advanced time series analysis we concentrate on revealing the dynamic (geometric) properties of the transitions between stepping and freezing.

Dynamic modelling and analysis can help in understanding and characterising specific features and properties of Parkinsonian gait. This in turn could inform future rehabilitation and prevention interventions as well as strategies that people with Parkinson's might benefit from through informal use in daily life. A variety of mathematical modelling and data analysis approaches have been applied in the context of Parkinson's gait and motor control more generally as reviewed in [9]. Data-driven prediction and detection of the FOG phenomenon have been extensively addressed in previous work, including by us in the context of stepping in place force platform data [10], and others (see Table 1 in [11] for list of studies based on wearable sensors data [11]), and more recently plantar pressure data [12, 13]. A stochastic model of gait consisting of a random walk on a chain has been also proposed and applied to describe alterations in gait dynamics from childhood to adulthood [14]. To our knowledge none of the studies to date have combined mathematical modelling and data analysis to investigate dynamic properties of the FOG phenomenon in Parkinson's Disease, which is the main focus of our work.

Summary of results We analyse the transition between stepping and freezing observed during stepping-in-place experiments performed by Nantel et al [7]. In our analysis, after phase space embedding, stepping motion appears as large-amplitude oscillations while freezing appears as an approximate equilibrium or irregular drifting. We explore whether there exists a preferred phase at which the stepping trajectory escapes from the regular periodic stepping behavior to transition into freezing. The first step toward an answer to this question is to identify for each freezing event a location in phase space (after embedding) at which the transition occurs. At first sight this appears to depend strongly on the parameter choices in our data processing, such as, for example, a threshold for freezing that we could define. To address this sensitivity we seek to develop a recurrence-based method for identifying a unique phase and amplitude of the point at which the transition occurs that does not depend on method parameter choices. Our method consists of three steps. First, we apply a Hilbert Transform to reconstruct a two-dimensional embedding from the scalar experimental times series data as a signal in the complex plane. Second, we construct a Markov chain by discretizing the complex plane and counting the empirical transition probabilities of the Hilbert Transform output between the boxes from the discretization. Over a large range of method parameters the Markov chain has a clearly identifiable largest communicating class corresponding to regular periodic stepping (the *stepping class*.) In the final step we determine which boxes have a large transition probability out of the stepping class. We take the location (phase and amplitude) of these boxes in the complex plane as the threshold for transition to freezing. With the help of this naturally emerging threshold we obtain a mean escape time characterising the transition from stepping to freezing. We can also investigate if these transitions occur independent of phase or in some range of preferred phases (angles) along the oscillatory cycle preceding the freezing episodes.

This thesis presents the method for identifying the transition's phase and amplitude for each event, demonstrates that it is robust with respect to choice of method parameters, and tests it on the small patient data set that we have currently available. In principle, pinpointing the transition will allow us to determine whether transitions occur preferably at universal or patient-specific phase ranges, or whether they are uniformly distributed (we treat the latter as the null hypothesis). This would help us to further understand the underlying dynamics of FOG.

This thesis is organized as follows. This chapter presents the general motivation of the research. From the next chapter (Chapter 2), detailed related works will be thoroughly surveyed. After describing the type of data we use for our method development in Chapter 3 Section 3.1, we present our null model of the transition between stepping and freezing in Chapter 3 Section 3.2. Chapter 5 then describes our method to construct a Markov chain from the data and the analysis of the Markov chain's properties we expect for our data. The new algorithm we develop is applied to patients' stepping data in Section 5.3 and our findings regarding phase dependence of the transition into freezing are given in Chapter 5 Section 5.4. In Chapter 6 we study the robustness of the Markov chain model we designed in Chapter 5. Chapter 7 follows with conclusions and outlook.

2. Background on mathematical techniques

In the literature there are several attempts at creating models for gait from first principles, describing the neural system and its connection to body mechanics (reviewed in Chapter 1). An alternative approach to modelling is to use data and construct simple abstract models that fit the data. These modeling techniques rely on reconstructing the state of an abstract dynamical system from time series by *embedding*. This chapter introduces embedding techniques(delay embedding, Hilbert Transform and Koopman operators) and associated notation we will use later. We also describe the False-Nearest Neighbor criterion, which helps decide if the dimension of the embedding space is large enough. Since one of our target model classes are discrete-space Markov chains we also introduce some basic concepts for Markov chains.

2.1 Embedding Methods

2.1.1 Delay Embedding

When modelling the gait of a patient (for example with ordinary differential equations), a sensible model would contain state variables that account for the measurement outputs collected in the experiments. One technique for reconstructing a higher-dimensional state from one or more measurements is delay embedding [15]. This method naturally arises in the representation of the evolution of partially observed states in dynamical systems. The Takens Embedding Theorem is a fundamental result in dynamical systems theory that provides a theoretical basis for reconstructing the state space of a dynamical system from time series data [16]. Assume that the underlying dynamical system is governed by an autonomous ordinary differential equation (ODE) $x'(t) = f(x(t))$ with $x(t) \in \mathbb{R}^n$ of some dimension n, and that the measurement is some function of the state $y(t) = g(x(t))$, where the dimension of y is lower than n. Let us assume $y(t) \in \mathbb{R}^1$ for illustration. The state variable constructed by delay embedding is $\mathbf{X}(t) = \{y(t), y(t+\tau), y(t+2\tau), ..., y(t+(N-1)\tau))\}$, where N is the embedding dimension and τ is the delay time. In experiments data is often sampled at discrete times with time step δ_t, resulting in a sequence y_i of measurements taken at times $t_i = t_0 + i\delta_t$. In this case a delay embedding is of the form

$$X_i = (y_i, y_{i+\tau}, \dots, y_{i+(N-1)\tau}) \tag{2.1}$$

where the embedding delay τ is an integer. In this case the real time delay is $\tau\delta_t$.

Criterium for the delay time τ We demonstrate the effect of τ for the simple example ODE $x''(t) = -\omega^2 x(t)$ to show a possible criterion for choosing a delay time τ. We choose as output $y(t) = x(t)$, which is in $\mathbb{R}^1$. Figure 2.1 shows the trajectory $X(t) = (y(t), y(t-\tau)) = (x(t), x(t-\tau))$ after delay embedding of $x''(t) = -\left(\frac{\pi}{50}\right)^2 x(t)$ in 2 dimensions for different values of τ. We can see that in this case the time delay embedding with $\tau = T/4$, where T is the period with $T = 2\pi/\omega$, produces a nice circle instead of a sharp ellipse.

In many applications (also in ours as Figure 3.1 shows) the output $\mathbf{X}(t)$ (or $\mathbf{y}(i)$) is a combination of oscillations with different amplitudes and frequencies. A good embedding would be one where the most prominent part of the oscillations follows nearly a circle in the phase space. This criterion can be made precise using the Fourier transform of the time series.

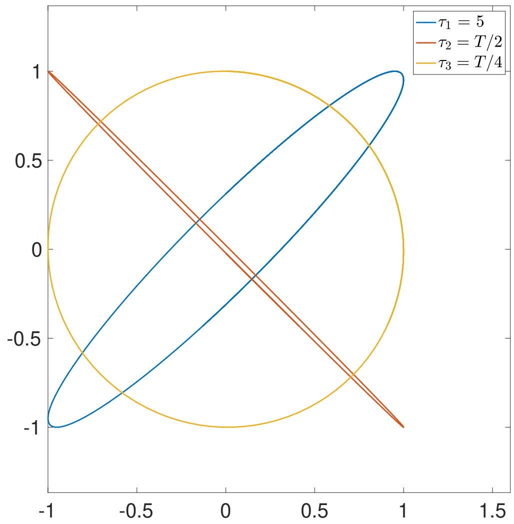


Figure 2.1: The trajectory after delay embedding of $x''(t) = -\left(\frac{\pi}{50}\right)^2 x(t)$ for delay time $\tau = 5$ (blue), $T/2$ (red), $T/4$ (orange) respectively.

2.1.2 Embedding based on the Hilbert Transform

Notation of Hilbert Transform and basic properties The Hilbert Transform (HT) arose in Hilbert's 1905 work on a problem Riemann posed concerning analytic functions [17, 18] which has come to be known as the Riemann–Hilbert problem. Hilbert's work was mainly concerned with the Hilbert transform for functions defined on the circle [19, 20]. The Hilbert transform has been extensively used to study phase demodulation in synchronization studies [21, 22]. It could be also used to perform a two-dimensional embedding of an observed one dimensional signal. This embedding then is usually treated as an analytic signal and the phase is extracted as the argument of the complex state (or, more generally, as an angle of rotation with respect to some point on the embedding plane). The Hilbert transform is an integral transformation (similar to the Laplace and Fourier transforms), which is introduced by evaluating an integral under special integral conditions in mathematics and physics:

$$H\left[x\left(t\right)\right]=\tilde{x}(t)=\pi^{-1}\int_{-\infty}^{\infty}\frac{x\left(\tau\right)}{t-\tau}\mathrm{d}\tau \tag{2.2}$$

We refer to the resulting trajectory $(x(t),\tilde{x}(t))$ in the plane as Hilbert embedding of $x(t)$. Note that we have to consider this integral as the Cauchy principal value [23], which permits assigning a value to the integral despite the singularity at $\tau=t$ and $\tau=\pm\infty$. From the mathematical definition alone, it is difficult to understand the physical meaning of transformation. Physically, HT is equivalent to a special linear filter. All the spectral amplitudes have not changed, but the phase is shifted $-\frac{\pi}{2}$ which is defined by convolution.

Hilbert Transform to complex signal Hilbert transform can be used to provide a signal amplitude, instantaneous phase and frequency information for a (real) scalar oscillatory time series. To estimate these characteristic attributes of a time series $x(t)$ coming from a measurement obtained over a finite period of time, we construct $X(t)=x(t)+\mathrm{i}\tilde{x}(t)$, where the real part is the original time series and the imaginary is the Hilbert Transform of $x(t)$, as defined in (2.2) [1]. In order to extract the original time series $x(t)$ from the complex signal $X(t)$, one has to use a substitution $x(t)=0.5[X(t)+X^{*}(t)]$, where $X^{*}(t)$ is the complex conjugate signal of $X(t)$.

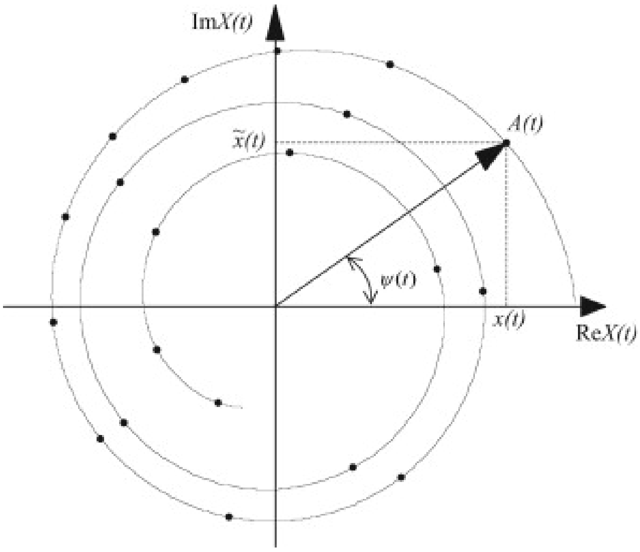


Figure 2.2: The signal in the complex plane after Hilbert Transform [1].

According to the analytic signal theory, a real oscillatory process $x(t)$, measured by, say, a transducer, is only one of the possible projections (the real part) of some analytic signal $X(t)$. Then the second projection of the same signal (the imaginary part) $\tilde{x}(t)$ will be conjugated according to the HT. An analytic signal has a geometrical representation in the form of a phasor rotating in the complex plane [1], as shown in Figure 2.2. A phasor can be viewed as a vector rooted at the origin of the complex plane with a length $A(t)$ and an angle $\psi(t)$ (see Figure 2.2), which are the polar coordinates for $X(t)$. The projection on the real axis is the original real signal, recoverable from the polar coordinates by $x(t) = A(t)\cos\psi(t)$. The polar coordinates of the complex signal $X(t)$,

$$X(t) = |X(t)|\left[\cos\psi(t) + i\sin\psi(t)\right] = Ae^{i\psi(t)},$$

construct an instantaneous amplitude (envelope, magnitude, modulus) for the original signal $x(t)$. The instantaneous amplitude of $x(t)$ is

$$A(t) = \pm |X(t)| = \pm \sqrt{x^2(t) + \tilde{x}^2(t)} = \pm \mathrm{e}^{\mathrm{Re}[\ln X(t)]},$$

and the instantaneous phase of $x(t)$ is

$$\psi(t) = \arctan \frac{\tilde{x}(t)}{x(t)} = \mathrm{Im}\left[\ln X(t)\right].$$

The change of coordinates from rectangular $(x, \tilde{x})$ to polar (A, ψ) gives $x(t) = A(t)\cos(\psi)$, $\tilde{x}(t) = A(t)\sin(\psi)$. For signals where an embedding dimension $N = 2$ is enough the Hilbert transform automatically picks the optimal delay when using it for the example (see Figure 2.1) we discussed in the delay embedding.

2.2 The False Nearest Neighbour (FNN) criterion

Takens' Delay Embedding Theorem [24] states that one can reconstruct trajectories of a n-dimensional ODE $\dot{x}(t) = f(x(t))$ from the time series of an observation $y(t) = g(x(t)) \in \mathbb{R}$ and its delayed values. More precisely, there exists an invertible map between the attractor $\mathscr{A}$ of the ODE and the set $\{g(x(t)), g(x(t-\tau)), \ldots, g(x(t-(m-1)\tau)) : t \in \mathbb{R}, x(\cdot)$ trajectories on attractor $\mathscr{A}\}$. For this to hold the dimension m should be sufficiently large and g and τ have to satisfy some technical conditions. For example, if the ODE consists of two uncoupled systems and the observation depends only one one of the subsystems then the other subsystem's state cannot be reconstructed [ref to Kantz book]. This theoretical result motivates the development of methods for determining the necessary dimension m from a time series of measurements when one suspects that a model for the underlying dynamical system from which the measurements come can be approximated by a low-dimensional ODE. One such method is the False Nearest Neighbour method proposed by Kennel *et al.* [15]. This is a method to estimate the necessary embedding dimension N (which is the minimum number of variables required to accurately reconstruct the phase space of a dynamical system from its time series data) for a scalar time series with the so-called False Nearest Neighbour (FNN) method to minimise the number of variables required to accurately reconstruct the phase space of a dynamical system. From a geometric point of view, a time series is the projection of the trajectories into a one-dimensional space from a high-dimensional phase space. In the process of projection, the trajectory will be distorted. When two points that are not adjacent in the high-dimensional phase space are projected on a one-dimensional axis, they may become two adjacent points, such that in the projection points may be close to each other that were not close in the phase space. We call them ”false neighbours”.

Reconstructing the phase space tries to recover the trajectory from the time series. As the embedding dimension N increases, the trajectory of chaotic motion will gradually unfold, and the false neighbours will also be gradually eliminated, so that the trajectory will be restored.

In a N-dimensional phase space each phase point $\mathbf{y}(i)$ from ((2.1)) has a nearest point $\mathbf{y}^{nbh}(i) = \{x^{nbh}(i), x^{nbh}(i+\tau), \ldots, x^{nbh}(i+(N-1)\tau)\}$ within a certain distance:

$$R_N(i) = ||\mathbf{y}(i) - \mathbf{y}^{nbh}(i)|| \tag{2.3}$$

The nearest neighbour relation is not necessarily symmetric. When the dimension of the phase space increases from N dimension to $N+1$ dimension, the distance between these two phase points will change to $R_{N+1}(i)$ which obeys the following rule [15]:

$$R_{N+1}^2(i) = R_N^2(i) + ||x(i+N\tau) - x^{nbh}(i+N\tau)||^2 \tag{2.4}$$

If $R_{N+1}(i)$ is much larger than $R_N(i)$, this contradicts the assumption that the evolution in the phase space is uniquely determined by the current state. Thus, the increase from R_N to R_{N+1} is evidence that the two points that were non-adjacent in the original phase space are adjacent when projected by the N-dimensional embedding of the output. The nearest point $\mathbf{y}^{nbh}(i)$ turned out to be a “false” nearest neighbour and the embedding dimension N is too small. Let us define the ratio between new-point distance and nearest-neighbour distance as

$$a(i,N) = \frac{||x(i+N\tau) - x^{nbh}(i+N\tau)||}{R_N(i)}. \tag{2.5}$$

We introduce a threshold R_ξ and count $\mathbf{y}^{nbh}(i)$ as FNN of $\mathbf{y}(i)$ if:

$$a(i,N) > R_\xi. \tag{2.6}$$

Kennel *et al.* [15] and Cao [25] suggest a value of threshold R_ξ between $[10,40]$.

The above method was designed assuming that the time series is the output of a (possibly chaotic) low-dimensional deterministic map or ODE. If this map or ODE is perturbed by noise and we have finite-length data of size M, we can also add the following rule [15]:

$$\text{if}, R_{N+1}(\mathrm{i})/R_A \geq 2, \tag{2.7}$$

then $\mathbf{y}^{nbh}(i)$ is the FNN of $\mathbf{y}(i)$, where

$$R_A^2 = \frac{1}{M}\sum_{i=1}^{M}[x(i)-\bar{x}]^2, \quad \bar{x} = \frac{1}{M}\sum_{i=1}^{M}x(i).$$

2.3 Markov chains and communicating classes

In this section we consider a discrete-space Markov chain with N states. Transition probabilities between states i and j are given in a transition matrix A in entry $A_{i,j}$. We let $\mathbf{u}_i$ be the ith unit row vector ($u_{ij} = 1$ if $i = j$, 0 otherwise).

Definition 2.3.1 (accessible)**.** A state j is said to be accessible from a state i (written $i \to j$) if there exists $n \geq 0$ such that $\mathbf{u}_i A^n \mathbf{u}_j^\top > 0$ [26].

Definition 2.3.2 (Communicating class)**.** A state i is said to communicate with state j (written $i \leftrightarrow j$) if both $i \to j$ and $j \to i$. The relation $\leftrightarrow$ is an equivalence relation. A communicating class is an equivalence class of $\leftrightarrow$, so, it is a maximal set of states G such that every pair of states in G communicates with each other [26].

Property 2.3.3 (Recurrent states of a Markov chain)**.** *Let f_i^n denote the probability of first return to state i at time n. Assuming we are in state i at time* 0*, let $f_i = \sum_{n=1}^{\infty} f_i^n$. Then f_i denotes the probability of eventual return to state i (starting in state i initially). If $f_i < 1$, then state i transient. If $f_i = 1$, then we say state i is recurrent [27].*

Property 2.3.4 (Positively recurrent states of a Markov chain)**.** *For recurrent states, the mean return time is given by $\mu_i = \sum_{n=1}^{\infty} n f_i^n$. If $\mu_i < \infty$, then state i is said to be positively recurrent. If $\mu_i = \infty$, then state i is called null recurrent [27].*

Definition 2.3.2 uses that states communicating with each other ($i \leftrightarrow j$) is an equivalence relation, which implies that one can partition the discrete state space into equivalence classes (which are then called communicating classes). We will use the indicator matrix to partition the states of a chain into communicating classes since we can learn much about the long-time behavior of a Markov chain merely from the indicator matrix of its transition matrix. Let us illustrate the partition process with an example. Assume that the Markov chain has 7 states and transition matrix A_1:

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/9 & 8/9 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 1/2 & 1/8 & 1/8 & 0 \\ 0 & 0 & 0 & 0 & 4/5 & 1/10 & 1/10 \\ 0 & 0 & 0 & 0 & 3/10 & 7/10 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The indicator matrix The indicator matrix of a transition matrix is constructed by setting each positive element to 1 while the other elements remain 0. More formally, given a transition matrix A, the elements of its indicator matrix Z are given by:

$$Z_{ij} = \begin{cases} 1 & \text{if } A_{ij} > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Thus, while A is a real-valued matrix (with entries in the interval $[0,1]$), Z is a binary matrix (with entries equal to 0 or 1). The transition diagram for the indicator matrix includes a directed edge from state i to state j if and only if $Z_{ij} = 1$. The indicator matrix

Z_1 of matrix A_1 is:

$$Z_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Definition 2.3.5 (Reachability)**.** State i can reach state j if $i = j$ or if there is a path of any length from i to j in the transition diagram.

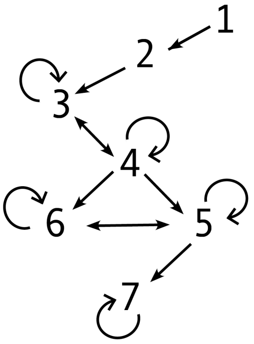


Figure 2.3: Transition diagram of matrix A_1

Figure 2.3 shows the transition diagram of matrix A_1. For simple cases, we can easily determine reachability (shown by adjacency matrix of a directed graph calculated from powers of Z) by inspection of the transition diagram. However, when the number of states is large, matrix methods are more convenient. When raising the matrix A to the power n for $t \geq 0$ (using the convention $A^0 = I$, the identity matrix), the entry A^n_{ij} gives the number of paths of length n from i to j. To characterize the reachability relation using matrix methods (powers of A), we may thus construct a binary reachability matrix

$Z_{\infty,ij}$ as

$$Z_{\infty,ij} = \begin{cases} 1 & \text{if } Z^n_{ij} > 0 \text{ for some } n \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

We observe we only need to check powers of n for Z that are less or equal to $n_{\max}$, where $n_{\max}$ is defined as the power where the number of non-zero entries in the matrix $\sum_{n=0}^{n_{\max}} Z^n$ equals that of $\sum_{n=0}^{n_{\max}+1} Z^n$. The Matlab function `transclosure` finds the reachability matrix $Z_{\infty,ij}$ for arbitrary transition matrices. For example, the reachability matrix $Z_{\infty 1}$ of matrix A_1 is

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

Communicating classes Since we can determine the reachability, we can now partition the states of the Markov chain into communicating classes, assigning states i and j to the same class if and only if each of these states can reach and be reached by the other. We construct a binary bidirectional reachability matrix $Z_{\text{bi}} = \min(Z_\infty, Z_\infty^T)$:

$$Z_{\text{bi},ij} = \begin{cases} 1 & \text{if } Z_{\infty,ij} = 1 \text{ and } Z_{\infty,ji} = 1, \\ 0 & \text{otherwise.} \end{cases}$$

If we draw the graph using $Z_{\text{bi},ij}$ as adjacency matrix, the communicating classes would appear as isolated groups (or clusters) that are themselves fully connected. That is, node i can reach and be reach by node j if and only if i and j belong to the same group. Consequently, if states i and j belong to the same communicating class, then

$$Z_{\text{bi},ik} = Z_{\text{bi},jk} \text{ for all } k \in \{1,...,n\}.$$

The binary bidirectional reachability matrix Z_{bi1} of matrix A_1 is:

$$Z_{\text{bi1}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

To determine the membership of each communicating class, we may thus list the unique rows of the matrix Z_{bi} or Z_{∞}. Each row of this matrix characterizes the membership of a different communicating class. We label this unique communicating class indicator matrix as matrix $U \in \{0,1\}^{u \times n}$ in ascending order, the column number n corresponding to the state number and the row number u are the total number of the communicating classes of matrix A. The unique matrix U_1 of matrix A_1 is

$$U_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The matrix U_1 indicates the communicating classes of matrix A_1 are $\{1\}, \{2\}, \{3,4\}, \{5,6\}, \{7\}$.

2.4 Koopman Operator

This section defines what the Koopman operator is and sketches the underlying idea very briefly, but that we ended up not using the technique. However, later in Section 5 we construct Markov chains based on empirical transition matrices, and the transpose of such Markov-chain transition matrices are approximations of the Koopman operator.

Koopman-operator based data analysis The problem of identifying governing equations of dynamical systems from time-series data has attracted considerable interest in many fields such as biology, engineering, and finance. It is also closely related to network inference, which aims to reconstruct the interactions between different components of a system, a problem of paramount importance in biomedical applications [28, 29]. In many cases the identification problem is challenging due to the nonlinear nature of the systems which need to deal with long, highly sampled time series or sparsely sampled time series and provide a relationship between the system inputs and outputs.

Recently, there is growing interest in operator-theoretic approaches for the analysis of dynamical systems. Operator-theoretic approaches are based on the Perron–Frobenius operator or its adjoint, the Koopman operator (or composition operator) [30]. The core concept of Koopman operator analysis is that the evolution of a nonlinear dynamical system can be expressed using infinite-dimensional linear operators. This method was first proposed by Koopman in 1931 [30], after 2004, Mezic and his collaborators [31, 32] discussed the method, which was applied to reduce the order of stochastic or deterministic models, and found that the spectral decomposition of Koopman operators is related to the spatial mode of the dynamical system [33].

Koopman operators for discrete-time dynamical systems Since we are concerned with time series analysis, it is natural to consider the setting of discrete-time nonlinear dynamical system:

$$\mathbf{x}_{k+1} = \mathbf{F}_t(\mathbf{x}_k) \tag{2.8}$$

In Eq(2.8) $\mathbf{x}$ is the state variable. The subscript t is the sampling time step for the flow map $\mathbf{F}$.

The Koopman operator K_t is an infinite-dimensional linear operator that acts on observable function g as:

$$K_t g = g \circ \mathbf{F}_t \tag{2.9}$$

where $\circ$ is the composition operator, so that:

$$[K_t g](\mathbf{x}_k) = g\left(\mathbf{F}_t\left(\mathbf{x}_k\right)\right) = g(\mathbf{x}_{k+1}) \tag{2.10}$$

In other words, the Koopman operator K_t defines an infinite-dimensional linear dynamical system that advances the observation $g_k = g(\mathbf{x}_k)$ of the state $\mathbf{x}_k$ to the next timestep:

$$g(\mathbf{x}_{k+1}) = K_t g(\mathbf{x}_k). \tag{2.11}$$

Since for any observable function g_1, g_2 and coefficient α, β, we have $K_t(\alpha g_1 + \beta g_2)(\mathbf{x}) = \alpha K_t g_1(\mathbf{x}) + \beta K_t g_2(\mathbf{x})$, K_t is a linear operator. Hence, it makes sense to investigate the spectrum of K_t, looking for special observables $\varphi(\mathbf{x})$ such that

$$K_t \varphi(\mathbf{x}) = \varphi(F_t(\mathbf{x})) = \lambda \varphi(\mathbf{x}) \tag{2.12}$$

where λ is the corresponding eigenvalue of K_t and $\varphi(\mathbf{x})$ is the eigenfunction. Let us assume that all the observables of the dynamical system lie in the linear span of such Koopman eigenfunctions, that is:

$$g(\mathbf{x}) = \sum_{i=1}^{\infty} \varphi_i(\mathbf{x}) c_i \tag{2.13}$$

where the c_i's are coefficients of observable g in the eigenfunction expansion, called the Koopman modes. Then we can describe the evolution of observables as:

$$\begin{aligned} K_t g(\mathbf{x}_k) &= \sum_{i=1}^{\infty} K_t \varphi_i(\mathbf{x}_k) c_i \\ &= \sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}_k) c_i \end{aligned} \tag{2.14}$$

because of expansion of g in the eigenbasis $\{\phi_i\}$ and the assumption that λ_i are eigenvalues of K_t for φ_i. But also

$$g(\mathbf{x}_{k+1}) = \sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}_k) c_i. \tag{2.15}$$

Thus, if one can identify from data a set of eigenfunctions and eigenvalues of K_t that spans a sufficiently large space of observables, then one may infer the evolution of an observable g in this space from its coefficients with respect to the identified eigenfunctions.

The Koopman operator K_t is unitary, and forms a one-parameter family (in t) of unitary transformations in the Hilbert space of square-integable observables. The Koopman operator is the dual, or left-adjoint, of the Perron-Frobenius operator, or transfer operator [34]. The Perron-Frobenius operator T of a map g acts on a probability distribution P by assigning TP as the probability distribution after applying g to P: $[TP](x \in A) = P(g(x) \in A)$, so $[TP](A) = P(g^{-1}(A))$. In Section 5.2.6 we will explain what Perron-Frobenius and Koopman operator are for the simple case of a discrete-time discrete space Markov chain.

3. Description and Initial Analysis of Experimental Data

We aim to investigate the FOG phenomenon in PD patients using stepping-in-place experiments on force plates. The collected data is subsequently analyzed to identify stepping and freezing episodes and delve into the underlying dynamics of FOG using mathematical modelling.

The methods developed in the thesis will be tested on FOG data obtained by stepping-in-place experiments by Nantel *et al.* [7]. Section 3.1 describes the experimental design and data collection process followed by [7]. After presenting the stepping-in-place task, we describe how the force data and time were non-dimensionalized for our analysis. In the following Section 3.1 and Section 3.2, we introduce a null model based on the Hopf bifurcation normal form with additive noise. This model represents the hypothesis that stepping is a harmonic oscillation coming from a stable limit cycle in an autonomous nonlinear oscillator, and the transition from stepping to freezing is due to uniformly distributed random small disturbances pushing the oscillator out of the basin of attraction of the limit cycle (toward an equilibrium). For this model we can investigate the dynamics exhibited during these transitions, for example, the distribution of escape times.

3.1 Stepping-in-place experiments

Nantel *et al.* [7] performed experiments recording freezing of gait using a repetitive stepping-in-place task on force plates which could identify freezing episodes in subjects with Parkinson's Disease. The task consisted of alternatingly raising the legs at a self-selected pace for 90 s per trial. Ground reaction forces were measured with a sampling frequency of 100 Hz on two force plates. The series of experiments generated 6 trial data sets of 90 s force time series for each of the 34 subjects. In figures and descriptions we label data set j for subject k as STk_j, where $k \in \{1,\ldots,34\}$, $j \in \{1,\ldots,6\}$. For example,

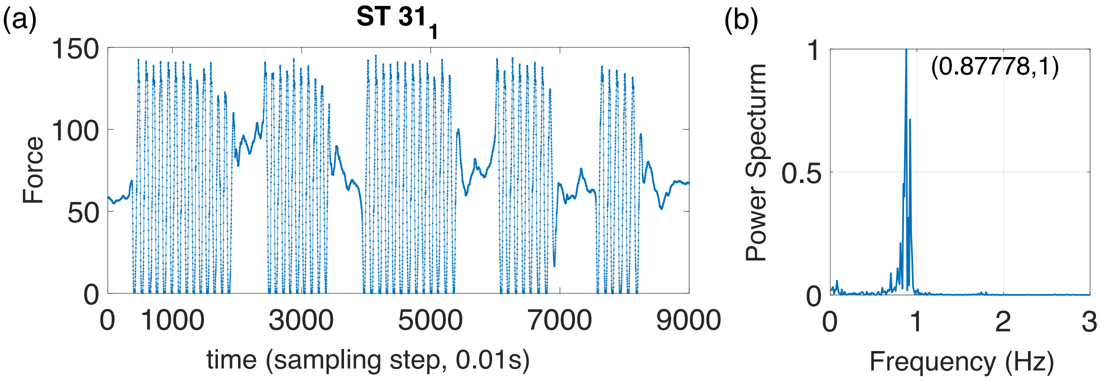


Figure 3.1: Panel (a): Time series of the left foot vertical force (as a percentage of body weight) of the subject numbered ST 31, data set 1, sampling time step $\delta t = 0.01$; Panel (b): Power spectrum (scaled to maximum equal to unit) with strong dominant peak at $f_{\text{stp}} \approx 0.88$Hz;

Figure 3.1(a) shows a typical time series generated by the data collection. Its title ST31_1 indicates that the time profile is from data set 1 of subject 31. The time series is the left foot vertical force, given as a percentage of the body weight, of subject ST 31. The sampling time step δt equals 0.01 seconds. This results in 9000 discrete time steps per trial. We observe several sudden transitions in the force magnitude from a large-amplitude oscillatory behavior to small-amplitude irregular fluctuations and drifting at approximately half the vertical force, which indicates that both feet are on the ground and the patient experiences a freezing episode. Episodes of standstill (involuntary freezing or intentional) are uniquely identified by absence of zero-force periods for longer than the identified stepping frequency shown in Figure 3.1(b). An expert manually identified freezing episodes in the time series data during the experiment in order to distinguish freezing episodes from episodes when the subject is intentionally standing still, as seen at the start and end of many time series.

During our analysis we non-dimensionalize the time series $\mathrm{Force}(t)$ of force data to the interval $[-1,1]$, shifting the force at freezing to approximately the value 0, by applying the transformation from $\mathrm{Force}(t)$ to $\mathrm{Force}_{\mathrm{scal}}(t)$ with

$$\mathrm{Force}_{\mathrm{scal}}(t) = \frac{\mathrm{Force}(t) - \mathrm{mean}(\mathrm{Force}(t))}{\max_t |\mathrm{Force}(t) - \mathrm{mean}(\mathrm{Force}(t))|}. \tag{3.1}$$

We report time in multiples of the sampling time step δt, such that all times t are integers. When presenting experimental data in the paper we label subject number (ST), data set and time interval (in multiples of $\delta t = 0.01$). For example, Figure 3.1 shows data from ST 31, data set 1, and the entire time interval $\{1,\dots,9000\}$ (corresponding to the full 90 s). In the force measurements, stepping is seen as a regular periodic behavior similar to a limit cycle oscillation, while freezing resembles an equilibrium or irregular drifting.

Stepping in place has been used as a protocol because it is extremely challenging to evaluate FOG using traditional force platforms, due to their limited size and the practical difficulties of observing unpredictable FOG events when asking participants to walk over them. For this reason, the approach of [7] has been to ask participants to step in place when standing on two force platforms. This approach provides highly accurate dynamic force information for each foot separately. Because these data relate to a protocol where patients have been denied the opportunity to walk forward, the temporal aspects of the task are more prominent (spatial features such as stride length are absent). Therefore, we focus our analysis on temporal features, i.e dynamical properties of scalar force data.

Full data sets and processing scripts are available at the following link `https://figshare.com/s/a14be7360925639736ba`.

3.2 Nullmodel: transitions independent of phase as escape from limit cycle

Time profiles during regular stepping exhibit strong periodicity, as Figure 3.1(a) shows. Figure 3.1(b) depicts the corresponding power spectrum characterized by a distinct dominant peak, corresponding to the stepping frequency f_{stp} of the subject during the experiment (in this case $f_{\mathrm{stp}} \approx 0.88\mathrm{Hz}$, such that the angular stepping frequency is $\omega_{\mathrm{stp}} \approx 1.76\pi/s$). The False Nearest Neighbour (FNN) criterion [35] for time delay embedding with Δ_t equal to a quarter of the dominant period, $T_{\mathrm{stp}}/4 = \pi/(2\omega_{\mathrm{stp}})$, demonstrates (as seen in Figure 3.2(right)) that two dimensions predict 65% of the signal correctly (that is, 35% of the predictions from nearest neighbours mismatch according to the FNN criterion), while 10% is mismatched with three embedding dimensions. The third embedding dimension plays a role after freezing events have occurred, as the time-delay embedded phase portraits in panel (a) and (b) of Figure 3.2 illustrate.

A phenomenological model for the escaping-from-a-limit-cycle behavior visible in Figure 3.2 is a (for example, two-dimensional) ODE, which is bistable with coexisting stable periodic limit cycle, corresponding to periodic stepping, and stable equilibrium state, corresponding to freezing, perturbed by small noise. Noisy fluctuations and/or

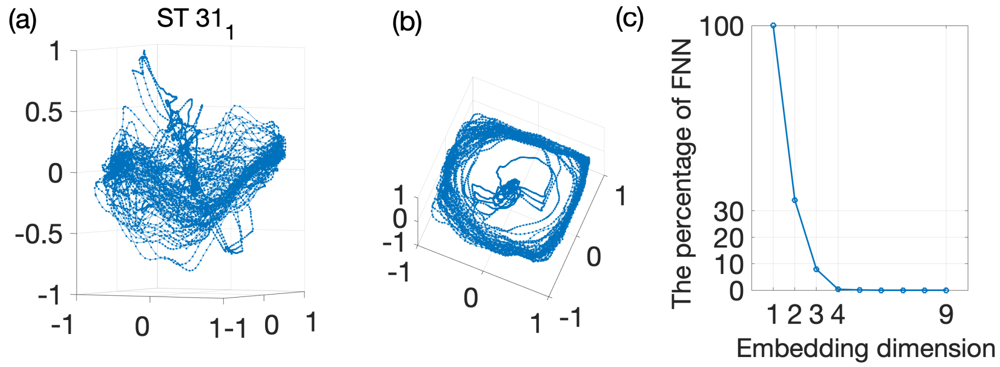


Figure 3.2: Time delay embedding analysis of time profile shown in Figure 3.1: Panel (a) and (b): Delay embedding in 3 dimensional space; Panel (c): False nearest neighbour percentage for different embedding dimensions and time delay.

external inputs could drive transitions between these two stable states representing transitions from stepping into freezing. The generalized Hopf bifurcation normal form model [36] exhibits such dynamics and hence can be used for describing some aspects of the behavior shown by the time-delay embedding of the data Figure 3.2.

3.2.1 Generalized Hopf Normal Form

The general minimal Hopf bifurcation model has the following form [36]:

$$\dot{y}_1(t) = \beta_1 y_1(t) - c y_2(t) + \beta_2 y_1(t)(y_1^2(t) + y_2^2(t)) - y_1(t)(y_1^2(t) + y_2^2(t))^2 \tag{3.2a}$$

$$\dot{y}_2(t) = c y_1(t) + \beta_1 y_2(t) + \beta_2 y_2(t)(y_1^2(t) + y_2^2(t)) - y_2(t)(y_1^2(t) + y_2^2(t))^2 \tag{3.2b}$$

where $y = (y_1, y_2)^T \in \mathbb{R}^2, \beta \in \mathbb{R}^2$. When transforming from Cartesian coordinates (y_1, y_2) to polar coordinates, we may treat $(y_1, y_2) \in \mathbb{R}^2$ as a complex variable $Z = Re^{i\theta} = y_1 + iy_2$. Then the form of (3.2) in polar coordinates is:

$$\dot{R} = \beta_1 R + \beta_2 R^3 - R^5, \qquad R > 0 \tag{3.3a}$$

$$\dot{\theta} = ct \tag{3.3b}$$

The local bifurcation diagram in the parameter plane β_1-β_2 of the normal form is presented in Figure 3.3. The different regions and bifurcations near the Bautin bifurcation are the following. In region 3 (Bistability) the system has a stable fixed point and a stable limit cycle coexisting due to the subcritical nature of the Hopf bifurcation. The point $\beta_1 = 0$ separates two types of the Hopf bifurcation curve: the half-line $\mathrm{H}_- = \{(\beta_1, \beta_2) : \beta_1 = 0, \beta_2 < 0\}$ corresponds to a supercritical bifurcation that generates a stable limit cycle, while the half-line $\mathrm{H}_+ = \{(\beta_1, \beta_2) : \beta_1 = 0, \beta_2 > 0\}$ corresponds to a subcritical bifurcation that generates an unstable limit cycle. Two hyperbolic limit cycles (one stable and one unstable) exist in the region between the line H_+ and the curve $T = \{(\beta_1, \beta_2) : \beta_1 = -\frac{1}{4}\beta_2^2, \beta_2 > 0\}$, at which two cycles collide and disappear via a saddle-node bifurcation of periodic orbits [2]. We go around the origin in counter-clockwise direction, $\beta_1 = R\cos\phi, \beta_2 = R\sin\phi$, with $R \ll 1$. If we start at $\phi = -\pi$, the system has a single stable equilibrium and no cycles. Crossing the Hopf bifurcation boundary H_- from region 1 to region 2 implies the appearance of a unique and stable limit cycle, which survives when we enter region 3. Crossing the Hopf boundary H_+ makes an extra unstable cycle show up which is inside the first one, while the equilibrium regains its stability. The cycle near the equilibrium is unstable and the other cycle is stable, this two cycles disappear at the curve T through a fold bifurcation that leaves a single stable equilibrium. This phenomenon is also called the generalized Bautin bifurcation [36].

We notice that the dynamics in region 3 is bistable: a stable limit cycle with radius and a stable equilibrium coexist, as we proposed as phenomenological model for Figure 3.2.

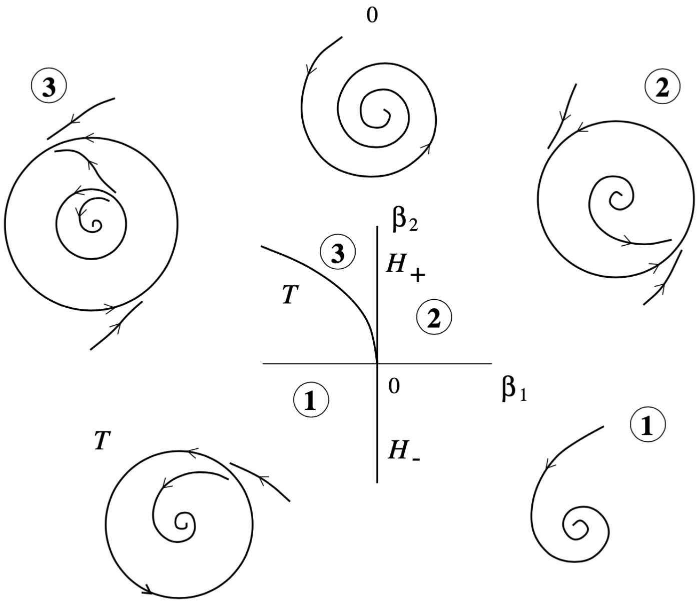


Figure 3.3: Bautin bifurcation diagram from `http://www.scholarpedia.org/article/File:Bautin.gif`): In a Bautin bifurcation an equilibrium point changes stability in a Hopf bifurcation, with not just one but potentially multiple limit cycles reated or destroyed as parameters β_1, β_2 change [2].

Thus, we expect that perturbing this dynamics with noise can generate the time series shown in Figure 3.1 and Figure 3.2. We model gait as a limit cycle of radius 1 and freezing as equilibrium at the origin. This matches how we normalized our experimental data where oscillations have amplitude equal to 1. We parametrise our model to have a stable limit cycle with radius equals to 1 as well, which implies that $\dot{R}=0$ for $R=1$, such that $\beta_2 = 1-\beta_1$. The angular frequency $\dot{\theta}$ should equal to $0.9667(2\pi)$ for consistency with our normalization of the experimental data. Let $c = 2\pi\omega$, then(3.3) will have the following form:

$$\dot{R} = \beta R + (1-\beta)R^3 - R^5, \qquad R > 0 \tag{3.4a}$$

$$\dot{\theta} = 2\pi\omega. \tag{3.4b}$$

Thus, all periodic orbits of (3.2) are circles around the origin and their radius is the equilibrium of (3.4a). The origin ($R = 0$) is a special case, corresponding to an equilibrium (a circle with radius 0). The equation (3.4a) for $\dot{R}$ can be written in the form of a flow following the (negative) gradient of a potential V,

$$\frac{\mathrm{d}R}{\mathrm{d}t} = -\frac{\partial V(R)}{\partial R} \tag{3.5}$$

where the potential function V is the integral of r.h.s of given by

$$V := -\frac{\beta}{2}R^2 - \frac{1-\beta}{4}R^4 + \frac{R^6}{6}. \tag{3.6}$$

The equilibrium points of the equation (3.5) are given by $\dot{R} = 0$, such that

$$\beta R + (1-\beta)R^3 - R^5 = 0, \qquad R > 0 \tag{3.7}$$

Apparently, $R = 0$ is one of the equilibrium points. The other equilibrium points are the solutions of the following equation:

$$\beta + (1-\beta)R^2 - R^4 = 0, \qquad R > 0. \tag{3.8}$$

Thus, (3.5) has 3 equilibria for $\beta \in (-\infty, 0) \setminus \{-1\}$, which are:

$$R_1 = 0, R_2 = \sqrt{-\beta}, R_3 = 1.$$

For $\beta = 0, -1$, there are two equilibria:

$$R_1 = 0, R_3 = 1.$$

For our stability analysis we consider the right-hand side of (3.4a),

$$f(R) = \dot{R} = \beta R + (1-\beta)R^3 - R^5, \tag{3.9}$$

and determine the derivative of $f(R)$:

$$f'(R) = \beta + 3(1-\beta)R^2 - 5R^4 \tag{3.10}$$

We can now get $f'(R_1) = \beta$, $f'(R_2) = -2\beta - 2\beta^2$, $f'(R_3) = -2\beta - 2$.

For $\beta > 0$, R_2 does not exist (is not real), $f'(R_1) > 0$, $f'(R_3) < 0$. Thus, R_1 is unstable and R_3 is stable. Therefore in the full two-dimensional system (3.4), the only attractor is a stable periodic orbit ($R = 1$) surrounding an unstable equilibrium at the origin. This equilibrium becomes stable in a subcritical Hopf bifurcation at $\beta = 0$ and in the regime $-1 < \beta < 0$ the system exhibits bistability with an attracting fixed point (origin, $f'(R_1) < 0$) and an attracting limit cycle (radius $R = R_3 = 1$, $f'(R_3) < 0$) separated by an unstable limit cycle ($R = \sqrt{-\beta}$, $f'(R_2) > 0$). For $\beta < -1$, $f'(R_1) < 0$, $f'(R_2) < 0$, $f'(R_3) > 0$, therefore R_3 is the unstable state, and R_1 and R_3 are stable states. The periodic orbits with $R_2 = \sqrt{-\beta}$ and $R_3 = 1$ meet each other in a transcritical bifurcation at $\beta = -1$, exchanging their stability (the periodic orbit with larger radius R is always stable). The derivative of the phase (angle) θ is the frequency of the oscillations, which we fix by setting $\omega = 0.9667$. Figure 3.4 shows the bifurcation diagram of system (3.4), where $-2 < \beta < 1$.

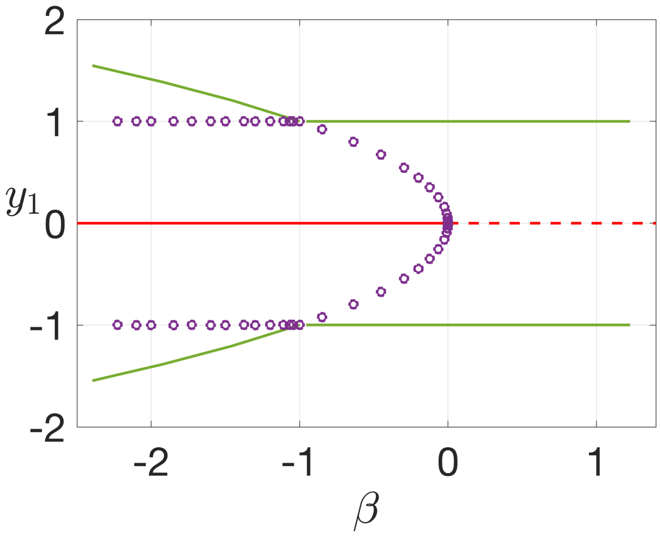


Figure 3.4: Bifurcation diagram in Cartesian coordinate, purple circles represent the unstable limit cycles, green lines represent the stable limit cycles, red line represents the stable equilibria and the red dashed line represents the unstable equilibria.

Additive noise effects The deterministic model cannot describe the escapes from the limit cycle, since in ODEs trajectories in the basin of attraction of a limit cycle will converge to it for $t \to \infty$. So, we need to include random perturbations into our differential equations. This is described mathematically by stochastic differential equations (SDEs), which we explain here for the simple case of *additive white noise*.

An ODE initial value problem solution $y(t)$ of $y' = f(y), y(0) = y_0$ is defined as the unique fixed point of the fixed point equation $y(t) = y_0 + \int_0^t f(y(s))\mathrm{d}s$. A SDE with additive noise for $y(t)$ is written in the form $\mathrm{d}y(t) = f(y(t))\mathrm{d}t + \sigma \mathrm{d}W(t)$ with initial value $y(0) = y_0$. The solution $y(t)$ is defined as the unique fixed point of $y(t) = y_0 + \int_0^t f(y(s))\mathrm{d}s + \sigma W(t)$, where $W(t)$ is the standard Wiener process (random walk). Since $W(t)$ is random, $y(t)$ is random as well. For any given realisation $W(t)$ (which one may generate independent of the SDE), one can treat the SDE as an ODE with the forcing $W(t)$. We model disturbances as a Gaussian white noise (the integral of which is

$W(t)$). We expect the cumulative effect of the noise to cause random transitions from walking to freezing. Hence, the model we consider is based on the generalized Hopf normal form model with additive white noise and has the following form:

$$\mathrm{d}y_1(t) = \left[\beta y_1 - 2\pi\omega y_2 + (1-\beta)y_1(y_1^2+y_2^2) - y_1(y_1^2+y_2^2)^2\right]\mathrm{d}t + \sigma \mathrm{d}W_1(t) \tag{3.11a}$$

$$\mathrm{d}y_2(t) = \left[2\pi\omega y_1 + \beta y_2 + (1-\beta)y_2(y_1^2+y_2^2) - y_2(y_1^2+y_2^2)^2\right]\mathrm{d}t + \sigma \mathrm{d}W_2(t). \tag{3.11b}$$

We set the initial condition as $y_1(0) = 1, y_2(0) = 0$, which is on the stable limit cycle of the ODE (with $\sigma = 0$). The inhomogeneities $W_j(t)$ are two independent standard Wiener processes. Here σ is the noise amplitude, σ^2 is the variance, where $\sigma > 0$ for the stochastic case. Since our goal is modelling the transition from stepping to freezing in the data, we assume that the unstable limit cycle is close to the stable limit cycle of the deterministic part, namely, $\sqrt{-\beta}$ is close to 1, but $\sqrt{-\beta} < 1$. Figure 3.5 demonstrates the dynamics of (3.11), which is computed in MATLAB using the Euler-Maruyama method for stochastic differential equations with the initial condition at $y_1 = 1, y_2 = 0$, Euler-Maruyama step size $h = 0.0003125$ and radius $R = \sqrt{y_1^2 + y_2^2}$. The realization trajectory spends some time near the stable limit cycle of the deterministic part (circling counterclockwise at radius $R = R_3 = 1$). Eventually it crosses the unstable limit cycle of the deterministic part first, then it crosses the threshold $R = 0.7$ and never come back. Finally the trajectory is attracted to the stable equilibrium, $R = 0$, which is the origin. The time series shows this transition from an oscillatory regime, to small, noise dominated fluctuations around the equilibrium state.

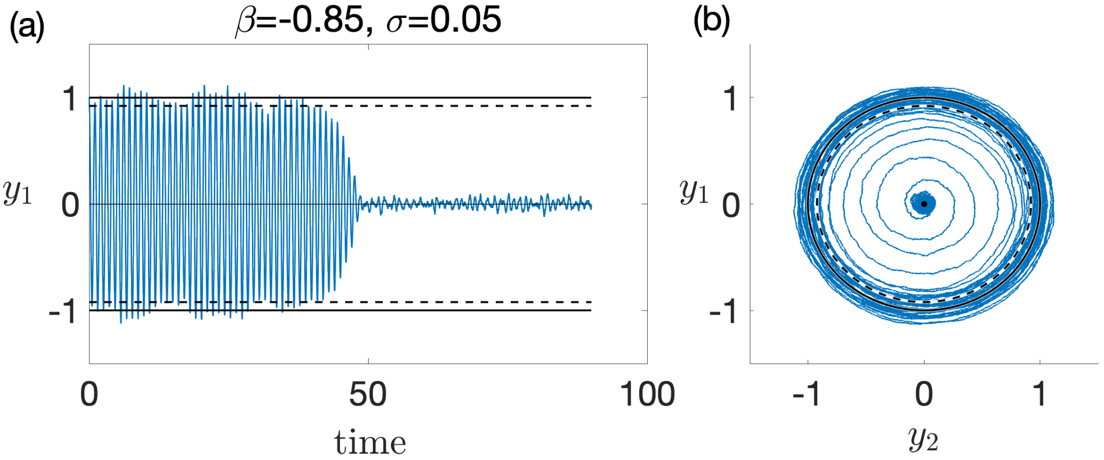


Figure 3.5: An example realization of the noise-driven dynamics of (3.11). Both pictures show the same realization in blue for $\beta = -0.85$ and $\sigma = 0.05$ in the phase space, for which the deterministic part contains a stable equilibrium at the origin, an unstable limit cycle (black dotted line), and stable limit cycle (solid line).

4. Escape Times for the Stochastically Perturbed Generalised Hopf Bifurcation Normal Form Model

4.1 Systematic study of escape time distribution

We expect that trajectories of an SDE with a bistable deterministic behaviour for $\sigma = 0$ and small noise amplitude σ will spend most time in the vicinity of the deterministic attractors. These attractors could be either stable steady states or periodic orbits, depending on the system in question. Occasionally, the system may transition between these attractors rapidly, a process driven by the perturbations from the small random forcing. Any positive amplitude σ of this forcing is sufficient to induce escapes from either the steady state or the periodic orbit, but the residence time near the deterministic attractor is dictated by the properties of the attractor in the deterministic system, such as linear stability and basin of attraction. The exit phenomenon, that is, escaping from a bounded domain in state space, is caused by random forcing on the evolution of such dynamical systems. In our case, the escape time t_{esc} corresponds to the time it takes for the realization trajectory to switch from being close to the stable limit cycle to being close to the origin. More precisely, when the amplitude R of the trajectory crosses the lower threshold ξ_{l}, where $\xi_{\mathrm{l}} \in D_1 = (0, R_2)$, or crosses the upper threshold ξ_{h}, where $\xi_{\mathrm{h}} \in D_2 = (R_3, \infty)$, the first escape time starting at $R(0) = R_0 = 1$ from a bounded domain D is defined as

$$t_{\mathrm{esc}} := \inf\{t \geq 0, R(t) < \xi_{\mathrm{l}} \text{ or } R(t) > \xi_{\mathrm{h}}\}, \tag{4.1}$$

Note that t_{esc} is a random variable that depends on the noise realization and reflects the influence of the noise on the nonlinear dynamics. If the noise amplitude is small, it does not matter where we start if we start anywhere safely in the basin of attraction. Figure 4.1 demonstrates the distribution of 1000 first escape time under different noises and different β values. We can see the histograms of first escape time from numerical simulations show the distributions which is similar to exponentially distributions if we cut off the last bin at $t = 90$s, which is what we expected distribution for using Guassian white noise [37]. The first bin does not have the most entries in each panel of Figure 4.1. Obviously, escaping immediately is not that likely. For larger noise amplitudes, the escape happens faster, which causes the shape to be more similar to an exponential tail [38].

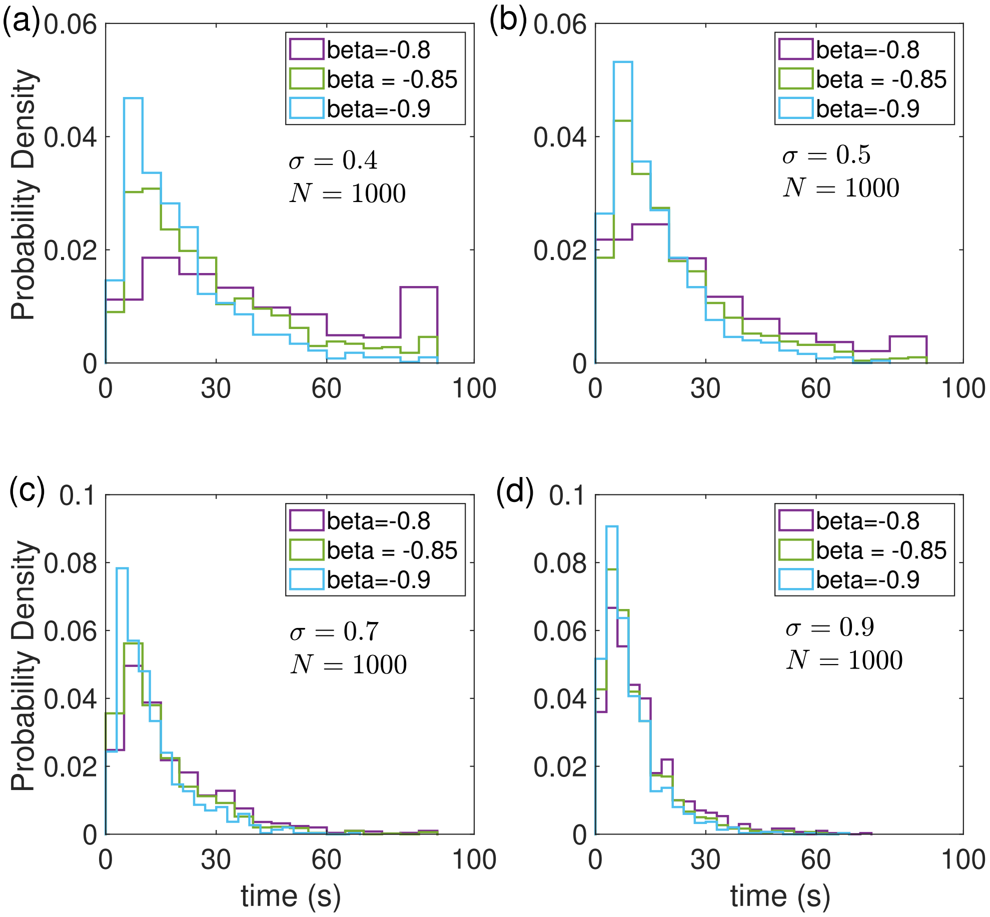


Figure 4.1: Frequency Distributions of First Escape Times from 1000 numerical simulations across various β values (with σ set to 0.4, 0.5, 0.7, and 0.9 in panels (a) (b) (c) and (d) respectively). Each panel depicts distinct β values in same pattern: $\beta = -0.8$ in purple, $\beta = -0.85$ in green, and $\beta = -0.9$ in blue.

For small enough noise the escape time has a cumulative distribution $Q(t) = \mathbb{P}\{t_{\text{esc}} < t\}$ with an exponential tail [38] and the mean escape time T (MET) W from the oscillatory attractor to the origin is

$$T = E(t_{\text{esc}}) = \int_0^\infty tQ'(t)\mathrm{d}t \tag{4.2}$$

where $Q'(t)$ is the density calculated by the derivative of $Q(t)$.

Differential equation for Mean First Escape Time To derive an expression for the mean first escape time, we transform SDE (3.11) following [39] into polar coordinates given by $z(t) = R(t)\exp[i\theta(t)]$:

$$\mathrm{d}R = \left[\beta R + (1-\beta)R^3 - R^5 + \frac{\sigma^2}{2R}\right]\mathrm{d}t + \sigma \mathrm{d}W_R \tag{4.3a}$$

$$\mathrm{d}\theta = 2\pi\omega \mathrm{d}t + \frac{\sigma}{R}\mathrm{d}W_\theta \tag{4.3b}$$

where W_R and W_θ are independent standard Wiener processes. The terms with $\frac{\sigma^2}{2R}$ and $\frac{\sigma}{R}$ arise from Itô's lemma. As the R equation is independent of θ we can consider the escape problem as a one-dimensional potential problem for $R(t)$. From equation (4.3), we can get,

$$\mathrm{d}R = -\frac{\partial V}{\partial R}\mathrm{d}t + \sigma \mathrm{d}W_R \tag{4.4}$$

where the potential function V is given by

$$V := -\frac{\beta}{2}R^2 - \frac{1-\beta}{4}R^4 + \frac{R^6}{6} - \frac{\sigma^2}{2}\ln R \tag{4.5}$$

The mean escape time $T(\beta,\sigma)$ from (4.2) can be found by considering solutions $R(t)$ of the SDE (4.3) and defining the mean first escape time (MFET) from an interval $(\xi_{\mathrm{l}},\xi_{\mathrm{h}})$ when starting from $R(0) = R_0$ inside the interval $(\xi_{\mathrm{l}},\xi_{\mathrm{h}})$:

$$W(R_0) := \mathbb{E}(\inf\{t > 0 : R(t) = \xi_{\mathrm{l}} \text{ or } R(t) = \xi_{\mathrm{h}}, \quad \text{given } R(0) = R_0\}). \tag{4.6}$$

In the following we take a fixed value of $\xi_{\mathrm{l}} = \xi = 0.7 < R(0) = 1$ and set $\xi_{\mathrm{h}} > R(0)$ sufficiently large. As the equilibrium with $R = 1$ is stable for the deterministic part, with basin of attraction stretching to $R \to \infty$, we do not expect the escape rate from

$(\xi_{\mathrm{l}}, \xi_{\mathrm{h}})$ to depend on ξ_{h}, as long as ξ_{h} is large (we choose $\xi_{\mathrm{h}} = 2$ for our computations). By the Dynkin formula for Markov processes in [40], $W(R_0)$ satisfies the following second-order differential equation:

$$\frac{\sigma^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}R^2}W(R) - V'(R)\frac{\mathrm{d}}{\mathrm{d}R}W(R) = -1 \tag{4.7}$$

with boundary conditions $W(\xi_{\mathrm{l}}) = 0$, $W(\xi_{\mathrm{h}}) = 0$. In our example $[\xi_{\mathrm{l}}, \xi_{\mathrm{h}}] = [0.7, 2]$.

General solution formula for (4.7)

We observe that (4.7) is linear and depends only on $\frac{\mathrm{d}W(R)}{\mathrm{d}R}$, such that we can reduce the second-order boundary value problem (4.7) to an integral. To this end, we first remove the factor in front of the highest-order term in equation (4.7), to get

$$\frac{\mathrm{d}^2}{\mathrm{d}R^2}W(R) - \frac{2V'(R)}{\sigma^2}\frac{\mathrm{d}}{\mathrm{d}R}W(R) = -\frac{2}{\sigma^2}. \tag{4.8}$$

Let $U(R) = \frac{\mathrm{d}}{\mathrm{d}R}W(R)$, then we can recover W as the integral of U and U satisfies

$$\frac{\mathrm{d}}{\mathrm{d}R}U(R) - \frac{2V'(R)}{\sigma^2}U(R) = -\frac{2}{\sigma^2} \tag{4.9}$$

The Integrating Factor of (4.9) is: $\exp\left(\int -\frac{2V'(R)}{\sigma^2}\mathrm{d}R\right) = \exp\left(-\frac{2V(R)}{\sigma^2}\right)$ such that we have

$$\frac{\mathrm{d}}{\mathrm{d}R}\left[\exp\left(-\frac{2V}{\sigma^2}\right)U\right] = -\frac{2}{\sigma^2}\exp\left(-\frac{2V}{\sigma^2}\right) \tag{4.10}$$

Intergrating eq (4.10) from ξ_l to R we get

$$\left[\exp\left(-\frac{2V}{\sigma^2}\right)U\right]_{\xi_l}^{R} = -\frac{2}{\sigma^2}\int_{\xi_l}^{R}\exp\left(-\frac{2V(y)}{\sigma^2}\right)\mathrm{d}y, \tag{4.11}$$

such that

$$\exp\left(-\frac{2V(R)}{\sigma^2}\right)U(R) = \exp\left(-\frac{2V(\xi_l)}{\sigma^2}\right)U(\xi_l) - \frac{2}{\sigma^2}\int_{\xi_l}^{R}\exp\left(-\frac{2V(y)}{\sigma^2}\right)\mathrm{d}y. \tag{4.12}$$

Using x as the argument of U (instead of R) the general solution U has the form

$$\begin{aligned} U(x) = \exp\left(-\frac{2V(\xi_l)}{\sigma^2}\right)U(\xi_l)\exp\left(\frac{2V(x)}{\sigma^2}\right) \\ - \frac{2}{\sigma^2}\exp\left(\frac{2V(x)}{\sigma^2}\right)\int_{\xi_l}^{x}\exp\left(-\frac{2V(y)}{\sigma^2}\right)\mathrm{d}y. \end{aligned}$$

Inserting now $W(x)$ as the integral of $U(x)$, we integrate (4.13) from ξ_l to R to get

$$\begin{aligned} [W(x)]_{\xi_l}^{R} =& \exp\left(-\frac{2V(\xi_l)}{\sigma^2}\right)U(\xi_l)\int_{\xi_l}^{R}\exp\left(\frac{2V(x)}{\sigma^2}\right)\mathrm{d}x \\ &- \frac{2}{\sigma^2}\int_{\xi_l}^{R}\int_{\xi_l}^{x}\exp\left[\frac{2(V(x)-V(y))}{\sigma^2}\right]\mathrm{d}y\mathrm{d}x. \end{aligned} \tag{4.13}$$

Since the boundary condition at ξ_l is $W(\xi_l) = 0$, this simplifies to

$$\begin{aligned} W(R) = \exp\left(-\frac{2V(\xi_l)}{\sigma^2}\right)W'(\xi_l)\int_{\xi_l}^{R}\exp\left(\frac{2V(x)}{\sigma^2}\right)\mathrm{d}x \\ - \frac{2}{\sigma^2}\int_{\xi_l}^{R}\int_{\xi_l}^{x}\exp\left[\frac{2(V(x)-V(y))}{\sigma^2}\right]\mathrm{d}y\mathrm{d}x. \end{aligned} \tag{4.14}$$

In this equation for $W(R)$, the constant $W'(\xi_l)$ is still unknown. This constant is determined by the other boundary condition, at $R = \xi_h : W(\xi_h) = 0$.

$$0 = W(\xi_h) = \exp\left(-\frac{2V(\xi_l)}{\sigma^2}\right) W'(\xi_l) \int_{\xi_l}^{\xi_h} \exp\left(\frac{2V(x)}{\sigma^2}\right) \mathrm{d}x \\ - \frac{2}{\sigma^2} \int_{\xi_l}^{\xi_h} \int_{\xi_l}^{x} \exp\left[\frac{2(V(x) - V(y))}{\sigma^2}\right] \mathrm{d}y\mathrm{d}x. \quad (4.15)$$

We can isolate $W'(\xi_l)$, such that

$$W'(\xi_l) = \frac{\frac{2}{\sigma^2} \int_{\xi_l}^{\xi_h} \int_{\xi_l}^{x} \exp\left[\frac{2(V(x) - V(y))}{\sigma^2}\right] \mathrm{d}y\mathrm{d}x}{\exp\left(-\frac{2V(\xi_l)}{\sigma^2}\right) \int_{\xi_l}^{\xi_h} \exp\left(\frac{2V(x)}{\sigma^2}\right) \mathrm{d}x}. \quad (4.16)$$

Let us abbreviate $S(R) = \exp\left(\frac{2V(R)}{\sigma^2}\right)$, where in our case $V = -\frac{\beta}{2}R^2 - \frac{1-\beta}{4}R^4 + \frac{R^6}{6} - \frac{\sigma^2}{2}\ln R$. Then the general solution for equation (4.7) is

$$W(R) = \frac{2}{\sigma^2} \frac{\int_{\xi_l}^{\xi_h} \int_{\xi_l}^{x} \frac{S(x)}{S(y)} \mathrm{d}y\mathrm{d}x}{\int_{\xi_l}^{\xi_h} S(x)\mathrm{d}x} \int_{\xi_l}^{R} S(x)\mathrm{d}x - \frac{2}{\sigma^2} \int_{\xi_l}^{R} \int_{\xi_l}^{x} \frac{S(x)}{S(y)} \mathrm{d}y\mathrm{d}x. \quad (4.17)$$

For our asymptotically polynomial potential ($V(R) \sim R^6/6$ for large R), the expression $W(R)$ in (4.17) has a well-behaved finite limit for $\xi_h \to \infty$, such that we may replace the upper integral bound ξ_h with $+\infty$ in (4.17). This would correspond to permitting arbitrarily large excursions toward large R before escape.

Figure 4.2 shows the MFET simulated for the polar coordinate R for σ between 0.01 and 0.1 with different β values. We use MATLAB to numerically calculate the MFET and set the ω to 0.9667 and the total time to 90 seconds, matching the experimental data. For small noise, $\sigma = 0.01$, sometimes 0.02, the MFET converge to 90s but we are expecting a higher MFET, this can be reasonably explained by the total time we set which is 90 seconds. We also expect an exponential tail for the MFET, and the position of the unstable limit cycle $R = \sqrt{-\beta}$ should be close to the stable limit cycle which is $R = 1$, so next we particulary put our attention on $\beta = -0.8, -0.85, -0.9$.

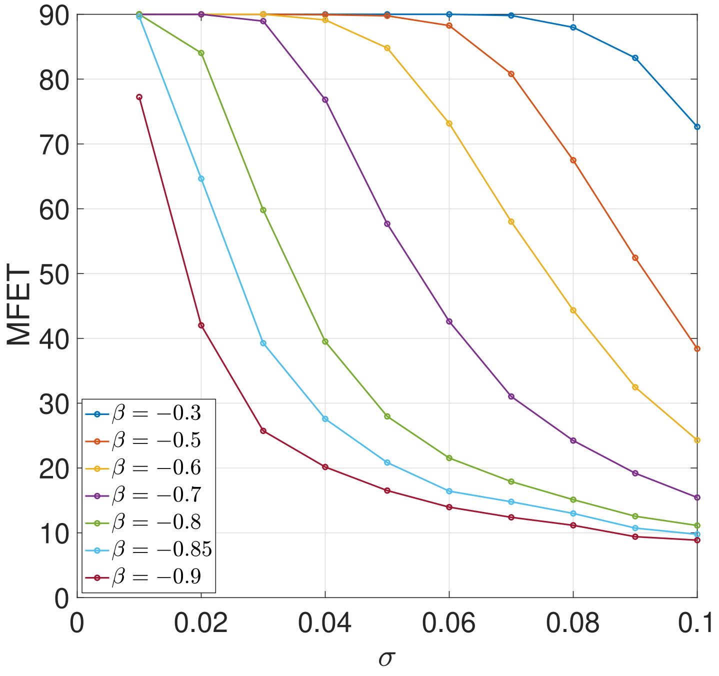


Figure 4.2: Mean first escape time in Polar coordinate for (4.3) 1000 Monte-Carlo simulations with different β values (with β set to -0.3, -0.5, -0.6, -0.7, -0.8, -0.85 and -0.9) for σ between 0.01 and 0.1.

We use the MATLAB chebfun package [41] to solve (4.7) directly as a Dirichlet boundary value problem, instead of evaluating the explicit integral (4.17). Figure 4.3 summaries the comparsion for MFET between the solution of (4.7) and the Monte-Carlo simulations of (4.3)(3.11) with sample size $N = 1000$ and $\beta = -0.8, -0.85, -0.9$. The mean escape time simulated from Polar coordinate for (3.11) agrees well with the solution of (4.7). The MFET for the SDE in polar coordinates, (4.3), is slightly smaller than the estimate from simulations as we stop simulation at and count any realization that does not escape before $t = 90$ as escaped at 90, which underestimates the MFET especially for small σ values. However, in Figure 4.3 the simulation result of (3.11) in Cartesian coordinates deviates strongly from the nearly exact value and the Monte-Carlo simulations of (4.3) in polar coordinates. Figure 4.4 investigates the source of the discrepancy of the escape statistics in simulations between polar and Cartesian coordinates systematically.

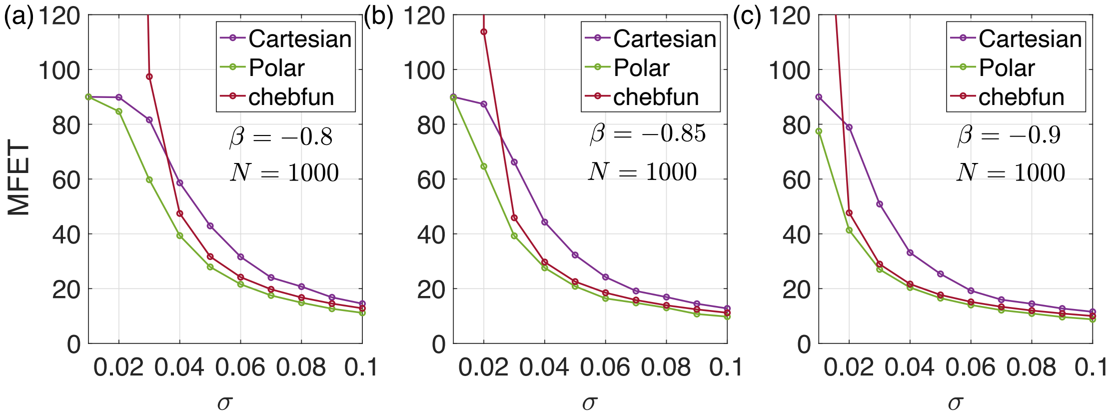


Figure 4.3: The green line and the purple line are the mean first escape time from numerical simulations in Polar coordinate(4.3) and in Cartersian coordinate(3.11). The red line is the analytical results for the mean first escape time. Panels (a) (b) and (c) show the MFET for Euler Maruyama step size $h = 0.0003125$ under $\beta = -0.8, -0.85$ and -0.9 respectively.

Figure 4.4 shows the MFET in Polar coordinate and in Cartesian coordinate for $\beta = -0.85$ for three time steps in the Euler-Maruyama stochastic integrator. As we reduce the step size the results for the Cartesian coordinates approach the simulation results for polar coordinates. Thus, we conclude that the discrepancy is due to the discretization error caused by the positive step size in the Euler-Maruyama method.

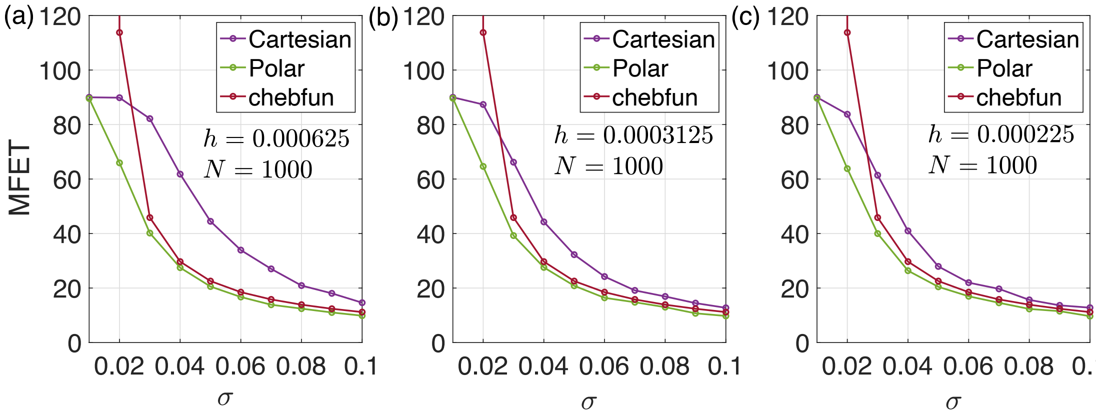


Figure 4.4: The green line and the purple line are the mean first escape time estimated from 1000 numerical simulations in polar coordinates (4.3) and in Cartesian coordinates (3.11). The red line shows the solution of (4.7) using `chebfun`. Panels (a) (b) and (c) show the MFET for $\beta = -0.85$ under Euler Maruyama step size h= 0.000625, 0.0003125, and 0.000225 respectively.

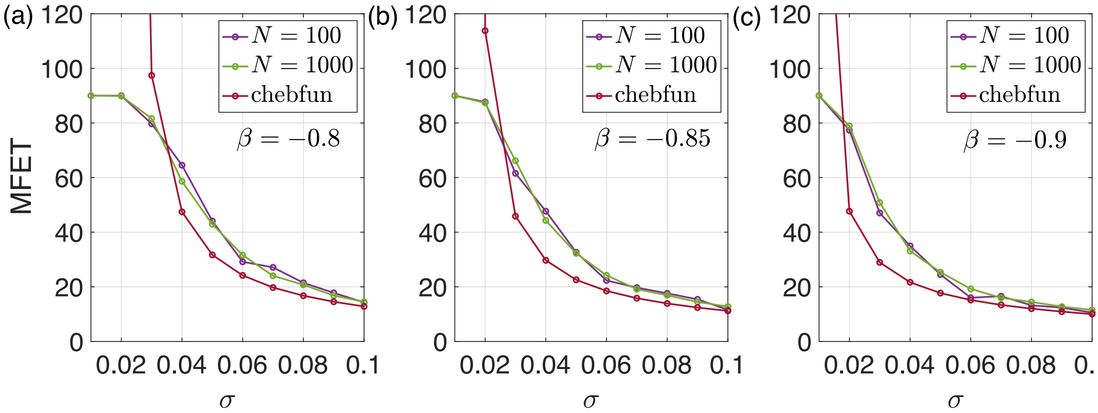


Figure 4.5: Mean first escape time from numerical simulations in Cartesian coordinate under sample size =100(green line), 1000(purple line). The red line is the analytical results for the mean first escape time. Euler Maruyama step size $h = 0.0003125$. Panels (a), (b), and (c) correspond to $\beta = -0.8$, $\beta = -0.85$, and $\beta = -0.9$ respectively.

We also investigate if the sample size has an effect on the difference of MFET distribution for different β values. Figure 4.5 demonstrates the MFET from numerical simulations in Cartesian coordinate under sample size 100 and sample size 1000. We can see that the sample size does not have a systematic influence on the MFETs. Figure 4.6 shows the uncertainty of the MFET estimate from numerical simulations for three values of β in Cartesian coordinates and in polar coordinates for sample size $N = 1000$. The symmetric error bars show the empirical standard deviation of MFET. This standard deviation is nearly independent of σ. The exception are small σ, for which many escapes are recorded as occuring at $t = 90$, the maximal time. Thus the standard deviation is small in the plot, not reflecting the true uncertainty.

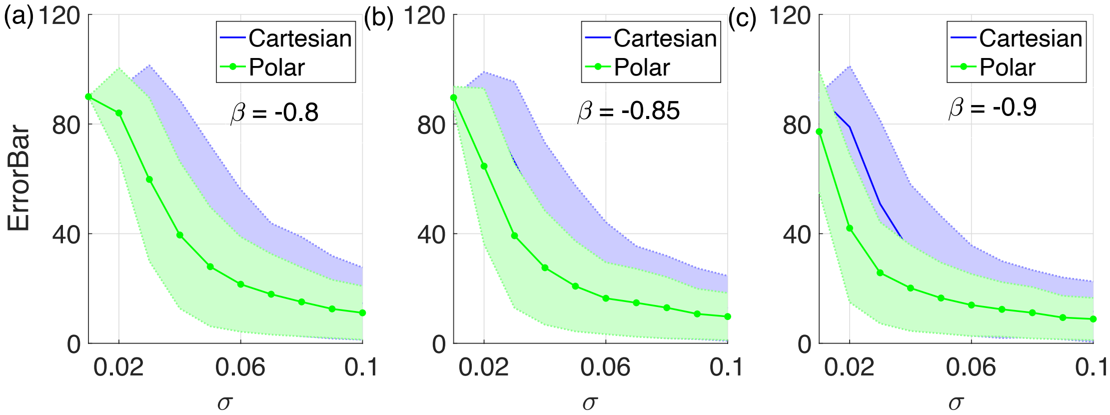


Figure 4.6: Error bar of mean first escape time from numerical simulations in Cartersian coordinate and in Polar coordinate under sample size $N = 1000$, the deep green and purple line in the middle is the mean first escape, shaded error bars generated by standard diviation which are drawed symmetrically. Panels (a), (b), and (c) correspond to $\beta = -0.8$, $\beta = -0.85$, and $\beta = -0.9$ respectively.

4.2 Conclusion for mean first-escape time from limit cycle

In this chapter, we delved into the dynamics of Freezing of Gait (FOG) in Parkinson's Disease through a mathematical framework based on a SDE with a bistable deterministic part and small-amplitude additive noise. In this SDE the transition from stepping to freezing corresponds to a transition from fluctuations around a stable limit cycle of the deterministic part to small fluctuations around the origin. The time it takes until this transition occurs is a typical escape time problem. Our SDE model is based on the generalised Hopf bifurcation normal form model, which provided a stochastic approach to modeling the transition from stepping to freezing. This model allowed us to account for the inherent randomness associated with transition. We explored the escape time problem, which focused on how the system deviates from the steady state under the influence of random white noise. Moreover, we found that the accuracy of our simulations was significantly affected by the step size of the numerical method employed.

5. Time Series Analysis and Modelling of the Freezing of Gait Phenomenon

In Chapter 3, we examined how a trajectory departures from the stable limit cycle in the Hopf Normal Form equation (3.11) due to white noise for different parameter values (β,σ). A notable assumption behind this theoretical model is that transition from the stable limit cycle, or ”stepping,” to the stable equilibrium, or ”freezing” is random and uniformly distributed around the circle. However, this uniform randomness may not be consistent with experimental data for cases of gait freezing in Parkinson’s disease.

One question arising when treating stepping as oscillations (through delay embedding or Hilbert transform) is whether there exists a unique or preferred angle at which the trajectory approaches this stable equilibrium. Addressing this question will enhance our understanding of the dynamics underpinning the experimental system and serve to refine the model.

In this chapter, we direct our attention towards the stepping behaviour of Parkinson’s patients. To that end, we apply the Hilbert Transform to construct a trajectory in a two-dimensional phase space from Parkinson’s data, converting a one-dimensional signal into a complex analytical signal. This allows us to propose a method for timing the transition from stepping into gait freezing and identifying this transition within the constructed phase space.

The results in this chapter largely draws upon our published work, which can be found at `https://doi.org/10.1137/22M1484341`. In addition to the material covered in the article, this chapter includes detailed derivations of formulas and explores the relationship between time series and transfer operators(Section 5.2.6).

5.1 Robust location of escape from oscillations

5.1.1 Hilbert Transform embedding of stepping data time series

The results of delay embedding in Figure 3.2(a,b) show nearly planar oscillatory behavior during regular stepping in the stepping-in-place experimental data. This makes the Hilbert Transform embedding a natural choice when expressing stepping as an oscillator. Following [1], we extract for a real scalar signal $x(t)$ on an interval of length $T = 2\pi/\omega$ the Fourier coefficients x_k. Then the Hilbert Transform $\tilde{x}(t)$ has the Fourier coefficients $-\mathrm{i}x_k$ (below we use the overbar notation $\bar{x}_k$ for the complex conjugate of x_k):

$$x(t) = x_0 + \sum_{k=1}^{\infty}\left[x_k \mathrm{e}^{k\omega t\mathrm{i}} + \bar{x}_k \mathrm{e}^{-k\omega t\mathrm{i}}\right], \qquad \tilde{x}(t) = \sum_{k=1}^{\infty}\left[-\mathrm{i}x_k \mathrm{e}^{k\omega t\mathrm{i}} + \mathrm{i}\bar{x}_k \mathrm{e}^{-k\omega t\mathrm{i}}\right].$$

Then the two-dimensional embedding is the complex scalar signal

$$X_{\text{hilbert}}(t) = x(t) + \mathrm{i}\tilde{x}(t) \in \mathbb{C}.$$

For discrete finite time series of stepping data the corresponding discrete Fourier Transform (based on FFT) is used. The signal $X_{\text{hilbert}}(t)$ oscillates around a non-zero mean with an amplitude that is determined by the dimensions of the force measurements. To obtain non-dimensionalized quantities for further analysis we use the re-scaled signal $X(t)$:

$$X(t) = \frac{X_{\text{hilbert}}(t) - \text{mean}(X_{\text{hilbert}})}{\max_t |X_{\text{hilbert}}(t) - \text{mean}(X_{\text{hilbert}})|}. \tag{5.1}$$

Using the representation of the signal $X(t)$ in its trigonometric or exponential form one can determine its instantaneous amplitude R and phase ψ:

$$X(t) = |X(t)|\left[\cos\psi(t) + \mathrm{i}\sin\psi(t)\right] = R(t)\mathrm{e}^{\mathrm{i}\psi(t)}.$$

In Figure 5.1(a) we show the Hilbert Transform embedding of the stepping time profiles in the complex plane. The part of the approximate phase portrait following an approximate ellipse corresponds to regular stepping episodes and the excursions toward the origin correspond to freezing episodes. Figure 5.1(b) shows the instantaneous amplitude $R(t)$ and phase $\psi(t)$ of $X(t)$, and the original normalized data, respectively. The red dots in the Figure 5.1(b) mark the force maxima (as determined by graph in the bottom panel) of each oscillation also in time-amplitude and time-phase plots. We can see that these peaks occur near $\psi = 2\pi$. They correspond to the phase at which the left foot of the subject (in this case ST31) reaches the ground.

5.1.2 Selection of transition intervals for individual freezing events

As we aim to study transitions from stepping to freezing we identify, for each freezing event k labelled by a domain expert, a *transition interval* $[t_{\text{start},k}, t_{\text{end},k}]$ that contains this transition. Figure 5.2 highlights two examples (labelled $k = \{A, B\}$) from the time series

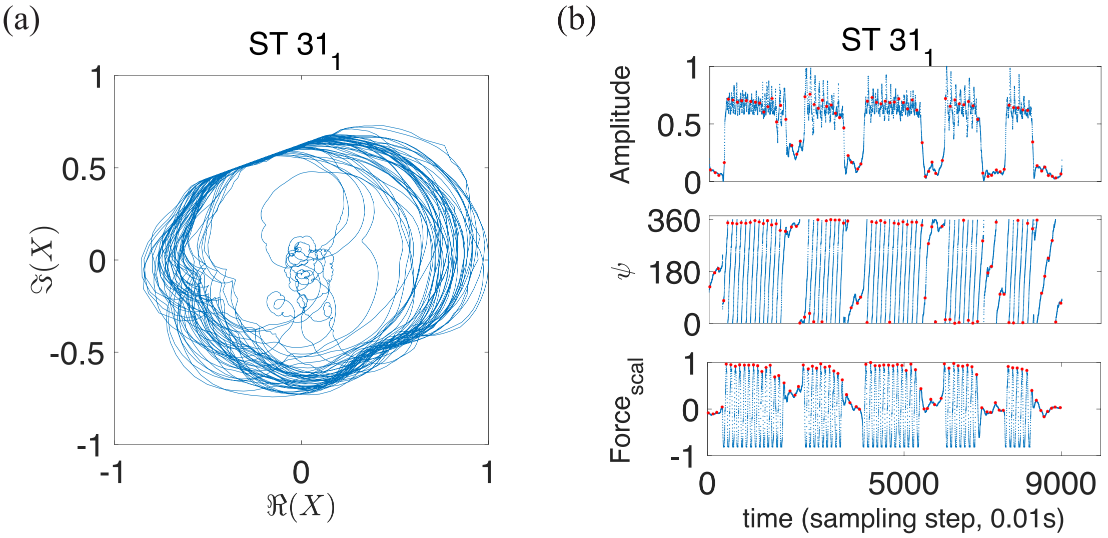


Figure 5.1: Panel(a): Embedding with scaled Hilbert Transform applied to stepping data from Figure 3.1 (see also right, bottom panel). Panel (b): Amplitude and Phase (ψ) of embedding. The red dots are locally maximum forces and their corresponding amplitude and phase.

shown in Figure 3.1 (ST31, data set 1). The interval boundaries are chosen such that $[t_{\text{start},k}, t_{\text{end},k}]$ contains at least 4 stepping oscillations and $t_{\text{end},k}$ is inside the part of the time series identified as freezing. For example, Figure 5.2 illustrates the highlighted events A and B extracted from ST31 (data set 1). Their transition intervals contain stepping and the beginning of the respective freezing episodes, in this case,

$$[t_{\text{start},A}, t_{\text{end},A}] = [1300, 1980], \qquad [t_{\text{start},B}, t_{\text{end},B}] = [4800, 5500].$$

As the choice of $t_{\text{start},k}$ and $t_{\text{end},k}$ is arbitrary, we will for all results below determine how they depend on the choice of $t_{\text{start},k}$ and $t_{\text{end},k}$ for each event k.

Figure 5.3(a) shows the data with Hilbert Transform for a transition interval that corresponds to freezing events extracted manually from the data (ST31, data set 1, time interval $[1300, 1980]$). Following our convention, each panel's label shows the

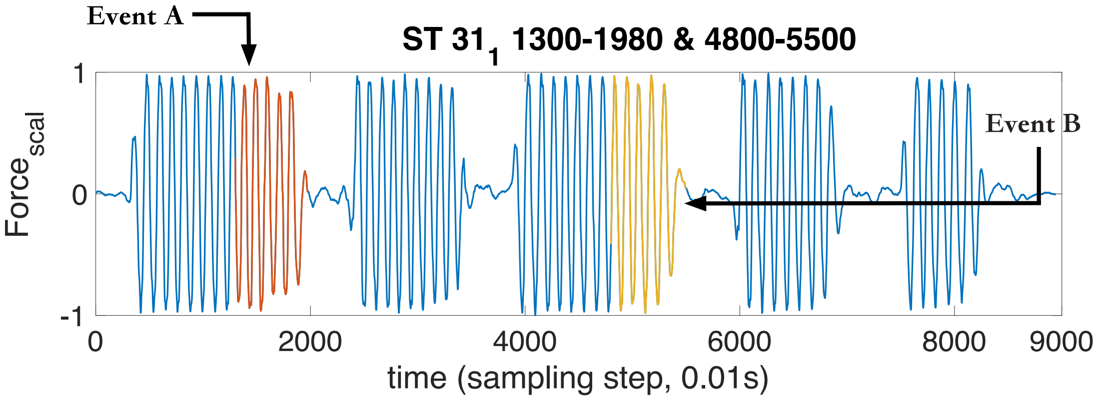


Figure 5.2: Sample events extracted from the original time series. Event A represents a portion of time series from ST31, data set 1 in subscript and time interval $[1300, 1980]$. Event B represents another portion of time series from ST31, data set 1 in subscript and time interval $[4800, 5500]$.

subject number with subscript indicating the number of the data set (we note that each subject has repeated the experiment several times and hence we have more than one data recordings set for each participant), followed by the pair of transition interval boundaries $[t_{\mathrm{start},k}, t_{\mathrm{end},k}]$. As expected for transition intervals corresponding to a single freezing episode, the embedded trajectories initially follow an ellipse during regular stepping before they approach the area near the origin of the complex plane (only once, in contrast to Figure 5.1(a), since each transition interval contains exactly one freezing episode and respectively one transition).

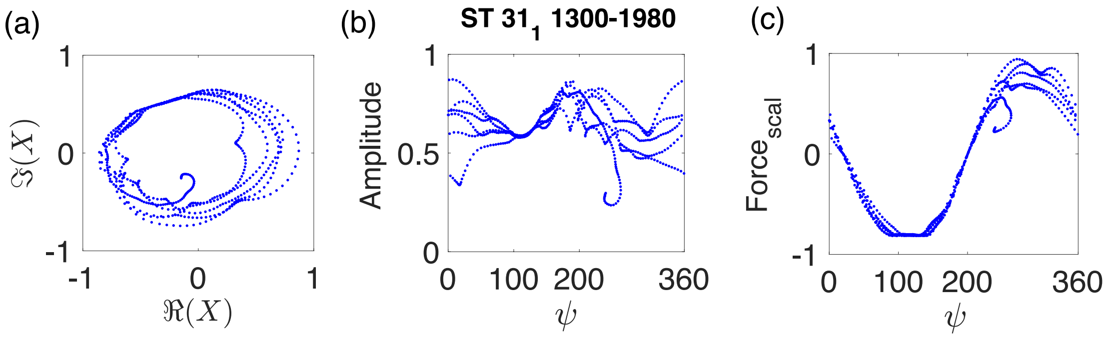


Figure 5.3: Embedded time series shown in different coordinate systems. Panel (a): embedded trajectory with Hilbert Transform. Panel (b): x-axis is angle ψ in degrees, y-axis is amplitude. Panel (c): x-axis is angle ψ in degrees, y-axis is scaled force.

Figure 5.3(b) and Figure 5.3(c) show the same set of embedded trajectories in two further coordinate systems, polar coordinates (x-axis is phase ψ in degrees, y-axis is amplitude) in Figure 5.3(b), and Cartesian coordinates x-axis is angle ψ in degrees, y-axis is scaled force: $\mathrm{Force}_{\mathrm{scal}}$ given in (3.1)) in Figure 5.3(c).

The representative transition in Figure 5.3 appears at first sight well described by the spontaneous transitions to the origin as they occur in the null model established by the Hopf normal form (4.3), up to a coordinate transformation that maps the ellipse followed during regular stepping onto the unit circle. However, we hypothesize a breaking of rotational phase symmetry, which would be a qualitative and practically relevant difference to the Hopf normal form. The following sections will describe a method to identify the location of the transition in the embedded phase plane uniquely. Then we will apply this method to all freezing transitions in the available data sets to investigate our hypothesis provisionally.

5.1.3 Motivation for discrete-time discrete-space Markov chain

Figure 5.4 and Figure 5.5 show a survey of Hilbert embeddings for freezing transition intervals from patient data. Figure 5.4 uses Cartesian coordinates X and Figure 5.5 uses polar coordinates (ψ, R). Especially Figure 5.5 highlights that fixing a threshold (for, e.g., amplitude R in polar coordinates as R_{th}) and defining the time of transition (for example) as the first crossing of this threshold (e.g., first time $t_{\mathrm{th}} \in [t_{\mathrm{start},k}, t_{\mathrm{end},k}]$ when $R(t_{\mathrm{th}}) \leq R_{\mathrm{th}}$) will introduce an extreme dependence of the timing t_{th} and the angle coordinate $\psi(t_{\mathrm{th}})$ on this threshold value. Furthermore, the threshold value will have to be adjusted for different events (possibly even between events for the same subject and data set), making the collection of generalizable statistics across subjects and events impossible and reinforcing the need for a subject-specific or personalized approach. In contrast to the sensitivity of timing and angle of the transition to the threshold, the *detection* of the presence of a freezing event with the help of thresholds is robust. As mentioned in Figure 3.1, freezing is uniquely determined by the absence of zero-force periods for longer than the identified stepping frequency.

Embedded time series for different subjects

Figure 5.4 shows the embedding of time series for different subjects with freezing events, obtained by applying the Hilbert Transform. Figure 5.4 shows the same embedded trajectories as Figure 5.5 in polar coordinates (ψ, R), where ψ is in degree and R is scaled to $[0,1]$ after Hilbert Transform.

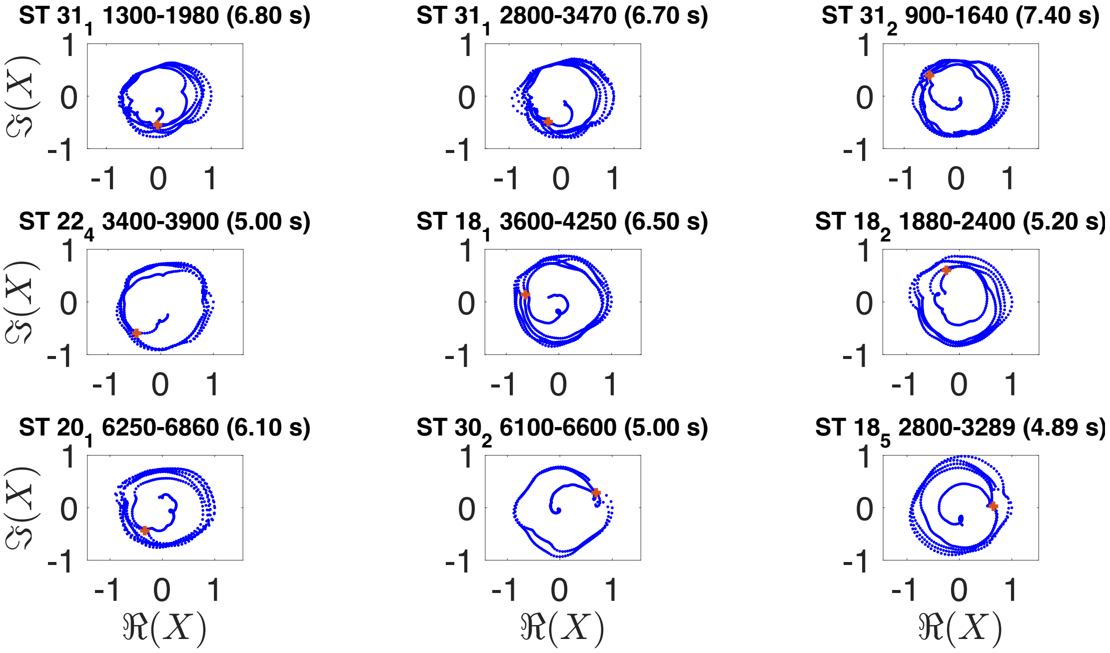


Figure 5.4: Embedded trajectories in $\mathbb{C}$ after scaling and Hilbert transform for 9 selected events and transition intervals. Each panel title specifies subject number, subject trial data set and sampling step numbers and resulting length of transition interval in seconds. See also step 1 of Figure 5.11. Red crosses indicate the cartesian coordinates $X_{\min}$ of boxes with minimal escape time from transition set.

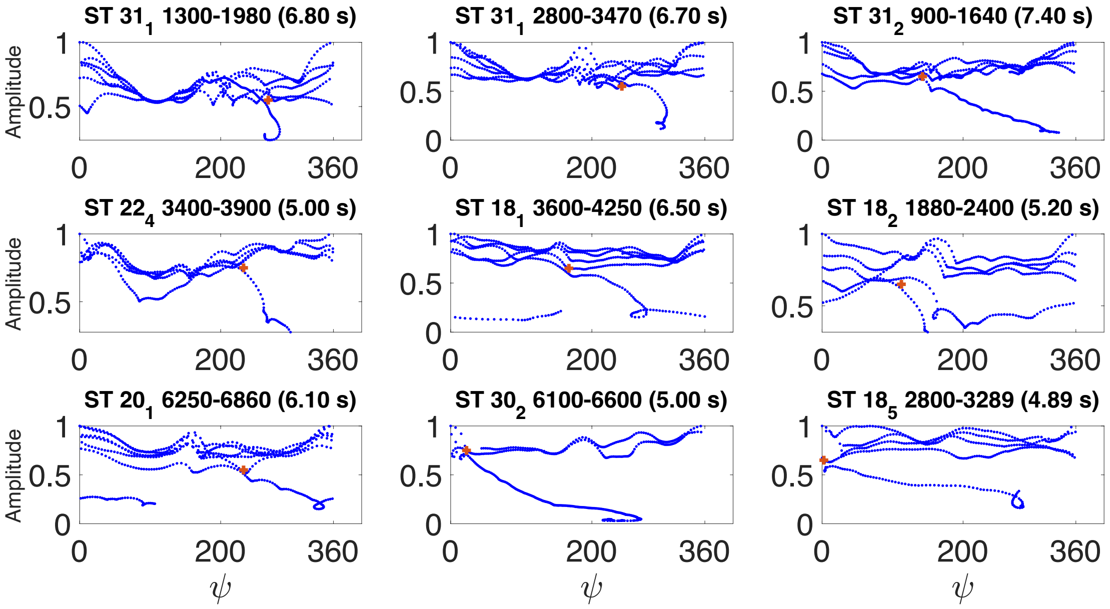


Figure 5.5: Same embedded trajectories as Figure 5.4 in polar coordinates (ψ, R), where ψ is in degree and R represents the amplitude. Red crosses indicate the polar coordinates $(R_{\min}, \psi_{\min})$ of boxes with minimal escape time from transition set.

We exploit the unique identifiability of the presence of freezing by constructing a discrete-time discrete-space Markov chain and its associated transition matrix for the variable X obtained from the Hilbert Transform embedding. This Markov chain divides the phase space into a transition set and an absorbing set (see below for precise definitions). The starting time into the freezing event can then be uniquely defined as the transition between these two subsets of the phase space. The red crosses in Figure 5.4 and Figure 5.5 show the transition points resulting from the procedure described below. They are clearly not determined by an amplitude coordinate, but we will have to test how strongly they depend on discretization parameters for our Markov chain (see Figure 1 in Appendix) and independent of the selected transition interval $[t_{\text{start},k}, t_{\text{end},k}]$ (see Figure 2 in Appendix).

5.1.4 Subdivision of the complex plane along polar coordinates

We subdivide the complex plane (the embedding space after Hilbert Transform) into boxes that are rectangular in polar coordinates as illustrated in Figure 5.6. Thus, the box boundaries are aligned with the radial direction and angular direction, respectively.

- By the scaling of our data the amplitude of $X(t)$ is always in $(0,1)$, such that we subdivide in the radial direction into P annuli of equal radial thickness $p = 1/P$. We enumerate starting from the origin.
- We subdivide in the angular direction anticlockwise, starting from positive x-axis into Q cones of equal size $q = 360°/Q$.

This results in $P \times Q$ discrete states in total, corresponding to the near rectangular boxes described above. A box $B_{k,\ell} \subset \mathbb{C}$ for $(k,\ell) \in \{1,\dots,P\} \times \{1,\dots,Q\}$ is then

$$B_{k,\ell} = \left\{ X \in \mathbb{C} : R(X) \in \left(\frac{k-1}{P}, \frac{k}{P} \right], \psi(X) \in \left[\frac{2\pi(\ell-1)}{Q}, \frac{2\pi\ell}{Q} \right) \right\},$$

where we use $R(X) \in (0,1]$ for the amplitude and $\psi(X) \in [0, 360°)$ for the argument of a point $X \in \mathbb{C}$. We enumerate the boxes in angle-first order such that the box with radius-angle index (k,ℓ) is at position $i = \ell + Q(k-1)$. The above subdivision defines index maps $\mathrm{ind}_{\mathrm{rad}} : \mathbb{C} \to \{1,\dots,P\}$, $\mathrm{ind}_{\mathrm{ang}} : \mathbb{C} \to \{1,\dots,Q\}$ and $\mathrm{ind} : \mathbb{C} \to \{1,\dots,PQ\}$ as follows:

$$\mathrm{ind}_{\mathrm{rad}}(X) = \mathrm{ceil}(R(X)P) \tag{5.2}$$

$$\mathrm{ind}_{\mathrm{ang}}(X) = \mathrm{floor}(\psi(X)Q/(2\pi)) + 1 \tag{5.3}$$

$$\mathrm{ind}(X) = (\mathrm{ind}_{\mathrm{rad}}(X) - 1)Q + \mathrm{ind}_{\mathrm{ang}}(X). \tag{5.4}$$

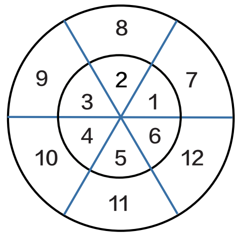
For box number i we may recover the corresponding annulus k and cone ℓ via

$$\ell = i - Q(\text{floor}(i/Q) - 1), \qquad k = \text{floor}(i/Q), \tag{5.5}$$

such that

$$\text{ind}_{\text{ang}}(X) = \text{ind}(X) - Q \times (\text{floor}(\text{ind}(X)/Q) - 1), \quad \text{ind}_{\text{rad}}(X) = \text{floor}(\text{ind}(X)/Q)$$

for every $X \in \mathbb{C}$.



q = 60°   *p* = 0.5

Figure 5.6: Subdivision and discretization of the unit circle in the complex plane. The numbering is anticlockwise and from the pole to the edge of the space with $P = 2$, $Q = 6$ and, hence, $p = 0.5$ and $q = 60°$. Note that typical tested discretizations are much finer: Figure 1 tests the range $q \in \{3°, 5°, 10°, 15°, 20°, 30°\}, p \in \{0.05, 0.1, 0.15, 0.2\}$.

In principle, we may consider the set $\{1, \ldots, PQ\}$ as the state space of our discrete-time discrete-space Markov chain. However, each box index corresponds to an approximate location inside the complex unit circle, given by its mid point with respect to radius and angle:

$$X_{\text{c},i} = \frac{k - 1/2}{P} \exp\left(\frac{2\pi\text{i}(\ell - 1/2)}{Q}\right), \tag{5.6}$$

where k and ℓ are related to i via (5.5). Thus, X_c is a vector in $\mathbb{C}^{PQ}$, and for each $X \in \mathbb{C}$

we can find the midpoint of the box in which X is located by

$$\begin{aligned} X_{\mathrm{c}}(X) &= X_{\mathrm{c},\mathrm{ind}(X)}, \quad \text{with coordinates} \\ R_{\mathrm{c}}(X) &= \frac{\mathrm{ind}_{\mathrm{rad}}(X)-1/2}{P} \\ \psi_{\mathrm{c}}(X) &= \frac{2\pi\mathrm{i}(\mathrm{ind}_{\mathrm{ang}}(X)-1/2)}{Q}. \end{aligned}$$

We may use X_{c} as the state space of the discrete-space Markov chain instead of the index in $\{1,\ldots,PQ\}$.

5.1.5 An empirical Markov chain transition matrix for a single transition interval

For the discrete state space $X_{\mathrm{c}} \sim \{1,\ldots,PQ\}$, defined in (5.6), we use the embedded trajectory $X(t) \in \mathbb{C}$ with t in a single transition interval $[t_{\mathrm{start}}, t_{\mathrm{end}}]$ to construct a provisional transition matrix A_{c} for probability distributions $P_{\mathrm{prob}} : \{1,\ldots,PQ\} \to [0,1]$:

$$P_{n+1} = P_n A_{\mathrm{c}} \tag{5.7}$$

For a given embedded trajectory $X(t)$ with $t \in [t_{\mathrm{start}}, t_{\mathrm{end}}]$ (recall that t were integers) we define the transition count $\mathrm{CT} \in \mathbb{Z}^{PQ\times PQ}$ as follows:

$$\mathrm{CT}_{ij} = |\{t \in [t_{\mathrm{start}}, t_{\mathrm{end}}-1] : \mathrm{ind}(X(t)) = i \text{ and } \mathrm{ind}(X(t+1)) = j\}| \tag{5.8}$$

for $i, j \leq PQ$ (we use the notation $|S|$ for the number of elements in a set S in (5.8)). The provisional transition matrix A generated by $X(t)$ is then:

$$\{A_{\mathrm{c}}\}_{ij} = \begin{cases} \dfrac{\mathrm{CT}_{ij}}{\sum_{k=1}^{PQ} \mathrm{CT}_{ik}} & \text{if } \sum_{k=1}^{PQ} \mathrm{CT}_{ik} > 0, \\ 0 & \text{otherwise} \end{cases} \tag{5.9}$$

The entry $\{A_{\mathrm{c}}\}_{ij}$ is then the empirical probability (as determined by counting) that $X(t+1)$ is in state j if $X(t)$ is in state i. By construction the row sums of A satisfy $\sum_j \{A_{\mathrm{c}}\}_{ij} = 1$ or $\sum_j \{A_{\mathrm{c}}\}_{ij} = 0$. We call

$$X_{\mathrm{emp}} = \{X_{\mathrm{c},i} : \sum_{k=1}^{PQ} \mathrm{CT}_{ik} > 0\} \quad \text{(in complex coordinates)}, \tag{5.10}$$

$$\mathscr{I}_{\mathrm{emp}} = \{i : \sum_{k=1}^{PQ} \mathrm{CT}_{ik} > 0\} \subseteq \{1, \ldots, PQ\} \quad \text{(in integer coordinates)}, \tag{5.11}$$

the state space with empirical support. We call its size

$$n_{\mathrm{emp}} = |\mathscr{I}_{\mathrm{emp}}|.$$

We construct an empirical transition matrix A which is restricted to $\mathscr{I}_{\mathrm{emp}}$. The transition matrix A for state space X_{emp} allows us to either generate surrogate time series by starting from a state $X_{\mathrm{emp},i}$ at time 0 and then generating the state at time $n+1$ by performing a multi-outcome Bernoulli trial with outcomes in X_{emp} and probabilities $A_{i,\mathscr{I}_{\mathrm{emp}}}$, and then changing the state at time $n+1$ to the outcome of the trial. Figure 5.7(b) depicts such a surrogate time series. Figure 5.7(c) shows the same time series, superimposing it onto the real component of the embedded trajectory, which is the coordinate of the

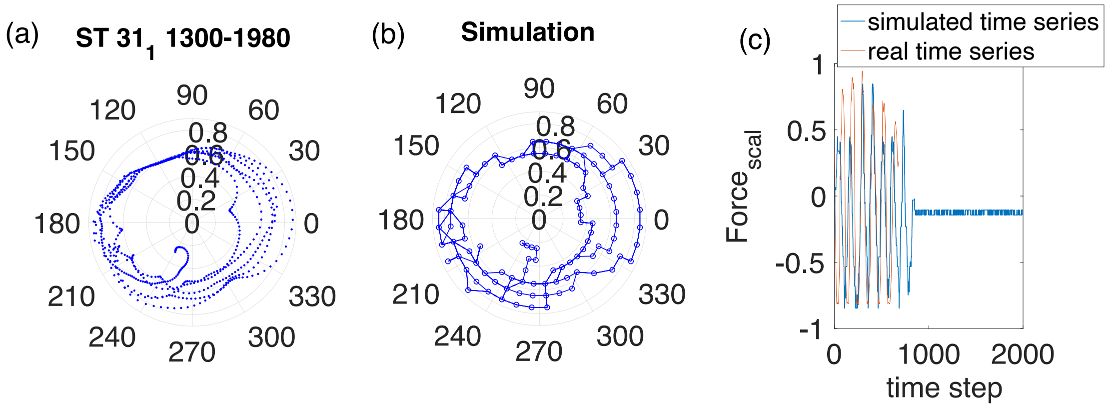


Figure 5.7: Embedded trajectory and surrogate time series for ST31, data set 1, transition interval $[1300, 1980]$ and discretization parameters $(p, q) = (0.1, 5^\circ)$. Panel (a) shows the embedded trajectory $X(t)$. Panel (b) demonstrates a single surrogate time series while Panel (c) shows the projection onto the real part, the coordinate of the embedding corresponding to the scaled forced data (blue: surrogate, red: data).

original signal.

5.2 Properties of the constructed Markov chain

In section 5.1.3 we have proposed a general method to derive a Markov chain transition matrix from time series data. In this section we will first classify the long-time behavior of a Markov chain by constructing the zero-pattern matrix to partition the chains into communicating classes as described in Section 2.3. Then we will divide the states into *Transition set* (corresponding to stepping) and *Absorbing set* (corresponding to freezing). Finally we will determine the mean escape time starting for each state in the Transition set.

Our construction is valid for a general Markov transition matrix $A = A_{ij} \in [0,1]^{n\times n}$, which has the following general properties given by G.Rudolph [42]:

- The eigenvalues of matrix A have modulus less or equal to 1.
- All its row sums are equal to 1, $\sum_j a_{ij} = 1$, thus $\mathbf{e} = (1,1,...1)^T$ is the right eigenvector corresponding to a unit eigenvalue.
- The entries are all non-negative.

We have $\pi_i A = \lambda_i \pi_i$, where π_i are the left eigenvectors, and λ_i are the eigenvalues of the Markov transition matrix A.

5.2.1 Partition of state space into classes

Markov chains as constructed from our data have transition matrices where most non-zero diagonal entries, because the boxes $B_{k,\ell}$ are large in phase space compared the distance covered during a single sampling time step. So, we expect each state to be accessible from itself, even if we use $n > 0$ in the above definition. Later in Subsection 5.2.2 we make the specific assumptions (ergodicity and aperiodicity within each communicating class) covering this issue.

Communicating classes only depend on the positivity of entries A_{ij} such that we may introduce the indicator matrix Z and the reachability matrix Z_∞:

$$Z_{ij} = \begin{cases} 1 & \text{if } A_{ij} > 0, \\ 0 & \text{otherwise,} \end{cases}, \qquad Z_{\infty,ij} = \begin{cases} 1 & \text{if } A^n_{ij} > 0 \text{ for some } n > 0, \\ 0 & \text{otherwise.} \end{cases}$$

If we consider states i and j connected when $A_{ij} > 0$, then the transition matrix A induces a graph, for which Z is the adjacency matrix. Thus, we may use graph theoretic methods to determine the reachability matrix Z_∞ (in Matlab the routine `transclosure`). From Z_∞ we may construct the bidirectional reachability matrix $Z_{\text{bi}} = \min(Z_\infty, Z_\infty^\mathsf{T})$, that is,

$$Z_{\text{bi},ij} = \begin{cases} 1 & \text{if } Z_{\infty,ij} = 1 \text{ and } Z_{\infty,ji} = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Two states i and j are in the same communicating class if the rows $Z_{\text{bi},i,(\cdot)}$ and $Z_{\text{bi},j,(\cdot)}$ are identical. Let the *class indicator matrix* $Z_{\text{cc}} \in \{0,1\}^{n_{\text{cc}} \times n_{\text{emp}}}$ be the matrix consisting of the n_{cc} unique rows of Z_{bi} (typically $n_{\text{cc}} \ll n_{\text{emp}}$), then each row of Z_{cc} corresponds to a communicating class of transition matrix A, and state i is in communicating class k for A if $Z_{\text{cc},k,i} = 1$. Communicating classes are partially ordered: we write that class

$$k_1 \to k_2$$

(with $k_1 \neq k_2$) if states from class k_2 are reachable from states in class k_1 (if reachability is true for one pair of states in classes k_1 and k_2 it is true for all pairs of states).

5.2.2 Communicating classes of empirical transition matrices — stepping class, transition set and absorbing set

The transition intervals $[t_{\text{start}}, t_{\text{end}}]$ in Section 5.1.2 are chosen such that the time series $x(t)$ starts in the stepping regime (large scaled force oscillations), stays there for most of the time and ends in a freezing episode. Thus, we expect the states of the Markov chain to fall into several communicating classes (see Figure 5.7(a)). Appendix A studies systematically the dependence of the number of communicating classes, their ordering and their geometric properties (e.g., phase angles ψ_{c} of the boxes where transition between classes is most likely, see Section 5.4) on discretization parameters P and Q and transition interval $[t_{\text{start}}, t_{\text{end}}]$. Our results in Appendix A show that this dependence is weak. Thus, communicating classes are a suitable object for studying the qualitative properties of our time series data. We expect communicating classes of our data from class indicator matrix Z_{cc} of the following types and with the following properties.

- **(Stepping)** We expect one large communicating class corresponding to a row in the class indicator matrix Z_{cc} with almost all the values with empirical support equal to 1. We denote this communicating class as the *stepping class*, naming it F_{step} and denoting its row index in Z_{cc} by k_{step}.

- **(Ergodicity)** By definition of a communicating class, all states j in the stepping class F_{step} are positively recurrent. We also assume that they are aperiodic. This is an assumption that the discretization box sizes in the Hilbert Transform plane $\mathbb{C}$ should not be too small compared to the sampling time step: at least some boxes from the discretization should contain several sampling time steps from the sampled time series such that the empirical probability of staying in the box during a time step is non-zero. With this aperiodicity assumption the stepping class is ergodic when one considers the conditional transition probabilities $P_{\text{step},i,j}$ under the condition that the Markov chain stays in the stepping class, defined by

$$P_{\text{step},i,j} = P(X_{n+1} = i \,|\, X_n = j \text{ and } X_{n+1} \in F_{\text{step}}) \text{ for } i, j \in F_{\text{step}}. \tag{5.12}$$

- **(Ordering relative to stepping class)** We also expect that all other classes can be related to F_{step} through the partial ordering (because F_{step} is large). That is, each class k satisfies $k \to k_{\text{step}}$ or $k_{\text{step}} \to k$.
- **(Transition set)** We collect the states in all communicating classes k with $k \to k_{\text{step}}$ (so coming before F_{step} in the partial ordering) in the so-called initial set F_0. The *transition set* is the union of stepping class F_{step} and initial set F_0, $F = F_0 \cup F_{\text{step}}$.
- **(Absorbing set)** We collect all communicating classes k with $k_{\text{step}} \to k$ (so coming after F_{step} in the partial ordering) in the so-called absorbing set E. We expect the absorbing set E to be non-empty.

Figure 5.8 shows transition set and absorbing set for the example transition matrix generated from ST31, data set 1, transition interval $[1300, 1980]$, and discretization parameters $p = 0.1$, $q = 5^\circ$. Panel (a) shows all boxes with empirical support in the form of circles at their centers $X_{\text{emp},i}$. The red circles are in boxes that belong to the absorbing set E, and the blue circles are in boxes that belong to transition set F. We observe that there is a unique state i in the absorbing set with $A_{ij} > 0$ for a state j in the transition set and highlight this point with a black cross. Figure 5.8(b) and (c) show the time series $x(t)$ (scaled force measurements) and its embedding in the unit circle of the complex plane. We color the points of the time series according to the class membership of the box they are in: red marking indicates that $\text{ind}(X(t))$ is in the absorbing set E. Blue

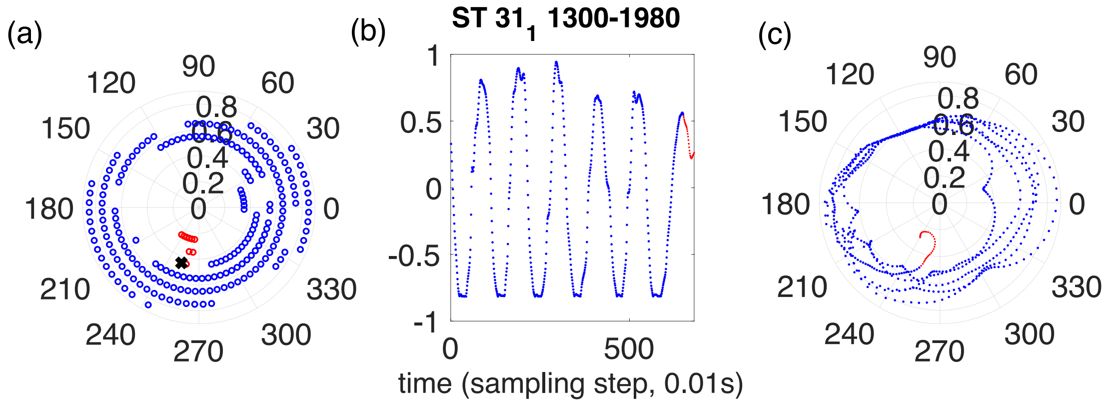


Figure 5.8: Panel (a): empirical support for Markov chain for time series ST31 and decomposition into transition set F (blue circles) and absorbing set E (red circles) with discretization parameters $(p,q) = (0.1, 5^\circ)$. The black cross is the "first" state, X_{tr}, in the absorbing set E (see Figure 5.2.4). Panel (b): underlying time series for panel (a), color coding the sampling points according to their location in transition set (blue) or absorbing set (red). Panel (c): same time series and color coding in phase plane obtained by Hilbert Transform.

marking indicates that $\text{ind}(X(t))$ is in the transition set F.

5.2.3 Mean escape time from the transition set

Suppose we can decompose the Markov chain given by transition matrix A into a Transition set F with m states and an Absorbing set E with $k = n - m$ states. Let $\mathbf{s} \in \mathbb{R}^{1\times m}$ be an initial probability distribution for states in the transition set F. Let p_i^t denote the probability that the Markov chain is in state i at step t from the starting distribution $\mathbf{s}$. We denote the probability distribution at step t by $\mathbf{p}^t$

$$\mathbf{p}^t = (p_1^t, p_2^t, \ldots, p_n^t).$$

Then the distribution $\mathbf{p}^t$ satisfies:

$$\mathbf{p}^0 = \mathbf{s}, \quad \mathbf{p}^t = \mathbf{s}A^t$$

We can extract the probability that the Markov chain is in the absorbing set at step t by defining the indicator function (or vector)

$$\boldsymbol{\xi} = (\xi_1, \xi_2, ..., \xi_n)^T$$

$$\xi_i = \begin{cases} 1 & i \in E \\ \\ 0 & i \in F \end{cases}$$

Let E^t denote the event that the Markov chain is in the Absorbing set at step t and F^t denote the event that the Markov chain is in Transition set at step t. Correspondingly, $P(E^t)$ is the probability that the Markov chain is in the absorbing set at step t and $P(F^t)$ is the probability that the Markov chain is in the transition set at step t. By applying the indicator vector $\boldsymbol{\xi}$, $P(E^t)$ can be calculated by:

$$P(E^t) = \mathbf{s}A^t\boldsymbol{\xi}$$

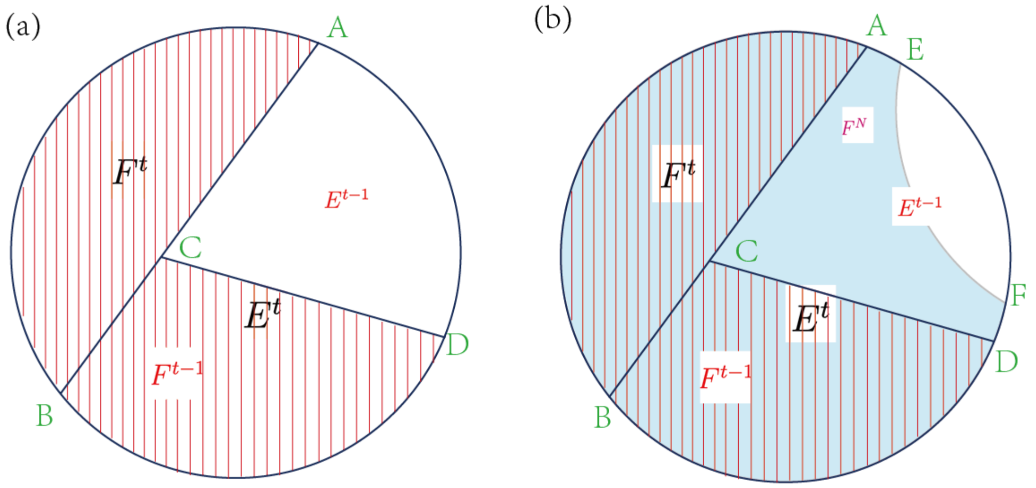


Figure 5.9: Venn diagram, a,b,c,d,e,f are the endpoints of the lines or arc which split the set. Left diagram (a): Event F^t is in the left part of line AB which means the Markov chain is in transition set at step t, then then right part of line AB demonstrates event E^t which means the Markov chain is in absorbing set at step t. The area ABCD covered by red lines demonstrates event F^{t-1} so the left area ACD is event E^{t-1}. Right diagram (b): Area in blue shows the event F^N and all other areas demonstrate the same event as diagram (a).

Suppose the support of the initial distribution is in the Transition set (so, $\mathbf{s}_\ell = 0$ for $\ell \in E$ $P(F^0) = 1$, $P(E^0) = 0$), we say the Markov chain has a **first escape** at step t when the Markov chain is in the Absorbing set at step t and still in the Transition set at step $t-1$. Venn diagram 5.9(a) illustrates the relations between the events involved, event F^t is to the left of line AB and event E^t is to the right of line AB, with $P(E^t) + P(F^t) = 1$. Regarding the event F^{t-1}, if at step t, the Markov chain is in the Transition set, then at step $t-1$ Markov chain must be in the Transition set as well. The other situation is at step t, the Markov chain is in the Absorbing set while it is in the Transition set at step $t-1$, then event F^{t-1} can be seen as the union of these two situations, shown as the area with red vertical stripes in Figure 5.9(a). Event E^{t-1} is the left area of the whole set. The probability of the Markov chain **first escape** at step t can be described as $P(E^t \setminus E^{t-1})$. We can see from the diagram (a) in Figure 5.9, $E^{t-1} \subseteq E^t$, therefore

$P(E^t \setminus E^{t-1}) = P(E^t) - P(E^{t-1})$. Hence, the mean escape time from the transition set, denoted MET(**s**), starting from initial distribution **s** equals

$$\mathrm{MET}(\mathbf{s}) = \sum_{t=1}^{\infty} tP(E^t \setminus E^{t-1}) \tag{5.13}$$

$$= \sum_{t=1}^{\infty} t\left[P(E^t) - P(E^{t-1})\right] \tag{5.14}$$

$$= \sum_{t=1}^{\infty} t\left[\mathbf{s}A^t\boldsymbol{\xi} - \mathbf{s}A^{t-1}\boldsymbol{\xi}\right] \tag{5.15}$$

$$= \mathbf{s}\sum_{t=1}^{\infty} tA^{t-1}(A-I)\boldsymbol{\xi} \tag{5.16}$$

By applying Taylor series $\sum_{t=1}^{\infty} tA^{t-1} = (I-A)^{-2}$, for regular $I-A$, equation (5.16) can be simplified as:

$$\mathrm{MET}(\mathbf{s}) = \mathbf{s}(A-I)^{-1}\boldsymbol{\xi} \tag{5.17}$$

However, since $\mathbf{e} = (1,1,...1)^T$ is the right eigenvector corresponding to a unit eigenvalue, $A\mathbf{e} = \mathbf{e}$, we will get $(A-I)\mathbf{e} = 0$ such that $\det(A-I) = 0$. Hence, $(A-I)$ is a singular matrix and has no inverse matrix. To solve this problem, we need the following preparatory work:

- For the convenience of presenting the derivation, change the order of all the absorbing states, so the first k states are in the Absorbing set E, states in matrix B

are corresponding to the transition set F, then we will have:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \left[\begin{array}{ccc|ccc} d_{11} & \cdots & d_{1k} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ d_{k1} & \cdots & d_{kk} & 0 & \cdots & 0 \\ \hline c_{11} & \cdots & c_{1k} & b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mk} & b_{m1} & \cdots & b_{mm} \end{array}\right] = \begin{bmatrix} D & \mathbf{0} \\ C & B \end{bmatrix}$$

where matrix D is a Markov transition matrix as well, $\sum_i d_{ji} = 1$.

- To be consistent with the reordering of the sequence of the absorbing states, we should also change the order of the starting distribution $\mathbf{s}$, then the first k states are in the Absorbing set. Since we are not starting from the Absorbing set, the first k elements should be 0, we can say $\mathbf{s} = (\mathbf{0}, \mathbf{s_1})$.

- For the reordered sequence of states the indicator vector $\boldsymbol{\xi} = (\xi_1, \xi_2, ..., \xi_n)^T$ will have the first k states in the Absorbing set. So, $\xi_i = 1$ when state i is in the Absorbing set, such that

$$\boldsymbol{\xi} = (\xi_1, \xi_2, ..., \xi_k, \xi_{k+1}, ..., \xi_n)^T$$

$$\xi_i = \begin{cases} 1, & i \le k \\ 0, & i > k \end{cases}$$

- Let $\boldsymbol{\xi} = \mathbf{e} - \boldsymbol{\zeta}$, where $\mathbf{e} = (1, 1, ..., 1)^T$ which is the right eigenvector corresponding to the unit eigenvalue. Therefore

$$\zeta = (\zeta_1, \zeta_2, ..., \zeta_n)^T$$

$$\zeta_i = \begin{cases} 0, & i \le k \\ 1, & i > k \end{cases}$$

is the indicator vector of the Transition set.

We can now express the mean escape time(MET) using the right-hand side (5.15):

$$\begin{aligned} \mathrm{MET}(\mathbf{s}) &= \sum_{t=1}^{\infty} t \left[\mathbf{s}A^t (\mathbf{e}-\zeta) - \mathbf{s}A^{t-1} (\mathbf{e}-\zeta)\right] && (5.18) \\ &= -\sum_{t=1}^{\infty} t \left(\mathbf{s}A^t\zeta - \mathbf{s}A^{t-1}\zeta\right) && (5.19) \\ &= -\mathbf{s}\left(\sum_{t=1}^{\infty} t \left(\begin{bmatrix} D^t & \mathbf{0} \\ C_1 & B^t \end{bmatrix} - \begin{bmatrix} D^{t-1} & \mathbf{0} \\ C_2 & B^{t-1} \end{bmatrix}\right)\right)\zeta && (5.20) \end{aligned}$$

C_i, D_i are matrices that do not matter, same later.

$$(5.21)$$

$$\begin{aligned} &= -\mathbf{s}\left(\sum_{t=1}^{\infty} t \begin{bmatrix} D^t - D^{t-1} & \mathbf{0} \\ C_1 - C_2 & B^t - B^{t-1} \end{bmatrix}\right)\zeta && (5.22) \\ &= -(\mathbf{0}, \mathbf{s_1}) \begin{bmatrix} D_1 & \mathbf{0} \\ C_3 & (B-I)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix} && (5.23) \end{aligned}$$

$$\mathrm{MET}(\mathbf{s}) = -\mathbf{s_1}(B-I)^{-1}\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix} \tag{5.24}$$

$$= -\mathbf{s_1}(B-I)^{-1}\mathbf{e}^{\mathsf{T}}. \tag{5.25}$$

Here B is a $m \times m$ matrix so that $(1,...,1)^{-1}$ is a $m \times 1$ matrix with all rows equal to 1. We can also conclude from equation (5.25) that the MET is independent of the transition probabilities in the absorbing set. Let us rename the transition matrix B, which is A restricted to the transition set F by $A_F \in \mathbb{R}^{m\times m}$. We can rewrite (5.27) as:

$$\mathrm{MET}(\mathbf{s}) = -\mathbf{s_1}(B-I)^{-1}\mathbf{e}^{\mathsf{T}} \tag{5.26}$$

$$= \mathbf{s_1}\,(I-A_F)^{-1}\,\mathbf{e}^{\mathsf{T}}. \tag{5.27}$$

5.2.4 Preferred transition states

Figure 5.10 shows the mean escape times for $i \in F$ and initial distributions equaling unit vectors,

$$\mathrm{MET}_i = \mathrm{MET}(\mathbf{u}_j) = \mathbf{u}_j(I-A_F)^{-1}\mathbf{e}^{\mathsf{T}}, \text{ where } i \text{ is the } j\text{th element of } F, \tag{5.28}$$

and $u_{j,j} = 1$ and $u_{j,\ell} = 0$ for $\ell \neq j$ ($\mathbf{u}_j \in \mathbb{R}^{1\times m}$). Thus, MET_i is the mean escape time when we are starting from a known state i in the transition set F. Figure 5.10(b) shows the top view and highlights the state $i_{\min}$ with minimal MET_i and its complex coordinate

X_{min}:

$$i_{\text{min}} = i \text{ such that such that } \text{MET}_i \text{ is minimal, and} \tag{5.29}$$

$$X_{\text{min}} = X_{\text{emp},i_{\text{min}}}, \quad R_{\text{min}} = R(X_{\text{emp},i_{\text{min}}}), \quad \psi_{\text{min}} = \psi(X_{\text{emp},i_{\text{min}}}). \tag{5.30}$$

The center of the box X_{min} is marked by a black '+' in Figure 5.10(b). The location of X_{min} is close to the (in this case unique) “first” state X_{tr} (with amplitude $R_{\text{tr}} = R(X_{\text{tr}})$, phase $\psi_{\text{tr}} = \psi(X_{\text{tr}})$ and index $i_{\text{tr}} \in E$) in the absorbing set into which one may transition from F, which is highlighted by a black cross in Figure 5.8(a). If i_{tr} is unique, then it is its own communicating class such that it comes indeed first in the partial ordering within E and follows directly in the partial ordering after F. The boxes X_{tr} and X_{min} are naturally close together whenever a unique $i_{\text{tr}} \in E$ exists in the sense of our partial ordering. While the mean transition times will clearly depend on the choice of boundaries for the transition interval $[t_{\text{start}}, t_{\text{end}}]$, the transition set F, the absorbing set E and X_{min} and X_{tr} may not (or only weakly) depend in t_{start} or t_{end}. Similarly, discretization parameters P and Q may only weakly affect X_{min} and X_{tr}. If these weak or non-dependences are true then a systematic collection of X_{min} (and ψ_{min}) from subject data sets will be able to determine whether the phase invariance implicitly assumed in the generalized Hopf normal form model is a valid assumption, or if (subject dependent or general) phases exist during which subjects are particularly vulnerable to freezing. Appendix A studied the dependence of X_{min} on all method parameters in details.

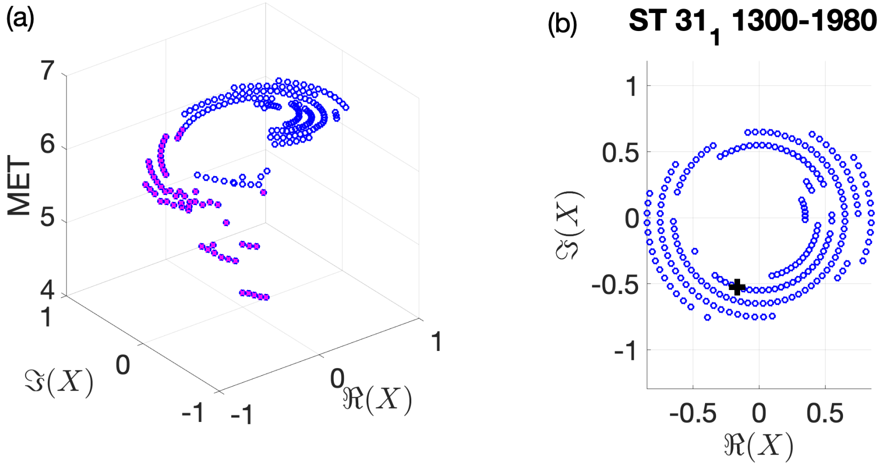


Figure 5.10: Panel (a): mean escape times from each state in transition set F for ST31 data set 1, time interval $[1300, 1980]$, discretization parameters $(p,q) = (0.1, 5^\circ)$. States marked by magenta crosses have MET below MET_F given in (5.41). Panel (B): top view, where black '+' marks $X_{\min}$, the state with minimal mean escape time.

5.2.5 Relationship between MET_i and the mean escape time from transition set F

The largest eigenvalue λ_1 (if there exist a largest in modulus eigenvalue) for the transition matrix A_F of the Transition set F is the escape rate (so, $\mathrm{MET}(\mathbf{s}) = 1/(1-\lambda_1)$ for a particular distribution $\mathbf{s}_F$, corresponding to the left eigenvector of A_F for λ_1. This distribution $\mathbf{s}_F$ is the so-called stationary distribution of the transition set F, that is, $\mathbf{s}_F$ is the limiting distribution of distributions $\mathbf{s}_N$,

$$\mathbf{s}_N = \mathbf{s}A_F^N/(\mathbf{s}A_F^N\mathbf{e}^\mathsf{T}) \to \mathbf{s}_F \quad \text{for } N \to \infty,$$

after N steps under the condition that the Markov chain stays in the transition set F for at least N steps. Therefore the probability of the Markov chain first escape from the Transition set F at step t under the condition that the Markov chain stays in F for at least

N steps can be concluded in the following formula (for $t > N$):

$$\begin{aligned} P(E^t \setminus E^{t-1} | F^N) &= \frac{P\left(\left(E^t \setminus E^{t-1}\right) \cap F^N\right)}{P(F^N)} \\ &= \frac{P(E^t) - P(E^{t-1})}{1 - P(E^N)} \end{aligned}$$

Note since $t - 1 \geq N$, we have $F^{t-1} \subseteq F^N$, which implies $E^t \setminus E^{t-1} \subseteq F^N$. Thus, $P\left(\left(E^t \setminus E^{t-1}\right) \cap F^N\right) = P(E^t \setminus E^{t-1}) = P(E^t) - P(E^{t-1})$. So, the mean escape time from F, assuming we have spent N steps in F, equals (comparing to (5.13)):

$$\mathrm{MET}(\mathbf{s})_F = \sum_{t=N+1}^{\infty} (t - N) \frac{P(E^t) - P(E^{t-1})}{1 - P(E^N)} \tag{5.31}$$

$$= \frac{\sum_{t=N+1}^{\infty} (t - N) \left[\mathbf{s}A^t \boldsymbol{\xi} - \mathbf{s}A^{t-1} \boldsymbol{\xi}\right]}{1 - \mathbf{s}A^N \boldsymbol{\xi}} \tag{5.32}$$

$$= \frac{\sum_{t=N+1}^{\infty} (t - N) \left[\mathbf{s}A^t \left(\boldsymbol{e} - \zeta\right) - \mathbf{s}A^{t-1} \left(\boldsymbol{e} - \zeta\right)\right]}{1 - \mathbf{s}A^N \left(\boldsymbol{e} - \zeta\right)} \tag{5.33}$$

$$= \frac{\sum_{t=N+1}^{\infty} -(t - N) \left[\mathbf{s}A^t \zeta - \mathbf{s}A^{t-1} \zeta\right]}{\mathbf{s}A^N \zeta} \tag{5.34}$$

$$= \frac{-\boldsymbol{s} \sum_{j=1}^{\infty} \left[j \left(A^j - A^{j-1}\right)\right] A^N \zeta}{\mathbf{s}A^N \zeta} \tag{5.35}$$

$$= \frac{-\mathbf{s}\sum_{j=1}^{\infty} j \begin{bmatrix} D_2 & \mathbf{0} \\ C_4 & B^j - A_F^{j-1} \end{bmatrix} \begin{bmatrix} D_3 & \mathbf{0} \\ C_5 & A_F^N \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{e} \end{bmatrix}}{\boldsymbol{s} \begin{bmatrix} D_3 & \mathbf{0} \\ C_5 & A_F^N \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{e} \end{bmatrix}} \tag{5.36}$$

$$= \frac{[0, \boldsymbol{s_1}] \begin{bmatrix} D_4 & \mathbf{0} \\ C_6 & (I - A_F)^{-1} \end{bmatrix} \begin{bmatrix} D_3 & \mathbf{0} \\ C_5 & A_F^N \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{e} \end{bmatrix}}{[0, \boldsymbol{s_1}] \begin{bmatrix} \mathbf{0} \\ A_F^N \mathbf{e} \end{bmatrix}} \tag{5.37}$$

$$= \frac{\boldsymbol{s_1} (I - A_F)^{-1} A_F^N \mathbf{e}}{\boldsymbol{s_1} A_F^N \mathbf{e}} \tag{5.38}$$

(using $\mathbf{0}$ for appropriately sized zeros matrices). We know that all eigenvalues of A_F have modulus < 1. We assume that A_F has one dominant eigenvalue $\lambda_1 (\lambda_1 > \lambda_2 > ... > \lambda_m)$ with left eigenvector π_1 such that $\boldsymbol{s_1} = \sum_i c_i \boldsymbol{\pi_i}$, so we can rewrite our the MET$(\boldsymbol{s})_F$ as:

$$\text{MET}(\mathbf{s})_F = \frac{\sum_i c_i (1 - \lambda_i)^{-1} \lambda_i^N \boldsymbol{\pi_i}}{\sum_i c_i \lambda_i^N \boldsymbol{\pi_i}} \tag{5.39}$$

$$= \frac{c_1 (1 - \lambda_1)^{-1} \lambda_1^N \boldsymbol{\pi_1} + \sum_{i \neq 1} c_i (1 - \lambda_i)^{-1} \lambda_i^N \boldsymbol{\pi_i}}{c_1 \lambda_1^N \boldsymbol{\pi_1} + \sum_{i \neq 1} c_i \lambda_i^N \boldsymbol{\pi_i}} \tag{5.40}$$

π_1 is scaled such that $\pi_1 \mathbf{1} = 1$. Here since the number of iterations N is large and $\lambda_1 > |\lambda_2| \geq \ldots \geq > |\lambda_m|$, we have that $\lambda_i^N \ll \lambda_1^n$ for $i \neq 1$. The eigenvalue λ_1 is called the mean survival probability (per time step) for F. Equation (5.40) can be simplified as:

$$\mathrm{MET}_F := \frac{1}{1-\lambda_1} = \mathrm{MET}\,(\mathbf{s}_N) + O((\lambda_{\mathrm{dec}}/\lambda_1)^N), \tag{5.41}$$

where λ_{dec} is the modulus of the second-largest eigenvalue of A_F. The mean survival time MET_F in the Markov chain's stepping class is expected to be approximately equal to the time from t_{start} (the start of the transition interval) to the transition. So, it should be on the same order but slightly shorter than the length of the a-priori chosen transition interval, $t_{\mathrm{end}} - t_{\mathrm{start}}$. The modulus of the second eigenvalue, λ_{dec}, is associated to the *mixing time*

$$\mathrm{MIX}_F := \frac{1}{1-\lambda_{\mathrm{dec}}} \tag{5.42}$$

within the stepping class. This mixing time is related to the time it takes to "forget the initial condition" while staying in the stepping class. We expect this time to be the time it takes to perform several steps (large scale oscillations in the data time series), so, on the order of several seconds.

The magenta crosses in Figure 5.10(a) highlight the boxes with center $X_{\mathrm{emp},i}$ for which MET_i are less than the overall mean $\mathrm{MET}_F = \frac{1}{1-\lambda_1}$. We observe that these magenta crosses are not uniformly spread around the unit circle but are mostly concentrated in a range of phases. We note that the escape time MET_F and its distribution $\mathbf{s}_F$ is guaranteed to exist due to the ergodicity of the stepping class F_{step}:

$$s_{F,i} = \lim_{N\to\infty} P(X_N = F_i \,|\, X_0 = k \text{ and } X_N \in F) \text{ for all } k \in F \tag{5.43}$$

(where we use the notation F_i to indicate the ith element of transition set F in the Markov chain state space $\mathscr{I}_{\text{emp}}$). In particular, the stationary distribution is reachable from all points in the transition set. Appendix C gives a simple example in which the Markov chain has a dominant eigenvalue so that the largest eigenvalue λ_1 for the transition matrix A_F of the Transition set F is the escape rate.

5.2.6 Time series analysis with transfer operators

The construction of the Markov chain after embedding with Hilbert Transform is a special case of deriving approximate transfer operators from a scalar time series, especially adapted by including knowledge about the underlying system.

- The Hilbert Transform reconstructs the approximate phase space (following delay embedding and false-nearest neighbour analysis that suggest dimension 2 is suitable for embedding time series dominated by regular oscillations).

- The subdivision into boxes chooses a basis of n_{emp} piecewise constant indicator functions on which a projection of the transfer operator is approximated.

- The empirical transition matrix A then encodes the approximate Kolmogorov forward operator (or Fokker-Planck or Perron-Frobenius operator) as $\mathbf{s} \mapsto \mathbf{s}A$ for discretized probability distributions $\mathbf{s} \in \mathbb{R}^{1\times n_{\text{emp}}}$, and the Kolmogorov backward operator (or Koopman operator) as $\mathbf{r}^{\mathsf{T}} \mapsto A\mathbf{r}^{\mathsf{T}}$for linear observables of the form $\mathbf{s} \mapsto \mathbf{s}\mathbf{r}^{\mathsf{T}}$ with $\mathbf{r}^{\mathsf{T}} \in \mathbb{R}^{n_{\text{emp}}\times 1}$ for probability distributions $\mathbf{s}$.

Hence, our analysis is a special case of the general operator-theoretic approaches for the data-driven analysis of dynamical systems, based on the Perron–Frobenius operator or its adjoint, i.e., the Koopman operator (composition operator) [30]. Mezic and his collaborators [31, 32] used Koopman operators to reduce the order of stochastic or deterministic models.

In Section 5.2.6 we show that in our case the eigenvector $\mathbf{r}^{\mathsf{T}}$ for the dominant eigenvalue λ_1 of the Koopman operator, restricted to the transition set (called A_F above) also encodes the escape times from F. This implies that other approaches to time series analysis that approximate the dominant mode of the Koopman operator should also be able to determine in which parts of the reconstructed phase space the system are most susceptible to freezing (escape from stepping) [43].

Approximation of escape preferences by dominant eigenvector of Koopman operator

The key quantity for determining states in the transition set that are most susceptible to transition is the vector of quantities $\mathrm{MET}(\mathbf{u}_j)$,

$$\mathsf{v}_{\mathrm{MET}} = \Big(\mathrm{MET}(\mathbf{u}_1),\ldots,\mathrm{MET}(\mathbf{u}_m)\Big)^{\mathsf{T}} = [I - A_F]^{-1}\,\mathbf{e}^{\mathsf{T}} \tag{5.44}$$

(see (5.28)), the expected escape times from transition set F, when starting in the jth state in F. Recall that $u_{j,j} = 1$, $u_{j,\ell} = 0$ for $\ell \neq j$, $\mathbf{u}_j \in \mathbb{R}^{1\times m}$, where m is the size of the transition set F. More precisely, the relative sizes of the entries in $\mathbf{v}_{\mathrm{MET}}$ matter: the larger $\mathbf{v}_{\mathrm{MET},j}$, the more likely one is to transition out of F from the jth state of F, compared to others. Thus, we may consider the vector $\mathbf{v}_{\mathrm{MET}}$ in the scaling $\mathbf{v}_{\mathrm{MET}}/\mathrm{MET}_F$, considering the escape times relative to the expected escape time MET_F (defined in (5.41)) starting from the stationary density $\mathbf{s}_F$, given in (5.43).

If the transition set F is long-lived (that is, $\mathrm{MET}_F \gg 1$) and the rate of mixing within the transition set F, $1-\lambda_{\mathrm{dec}}$, is large compared to the escape rate from F, $1-\lambda_1 = 1/\mathrm{MET}_F$, then the vector of escape rates $\mathbf{v}_{\mathrm{MET}}$ is approximately equal (up to scaling) to the right eigenvector $\mathbf{r}_F^\mathsf{T}$ of A_F corresponding to λ_1. Our precise assumptions are as follows.

- **(Transition set F is long-lived)** We assume that

  $$\lambda_1 \approx 1, \text{ more precisely, the mean escape time } \mathrm{MET}_F = \frac{1}{1-\lambda_1} \gg 1.$$

  We consider the spectral projection $\Pi_{\mathrm{qst}} = \mathbf{r}_F^\mathsf{T} \mathbf{s}_F \in \mathbb{R}^{m\times m}$ for A_F onto the right eigenspace of A_F corresponding to λ_1.

- **(Mixing is fast compared to escape)** Denoting by $\Pi_{\mathrm{dec}} \in \mathbb{R}^{m\times m}$ the spectral projection for A_F onto the complementary subspace characterised by a decay faster than λ_1, we assume that

  $$\left\| (I - A_F \Pi_{\mathrm{dec}})^{-1} \right\| \leq \frac{c_{\mathrm{dec}}}{1-\lambda_{\mathrm{dec}}}, \text{ where } c_{\mathrm{dec}} = O(1), \text{ and} \tag{5.45}$$

  $$\mathrm{MIX}_F = \frac{1}{1-\lambda_{\mathrm{dec}}} \ll \frac{1}{1-\lambda_1} = \mathrm{MET}_F\,. \tag{5.46}$$

As Π_{qst} and Π_{dec} are complementary spectral projections they satisfy $A_F\Pi_{\mathrm{qst}} = \Pi_{\mathrm{qst}}A_F$, $A_F\Pi_{\mathrm{dec}} = \Pi_{\mathrm{dec}}A_F$, $\Pi_{\mathrm{dec}}\Pi_{\mathrm{qst}} = \Pi_{\mathrm{qst}}\Pi_{\mathrm{dec}} = 0$. Splitting the vector $\mathbf{v}_{\mathrm{MET}}$ using the projections Π_{qst} and Π_{dec}, we observe that

$$\begin{aligned}
\mathbf{v}_{\mathrm{MET}} &= \frac{1}{1-\lambda_1}\Pi_{\mathrm{qst}}\mathbf{e}^\mathsf{T} + (I - A_F\Pi_{\mathrm{dec}})^{-1}\,\Pi_{\mathrm{dec}}\mathbf{e}^\mathsf{T}, \text{ such that} \\
\frac{\mathbf{v}_{\mathrm{MET}}}{\mathrm{MET}_F} &= \mathbf{r}_F + O\left(\frac{1-\lambda_1}{1-\lambda_{\mathrm{dec}}}\right)\Pi_{\mathrm{dec}}\mathbf{e}^\mathsf{T} = \mathbf{r}_F + O\left(\frac{\mathrm{MIX}_F}{\mathrm{MET}_F}\right)\Pi_{\mathrm{dec}}\mathbf{e}^\mathsf{T} \qquad (5.47)
\end{aligned}$$

(noting that $\Pi_{\mathrm{qst}}\mathbf{e}^\mathsf{T} = \mathbf{r}_F(\mathbf{s}_F\mathbf{e}^\mathsf{T}) = \mathbf{r}_F$), where the term $O(\mathrm{MIX}_F\,/\,\mathrm{MET}_F)$ is small by assumption (5.46). Expression (5.47) indicates the approximate relation between right eigenvector $\mathbf{r}^\mathsf{T}$ and the mean escape times.

5.2.7 Method summary

In summary, for each part of the empirically measured force time series where experimenters (domain experts) have flagged a freezing event we proceed in the following way:

1. Select a transition interval $[t_\mathrm{start}, t_\mathrm{end}]$ such that the data indicates stepping at t_start but freezing at t_end and such that several steps are included.

2. Embed the scalar time profile $x(t)$ into the unit circle in the complex plane using the Hilbert Transform and scaling (5.1) to obtain $X(t) \in \mathbb{C}$ for $t \in [t_\mathrm{start}, t_\mathrm{end}]$.

3. Subdivide the unit circle into $P \times Q$ boxes along polar coordinates (P radial annuli of equal thickness, and Q cones, see Figure 5.6 for illustration).

4. Generate an empirical discrete-time discrete-space Markov chain transition matrix A using (5.9). This transition matrix is typically supported only on a subset of the PQ boxes. Define the centers of these boxes $X_{\mathrm{emp},i}$, with angle coordinate $\psi_{\mathrm{emp},i} = \psi(X_{\mathrm{emp},i})$ for some $i \in \{1, \ldots, PQ\}$.

5. Identify the communicating classes for A and test if one can split them into a transition set F (including a stepping class, containing most of the states) and an absorbing set E.

6. Determine mean escape times MET_i from F into E when starting from any state i in F. Define X_min as the coordinates of the state $i_\mathrm{min} \in F$ with minimal MET_i (the angle coordinate is called $\psi_\mathrm{min} = \psi(X_\mathrm{min})$).

Appendix A investigates how robust (as far as we tested for the data available) our results are with respect to the choice of $[t_{\text{start}}, t_{\text{end}}]$ and discretization parameters P and Q.

5.3 Illustration of the methodology

To illustrate the methodology of using MET_i estimates to find preferred transition phases, we choose time series of two different patients. Figure 5.11 presents a graphical summary of the proposed methodology.

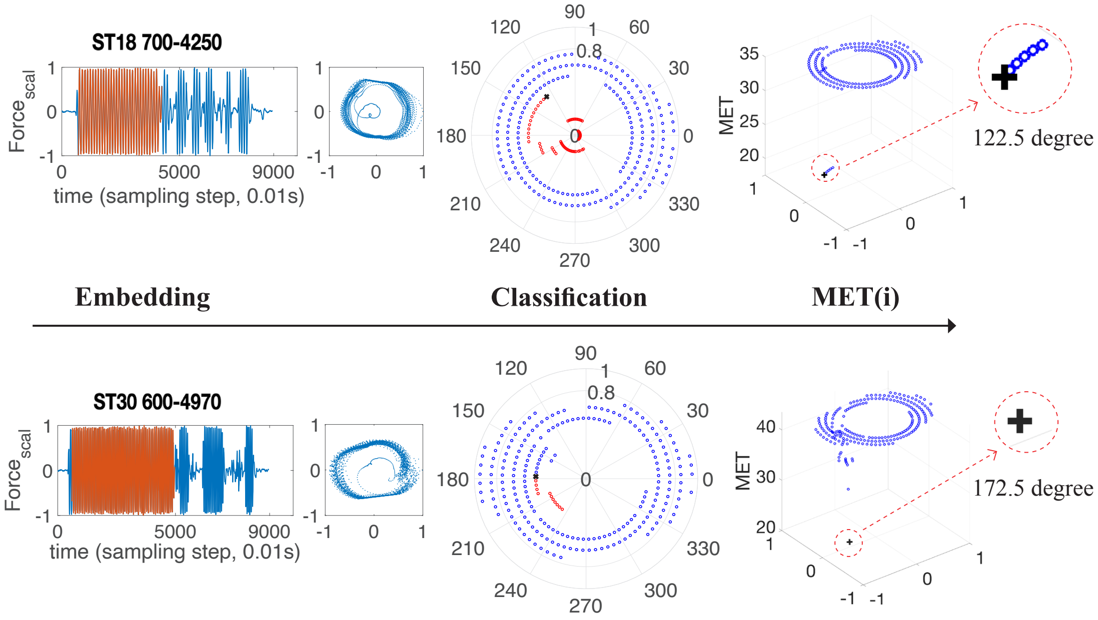


Figure 5.11: Illustration of the methodology for two subjects, ST18 and ST30. Step 1: choose the transition interval $[t_{\text{start}}, t_{\text{end}}]$ for embedding. Step 2: discretize complex plane into boxes along polar coordinate axes, obtain empirical transition probabilities and find communicating classes of resulting Markov chain. Step 3: determine MET_i, the mean escape time, from transition set F for each box.

The top panel starts with the time series from ST18, data set 1. In this case we use $[700, 4250]$ as the transition interval, which amounts to a total of 35.50 s where the sampling time step is 0.01 s between sampling points. After Hilbert embedding and discretization, all the points can be classified into transition set F or absorbing set E. There are a total 3455 data points in boxes contained in the transition set F, which corresponds to a time interval of 34.55 s. Therefore the empirical transition time in the data for this transition interval is 34.55 s. The mean escape time MET_F for the resulting Markov chain is 34.4139 s (obtained via $\mathrm{MET}_F = 1/(1-\lambda_1)$ following equation (5.41)), so is roughly the same, as expected. The mixing time is 7.7675 s (obtained via $\mathrm{MIX}_F = 1/(1-\lambda_{\mathrm{dec}})$, following equation (5.42)). The preferred phase given from ψ_{min} is 122.5°. The dominant eigenvalues for A_F are $\lambda_1 = 0.9997$ and (in modulus) $\lambda_{\mathrm{dec}} = 0.9987$ after rounding to four decimal places.

The bottom panel depicts time series from ST 30, data set 1. In this case we use $[600, 4970]$ as the transition interval, which amounts to a total of 43.70s with time step of 0.01 s between sampling points. There are a total 4352 points in transition set F, which correspond to 43.52 s, therefore for this transition interval the empirical transition time in the data is $\mathrm{MET}_F = 43.52$ s. The MET of the Markov chain is 43.1010 s (again, as expected, very close). The mixing time MIX_F is 5.8854 s. The preferred phase given from ψ_{min} is 172.5°. $\lambda_1 = 0.9998$ and $\lambda_{\mathrm{dec}} = 0.9983$ after rounding to four decimal places.

5.4 Transition phases for freezing events from stepping data

Figure 5.12 shows the most likely transition phases ψ_{min} derived based on eight events in our data set (subject 31, data sets 1 and 2), shown in Appendix B. It is clear that the

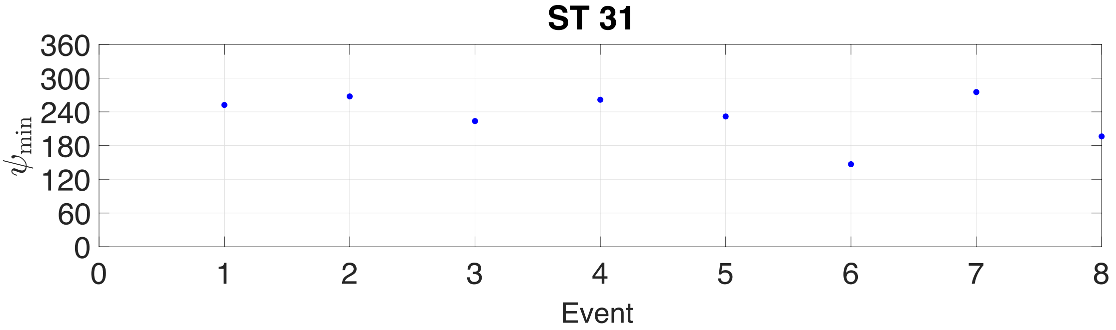


Figure 5.12: Panel (a) shows the state angle $\psi_{\min}$ for preferred transition for all freezing events of subject ST31 (in degree).

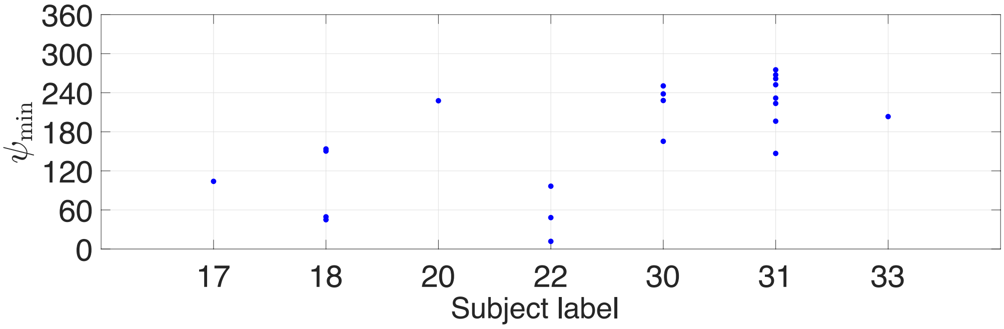


Figure 5.13: Phase $\psi_{\min}$ (preferred escape phase, in degree) of polar box i with minimal expected escape time MET_i of all subjects showing distinct stepping and freezing episodes in the data set. Number of events: 1 (ST17), 4 (ST18), 1 (ST20), 3 (ST22), 4 (ST30), 8 (ST31), 1 (ST33)

values of the preferred transition phases, $\psi_{\min}$ for the different events are clustered in the range 200°–280°. Figure 5.13 shows the preferred transition phases $\psi_{\min}$ for all events present in the experimental data. The x-axis denotes the patient number. We observe that the values for $\psi_{\min}$ corresponding to different subjects appear to be different. These observations suggest that although there might be a preferred transition phase for an individual subject it is not clear whether this holds true between subjects. Individual characteristics influencing the phase angle may be height, weight and the length of legs, etc. The limited number of repeating freezing episodes per subject in the data set does not allow us to be conclusive and reject the null hypothesis of transitions to freezing

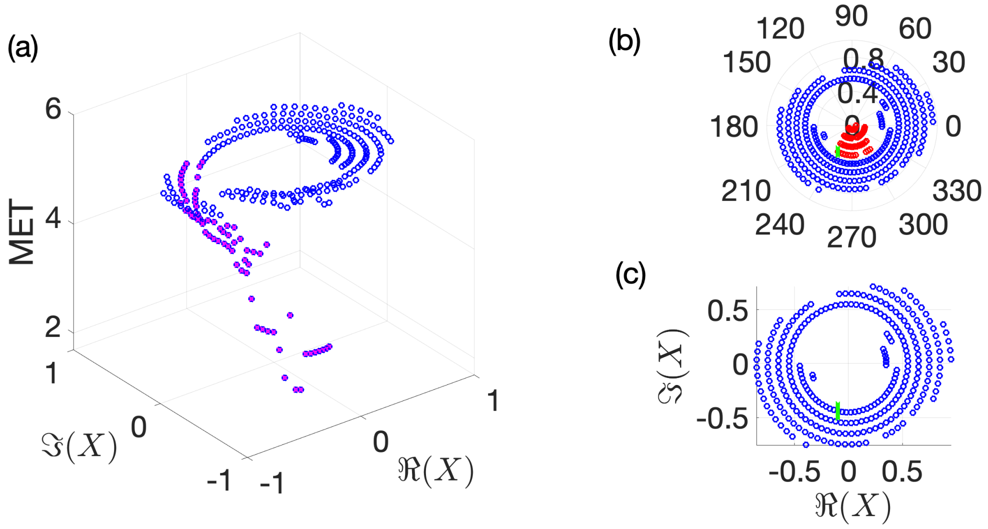


Figure 5.14: Communicating classes and mean escape times from transition set when combining all freezing episodes for subject ST31. Panel (a) shows the mean escape time calculated from each starting state on transition set F, the states marked in magenta crosses have a MET value below the overall expected MET given in (5.41). The preferred angle with minimal MET is marked as green cross in Panel (c). Panel (b) shows the state space of the trajectory after discretization and subdivision with $(p,q) = (0.1, 5^\circ)$. The red dots represents the states in absorbing set E, the green cross represents the ψ_{tr} which is the first states that the Markov chain touches the absorbing set E. The blue dots represents the states in transition set F.

occurring at uniformly distributed phase angles. Nevertheless, Figure 5.13 indicates a non-uniform distribution of phases and hence the possibility of phase dependence of the dynamics underlying the transitions from stepping into freezing we have hypothesized in this paper. Figure 5.14 illustrates the proposed methodology applied to data combining all events for patient #31. We observe that the preferred phase indicated by Figure 5.14 is consistent with the phase angles calculated independently and presented in Figure 5.13.

5.5 Conclusion

In this chapter we propose a methodology for timing the transition into freezing (of gait) and locating this transition in a reconstructed phase space. The methodology combines nonlinear time series analysis and mathematical modelling. We apply the developed methodology to real patient data collected as part of stepping in place experiments. Our approach is patient-specific and thus capable of studying properties of the FOG phenomenon on an individual level basis and potentially applicable on an individual level in the context of personalized interventions in the future [44]. Specifically, our analytical estimate $\frac{1}{1-\lambda_1}$ (equation 5.41) for the time it takes a patient to transition into freezing given a segment of stepping time series data immediately preceding a FE could be used in future algorithms for online FE prediction that could be transformtive for the quality of life of patients with Parkinson's disease as it would enable early warning signals to be calibrated (for each individual) and implemented in wearable devices and/or pressure sensing shoe insoles. Although our data set does not contain information on the step length, the use of insoles in [12, 13] allows the force generated during forward walking to be measured. In this regard the temporal characteristics evaluated here should still apply to forward walking. While measurements from wearable devices may be noisier and hence less accurate than the stepping data from Nantel *et al.*, in cases when the ratio between signal and noise is sufficiently large, phase space reconstruction based on embedding methods would be appropriate. The decomposability of the resulting state space of the Markov chain into transition set and absorbing set will be a good a-posteriori criterion whether the chosen embedding dimension is appropriate. This an interesting prospect for future work.

6. Test of Markov Chain Embedding Using Synthetic Data Produced by Normal Form

In Chapter 3, we studied a simple dynamical model for transition from periodic oscillation to steady state as seen in the Parkinson's Disease data of [7]. This was motivated by phase space reconstruction obtained by delay-embedding and Hilbert transform from our dataset. We built the stochastic model (3.11), which has a deterministic part equal to the generalised Hopf bifurcation normal form with its parameters set in its bistable regime (as described in Chapter 3). For this model we studied the probability of escape from the vicinity of the stable periodic orbit and the mean escape time in Chapter 4 [39, 37, 38].

In Chapter 5, we developed a Markov chain model to study freezing of gait in individuals suffering from Parkinson's disease. The Markov chain model had distinct absorbing sets in its state space, which correspond to freezing. This allowed us to determine a time and phase for the onset of freezing that is largely independent of thresholds and method parameters.

The Markov chain models we have constructed depend on each patient's experimental data. We claim to identify a preferred phase for freezing transition in this data using our model. This makes it necessary to test our method on a “null model” where we know that a preferred phase should not exist. The stochastically perturbed generalised Hopf bifurcation normal form (3.11) in Chapter 3 is such a null model. By testing our approach to determine a preferred phase for transition to freezing on the null model, we can evaluate the robustness of our Markov chain-based method and further understand its potential applicability in broader contexts. In this chapter, we first provide a detailed description how we generate synthetic data, how we choose the transition interval, which embedding methods we apply and where we introduce and control artificial bias.

While the trajectories of (3.11) have two components, the synthetic data sets for our Markov chain-based method is only a one-dimensional (scalar) time series. This implies that we have to choose a projection from $\mathbb{R}^2$ into $\mathbb{R}$. This projection will break the rotational symmetry of (3.11). We will explore the effect of varying this projection on the preferred phase of escape and the effect of different embedding methods for the data. Moreover, we study the effect of the projection angle on the preferred phase of escape to find out how to improve the accuracy and reliability of phase estimation.

6.1 Generation of synthetic data $y(t,\Omega) \in \mathbb{R}^2$

We consider again the stochastic differential equation to generate trajectories $y(t) \in \mathbb{R}^2$ that will serve as “synthetic data”:

$$\mathrm{d}y_1(t) = \left[\beta y_1 - 2\pi\Omega y_2 + (1-\beta)y_1(y_1^2+y_2^2) - y_1(y_1^2+y_2^2)^2\right]\mathrm{d}t + \sigma \mathrm{d}W_1(t), \tag{6.1a}$$

$$\mathrm{d}y_2(t) = \left[2\pi\Omega y_1 + \beta y_2 + (1-\beta)y_2(y_1^2+y_2^2) - y_2(y_1^2+y_2^2)^2\right]\mathrm{d}t + \sigma \mathrm{d}W_2(t), \tag{6.1b}$$

where we initially set the frequency Ω to a constant but later generalise it to a y-dependent function $\Omega = \omega(y)$, given in (6.3) below. System (6.1) has rotational symmetry if Ω is constant or depends only on the amplitude $R = \sqrt{y_1^2 + y_2^2}$. This means that for each trajectory $(y_1(t), y_2(t))$ and each rotation matrix $M(\theta_{\text{rot}})$ by angle θ_{rot}, the trajectory $M(\theta_{\text{rot}})(y_1(t), y_2(t))^{\mathsf{T}}$ is also a trajectory for the correspondingly rotated initial condition $M(\theta_{\text{rot}})(y_1(0), y_2(0))^{\mathsf{T}}$ and realisation $M(\theta_{\text{rot}})(W_1(t), W_2(t))^{\mathsf{T}}$.

The transformation of this stochastic differential equation into polar coordinates given by $z(t) = R(t)\exp[i\theta(t)]$ is:

$$\mathrm{d}R = \left[\beta R + (1-\beta)R^3 - R^5 + \frac{\sigma^2}{2R}\right]\mathrm{d}t + \sigma \mathrm{d}W_R, \tag{6.2a}$$

$$\mathrm{d}\theta = 2\pi\Omega \mathrm{d}t + \frac{\sigma}{R}\mathrm{d}W_\theta. \tag{6.2b}$$

where $\beta = -0.8$ determines attraction toward the origin, and $\sigma = 0.05$ is the noise amplitude. For each parameter combination we generate $N = 10,000$ realizations. The sampling time step $\delta t = 0.01$. The step size h for the Euler-Maruyama method is 0.00025. System (6.1) and system (6.2) identical to (3.11) and (4.3) in chapter 3, repeated here for easier recall. It is wroth to mention here that for numerical solution for SDEs there are different notions of convergence ”strong convergence”, ”weak convergence” (that is, the mean converges), ”almost sure convergence” etc depending on the type of σ(noise), in the system and the properties of the SDEs. Convergence refers to how well the numerical solution approximates the true solution as the h(stepsize), becomes smaller. For additive noise (with constant σ) the Euler-Maruyama has strong convergence with order h(stepsize). With factor σ/R (and assuming $R > R_{\min} > 0$), Euler-Maruyama has strong convergence of order $h^{1/2}$ (so, much worse, in other words we need 4 timers as many steps to halve the error) and weak convergence of order h [45].

Selection of β and σ In Chapter 3, Figure 4.2 illustrates the mean first escape time simulated for various values of σ between 0.01 and 0.1, and different values of β. From Figure 4.2, we observe how we have to choose β and σ in order to ensure that the mean escape time (which is the solution of the mean first escape time equation (4.7)) falls within the range of 20 to 30 seconds (equalling 2000 to 3000 time steps of $\Delta t = 0.01$s). A suitable choice of parameters are $\beta = -0.8$ and $\sigma = 0.05$. This selection of parameters achieves a balance needed between the exponential distribution of the first escape time and maintaining a mean first escape time within a desired range.

Selection of sample size N and step size h In the last section of Chapter 3, we have thoroughly examined the effect of sample size N and step size h on the performance of Euler's method. As illustrated in Figure 4.4 and Figure 4.5, we can conclude that a sample size of $N = 10,000$ is sufficiently large to distinguish differences between the resulting distributions and a Euler-Maruyama step size of $h = 2.5 \times 10^{-4}$ is sufficiently small to avoid numerical instability. These are reasonable choices to generate data similar to the experiments and maintain a high level of accuracy in our numerical approximations, while ensuring that the computational demands remain manageable.

Selection of δt We sample every $(\delta t/h)$th point of the Euler-Maruyama method's result to get $\delta t = 0.01$ which is equal to the sampling steps in the experimental data sets shown in Figure 3.1.

Introduction of state dependence of rotation frequency In chapter 3 we kept the rotation frequency Ω fixed (approximating the typical stepping frequency in experiments). However, when comparing time profiles of experimental data in Figure 3.1 (also Figure 6.1 below) and simulations of (3.11) (same as (6.1)) in Figure 3.5, we observe that the time profile of (6.1) shows visibly oscillatory decay all the way to the close vicinity of the origin, while the experimental time profile does not. For this reason we introduce Ω as a function of the state y, permitting us to manipulate rotation frequency in different parts of the phase space:

$$\Omega = \omega(y) = \frac{\Omega_{\text{fix}} \cdot (1 + \exp(-\alpha * (1 - R_{\text{drop}})))}{1 + \exp(-\alpha * (R - R_{\text{drop}}))} + \sum_i \tanh\left(y_{i\text{par}} \cdot y_i\right) \cdot y_{i\text{bias}} \tag{6.3}$$

where Ω_{fix} is the rotation frequency near the deterministic limit cycle of radius $R = 1$, and α is the logistic growth rate controlling the sharpness of the change of rotation frequency when the oscillation amplitude drops at radius $R = R_{\text{drop}}$. In the additional "bias term", $y_{i\text{par}}$ is the growth rate of the hyperbolic tanh function controlling the sharpness of switching between $y_{i,\text{bias}}$ and its negative. A non-zero bias level $y_{i\text{bias}}$ makes the rotation frequency different in the half-planes $\{y_i > 0\}$ and $\{y_i < 0\}$. Further, in (6.3) we used $R = \sqrt{y_1^2 + y_2^2}$. The following subsections will discuss in detail the three combinations of parameters in (6.3) that we will study to see how they affect the phase of transition as determined by the Markov chain.

6.1.1 Case of constant rotation frequency Ω

If the rotation frequency is constant then the coefficient $\omega(y)$ is a constant parameter, which we choose to be equal to the average frequency of stepping in our experimental datasets as obtained by inverse Fourier Transform (see Section 3), let $\alpha = 0$, $y_{i\text{bias}} = 0$:

$$\omega(y) = \Omega_{\text{fix}} = 0.8667, \tag{6.4}$$

where the unit is s^{-1}. Figure 6.2 Panel (a) presents an example time profile of system (6.1) with $\omega(y) = 0.8667$. The generated sample time series y_1 for (6.1) fixed $\Omega_{\text{fix}} = 0.8667$ looks qualitatively similar to experimental data (example repeated in Figure 6.1) until the transition. The time series y_1 in Figure 6.2 Panel(a) shows oscillations until $t \approx 5,500$, when the amplitude of the oscillations quickly decreases until $t \approx 6,000$. After $t \approx 6,000$ the time series fluctuates in an approximately Gaussian stationary distribution around $y_1 = 0$, which is consistent with what is observed in the experimental data. Also, the power spectrum distribution has one single peak, but overall, the general pattern remains consistent. A noticeable qualitative difference between the time profiles in Figure 6.2 Panel(a) and the experimental data in Figure 6.1 are the gradually decaying oscillations between $t = 5500$ and $t = 6000$ in Figure 6.2 Panel(a). These are not visible in the experimental data in Figure 6.1.

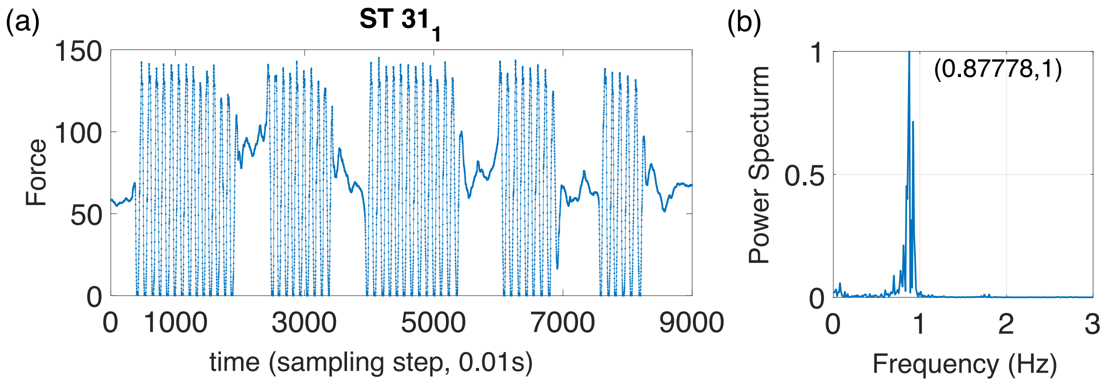


Figure 6.1: Panel (a): Time series of the left foot vertical force (as a percentage of body weight) of the experimental subject numbered ST 31, data set 1, sampling time step $\delta t = 0.01$; Panel (b): Power spectrum (scaled to a maximum equal to unit) with a strong dominant peak at $f_{\text{stp}} \approx 0.88$Hz.

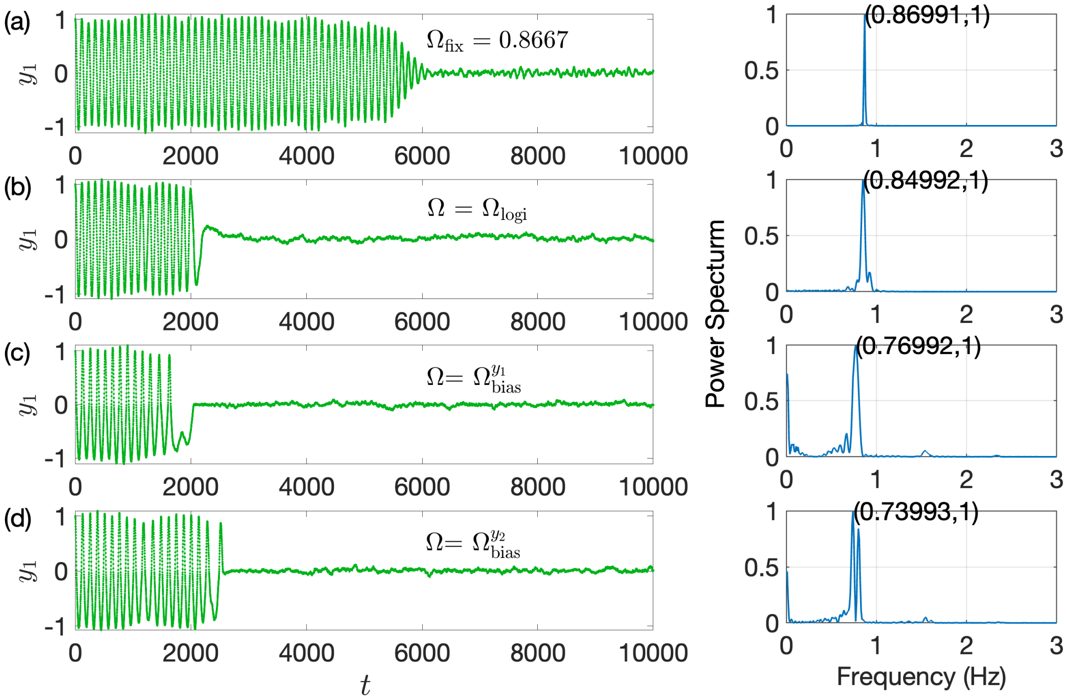


Figure 6.2: Panel(a) presents the generated measured time series y_1 for fixed $\Omega_{\text{fix}} = 0.8667$. Panel(b) presents the measured time series generated by logistic Ω function in (6.5). Panel(c) and Panel(d) presents the measured time series generated by Ω function with $y_{1\text{bias}}$ and $y_{2\text{bias}}$ respectively. Left side of each panel shows the corresponding power spectrum distribution and the black text shows the coordinate of the dominant frequency.

6.1.2 Case of purely amplitude-dependent rotation frequency

In (6.2), the change of radius R depends solely on R (not on phase θ). This reduction is valid as long as $\omega(y)$ only depends on R (not on y_1, y_2 individually). When assuming that ω depends on R and is no longer a constant, we expect the following effects:

- When R is within the range of 0.8 to 1.2, i.e. the trajectory rotates along the stable limit cycle, ω should be around $\Omega_{\text{fix}} = 0.8667$. This value is the average frequency ω observed in our experimental data during the stepping stage.
- As R begins to decrease below the radius of the unstable limit cycle, the frequency should decrease sharply (controlled by parameter α). This captures the change in

dynamics as the system transitions away from oscillations.

- When R small, the frequency should also be small, as $\omega(y) \ll 1$.

When choosing $y_{i\text{bias}} = 0$, function ω will become the following purely amplitude-dependent autonomous logistic rotation frequency function:

$$\Omega_{\text{logi}} = \omega(y) = \frac{\Omega_{\text{fix}} \cdot (1 + \exp(-\alpha \cdot (1 - R_{\text{drop}})))}{1 + \exp(-\alpha \cdot (R - R_{\text{drop}})}, \tag{6.5}$$

where we set $\alpha = 25$, $R_{\text{drop}} = 0.85$. This logistic function possesses some inherent properties that can help us understand its behavior. The left panel of Figure 6.3 displays Ω_{logi} as a function of R. As R increases, the logistic function converges to its maximum value, which in this case is 1. On the other hand, as R decreases from 1, it reaches the midpoint of the logistic function at R_{drop}, where α determines the steepness of the curve. The value of α influences how quickly the function transitions from one extreme to the other. When α is greater than 10, the steepness of the curve does not change significantly. This is because, at higher values of α, the transition between the minimum and maximum values of the function becomes increasingly rapid, and further increases in α result in only marginal changes in the steepness.

Given our requirement for a high steepness between the two conditions (normal oscillation and escape), we set $\alpha = 25$. This value ensures a rapid transition between the two states, effectively capturing the desired system dynamics. Additionally, we want R_{drop} to be close to the unstable limit cycle, which represents the point at which the dynamical system has already escaped. We set $R_{\text{drop}} = 0.85$, as it is close to the unstable limit cycle, given by $\sqrt{-\beta} = \sqrt{0.8} = 0.8944$. The numerator of this equation ensures that as R approaches 1, the frequency will be near Ω_{fix}, which aligns with our experimental data.

Figure 6.2 Panel (b) presents the Synthetic data $y_1(t, \Omega_{\text{logi}})$ generated by (6.1) with the Ω_{logi} (6.5). We can see that once y_1 does not oscillate, it rapidly decreases to 0, which is a characteristic behavior observed in our experimental data. In addition to the dominant frequency, there are few minor peaks present, which indicates that the time series is more complex than a single oscillatory behavior and contains other, smaller-scale oscillations or fluctuations. This reflects that this time series is more similar to our experimental data.

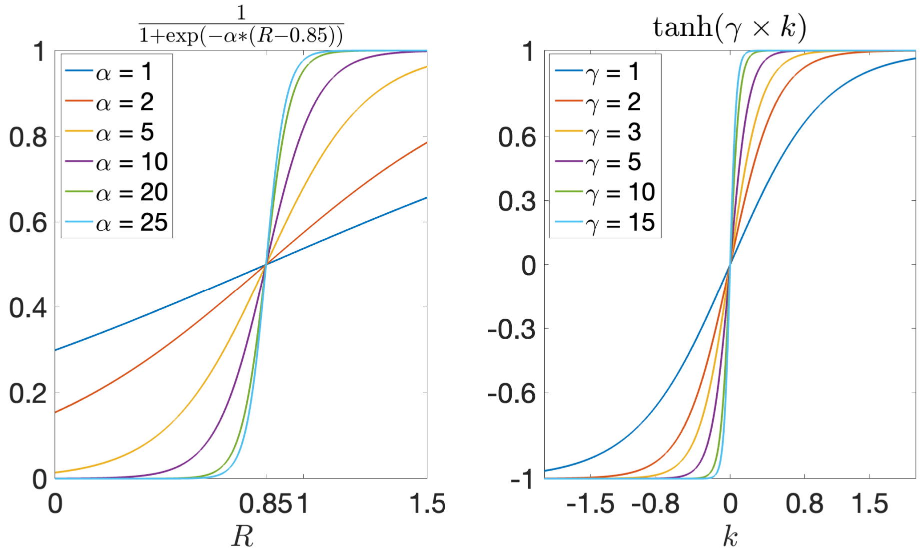


Figure 6.3: The left panel shows the curves for different values of α in the equation $1/(1+\exp(-\alpha \cdot (R-0.85)))$. The right panel shows the curves for different values of γ in the equation $\tanh(\gamma \cdot k)$.

6.1.3 Case of phase-dependent rotation frequency

Returning to our original objective, we aimed to test the validity of our methodology using a different dataset, while still focusing on automatically identifying absorbing sets that can be utilized to infer the stepping phase at which freezing events occur. To achieve this, it is crucial that our dataset exhibits a distinctly preferred phase, which will enable a more intuitive evaluation of the accuracy and practicality of our method.

By adding bias to the components y_1 or y_2, which can break the rotation symmetry of (6.1) to create a system with biases in the transition phase controllable by parameter $y_{i\text{bias}}$ and $y_{i\text{par}}$

$$\Omega_{\text{bias}}^{y_i} = \omega(y) = \frac{\Omega_{\text{fix}} * (1 + \exp(-\alpha \cdot (1 - R_{\text{drop}})))}{1 + \exp(-\alpha \cdot (R - R_{\text{drop}})} + \sum_i \tanh\left(y_{i\text{par}} \cdot y_i\right) \cdot y_{i\text{bias}}. \tag{6.6}$$

We consider the following two parameter combinations for $y_{i\text{bias}}$:

$$\begin{cases} y_{1\text{bias}} = 0.3, & y_{2\text{bias}} = 0 \\ y_{1\text{bias}} = 0, & y_{2\text{bias}} = 0.3, \end{cases}$$

where $y_{1\text{par}} = 20$, $y_{2\text{par}} = 20$. The right panel of Figure 6.3 illustrates the behavior of the function $k \mapsto \tanh(\gamma \times k)$ for different values of slope γ. Similar to the logistic function, the hyperbolic tangent (tanh) function also has an upper bound and a lower bound, in this case, 1 and -1. The slope parameter γ represents the steepness of the curve. One notable property of the tanh function is its oddness, which means that it is symmetric with respect to the origin. This property ensures that, as long as the function is steep enough, the values of tanh are mostly concentrated at the upper bound when the input variable is positive and at the lower bound when the input variable is negative. If $y_{2\text{bias}} = 0, y_{1\text{bias}} = 0.3$, in (6.1), as long as y_1 is greater than 0, the bias $\tanh\left(y_{1\text{par}} \cdot y_1\right) \cdot y_{1\text{bias}}$ is positive and between $(0, 0.3]$, causing the rotation frequency to be larger. Conversely, when y_1 is less than 0, the bias is negative and between $[-0.3, 0)$, resulting in a reduction of rotation frequency $\omega(y)$. This variation in rotation frequency between the half-planes permits us to introduce a controllable bias in the observed phase of transitions.

Panel (c) and Panel (d) in Figure 6.2 display the time series $y_1(t, \Omega_{\text{bias}}^{y_1})$ and $y_1(t, \Omega_{\text{bias}}^{y_2})$ which are generated by (6.1), where the rotation frequency is as given in (6.6). These panels demonstrate the effects of introducing bias into the y_1 and y_2 components of the system. In Panel(c), we set $y_{2\text{bias}} = 0$, which means that we only apply bias to the y_1 component. This panel illustrates the behavior of $y_{1\text{bias}}$ when only one component is biased, and we can observe how this modification impacts the dynamics of the system later after applying our methodology. In Panel(d), we apply bias only to the y_2 component while leaving the y_1 component unbiased. This panel provides a complementary view of the $y_{1\text{bias}}$ behavior when the other component is biased. The introduction of bias to the Ω function leads to a decrease in the dominant frequency and an increase in minor peaks of the power spectrum distribution compared to the Ω_{fix} and Ω_{logi} which are without bias, as shown on the right-hand side of panels (c) and (d) of Figure 6.2.

6.2 Comparative Analysis of Preferred Phases Using Synthetic Data

We have now generated four sets of synthetic data $y(t, \Omega) \in \mathbb{R}^2$ from the previous section using (6.1) that can be used to apply our methodology. Each set of data contains 10,000 simulations, but the number of simulations that can be effectively used is approximately 5,000-6,000 per dataset (depending on the β and σ we choose in (6.1)). This is due to two reasons:

- (No escape) Due to the random nature of the noise and the length of the time series we set (10,000 time steps with $\mathrm{dt} = 0.01$), the noise-induced escapes from oscillatory attractors to steady states does not occur in each realisation.

- (Escape too early) Since later we need to discretize our y_1-y_2 state space and then construct a Markov chain transition matrix, which is a good representation of escape, we also to ignore time series with an insufficient number of points at oscillatory attractors.

To choose our transition interval ($[t_{\text{start}}, t_{\text{end}}]$), we select the first point that hits amplitude $R = 0.5$ ($R = \sqrt{y_1^2 + y_2^2}$) as t_{end}, the 1999th point before t_{end} as t_{start}, resulting in a total length of $[t_{\text{start}}, t_{\text{end}}]$ of 2000 sampling points. These time series are used as models for the experimenter-labelled escape from the oscillatory attractor to the steady state, and any time series with $R \geq 0.5$ for all times or without 1999 points before t_{end} will be excluded. Thus, each dataset will be filtered for the set of time series where escape exists.

Projection to scalar time series The experimental data is a scalar time series, while (6.1) is a two-dimensional signal. To investigate the effect of the embedding method used we first project the two-dimensional $y(t)$ to a scalar time series using the **Projection Angle θ_r**.

Projection Angle θ_r The Hilbert Transform, $H(x(t))$, identifies the position of $x(t)$ in the complex plane by reconstructing an imaginary part consistent with the harmonic (cos and sin) components of periodic motion (see section 2.1.2 for complete description). However, our observation of $H(x(t))$ contains only the information from $x(t)$, or in our case, a projection $x(t) = y_1(t)$, and it cannot recover all the information contained in the y_1-y_2 plane. Given that our viewing angle is predetermined, with y_1 as the x-axis, $H(x(t))$ might exhibit a predictable and systematic observation bias that breaks rotation symmetry. To corroborate this hypothesis, we choose projections onto one-dimensional subspaces different from $\{y \in \mathbb{R}^2 : y_2 = 0, y_1 \text{ arbitrary}\}$. Possible projections are determined by the counterclockwise angle θ_r, $\theta_r \in [0,360)$ degree resulting in a projection $x(t) = y_{\text{new}}[\theta_r]$, using the rotation formula:

$$y_{\text{new}}[\theta_r] = y_1 \cos(\theta_r) + y_2 \sin(\theta_r) \tag{6.7}$$

We choose the scalar time profile $y_1(t)$ as our default measured time series $x(t)$ for $t \in [t_{\text{start}}, t_{\text{end}}]$ and following the steps of our method summarized in Section 5.2.7, we will obtain a set of angular coordinates ψ_{min}, which represents the escape phase for each dataset, denoted as $\Psi^{\text{EM}}_{\text{min}}(\Omega, \theta_r)$, where EM is the embedding method we choose and θ_r represents the degree of projection for the default $x(t) = y_1(t)$.

Embedding Method (EM) We study the effects of two different embedding methods and compare them to the full time series as benchmark:

- (Hilbert embedding) In Chapter 5 the scalar time profile $x(t)$ is embedded into the unit circle in the complex plane using the Hilbert Transform and scaling (5.1), resulting in $X(t) \in \mathbb{C}$ for $t \in [t_{\text{start}}, t_{\text{end}}]$. Subsequently, the complex plane can be subdivided into boxes. The EM superscript for the resulting Ψ is "hilbert" for this case, e.g., $\Psi^{\text{hilbert}}_{\text{min}}(\Omega, \theta_r)$.

- (Benchmark: no projection and no embedding) In (6.1), when y_1 and y_2 are positioned in the y_1-y_2 plane and scaled to $[-1,1]$, that is,

  $$X(t) = (y_1(t), y_2(t)) / \max_{t \in [t_{\text{start}}, t_{\text{end}}]} \left(\sqrt{y_1(t)^2 + y_2(t)^2}\right),$$

  we can also obtain the unit circle and then follow the subsequent steps to get ψ_{min}. Using y_1 and y_2 simultaneously allows us to fully utilize all the information contained in $y(t,\Omega)$. Furthermore, the trajectory in the y_1-y_2 plane is devoid of errors. The EM superscript of Ψ is "(y_1,y_2)", e.g., $\Psi_{\min}^{(y_1,y_2)}(\Omega,\theta_r)$, which is independent of θ_r as no projection is performed in this case. This serves as a benchmark to evaluate the efficacy and precision of the embedding process.

- (Delay embedding) To provide additional perspectives on the embedding process, we also compare a 2-D delay embedding approach where the delay Δt is set to one-fourth of the oscillation period, with details available in chapter 3, $X(t)$ is calculated using Equation (2.1). EM is called as ”delay”.

Figure 6.4 illustrates the application of our methodology to a an example time series $y_1(t)$, $y_2(t)$, adhering to the steps outlined in Section 5.2.7. Panels (a) and (b) show the time series in the y_1-y_2 plane, generated by a realisation of (6.1), and on the respective scalar time profiles, shown in green. The black portion represents our selected time interval $[t_{\text{start}},t_{\text{end}}]$ that encompasses the escape process, with the final point having a radius of 0.5 and a total span of 2000 points. The scalar time profile $y_1(t)$ for $t \in [t_{\text{start}},t_{\text{end}}]$ is embedded in the unit circle using the Hilbert transform ($H(y_1)$) and the 2D delay embedding ($\text{delay}(y_1)$), which is shown as the black trajectory in panels (c) and (e) in the y_1-y_2 plane, respectively. The black trajectory in panel (d) is the original time series, scaled to the unit circle in the y_1-y_2 plane, with $y_1(t)$ on the x axis and $y_2(t)$ for $t \in [t_{\text{start}},t_{\text{end}}]$ on the y axis. For ease of presentation, the subsequent appearances of $y_1(t)$, $y_2(t)$ for $t \in [t_{\text{start}},t_{\text{end}}]$ are abbreviated to y_1, y_2. The plane is divided into boxes along the polar axis to obtain the empirical transition probabilities and to identify the communication classes of the resulting Markov chain. In panels (c), (d), and (e), the transition set F is denoted by blue dots and the absorbing set E by red dots. The transition set F from each box determines $\psi_{\min}$, which is represented as a black cross in panels (c), (d), and (e).

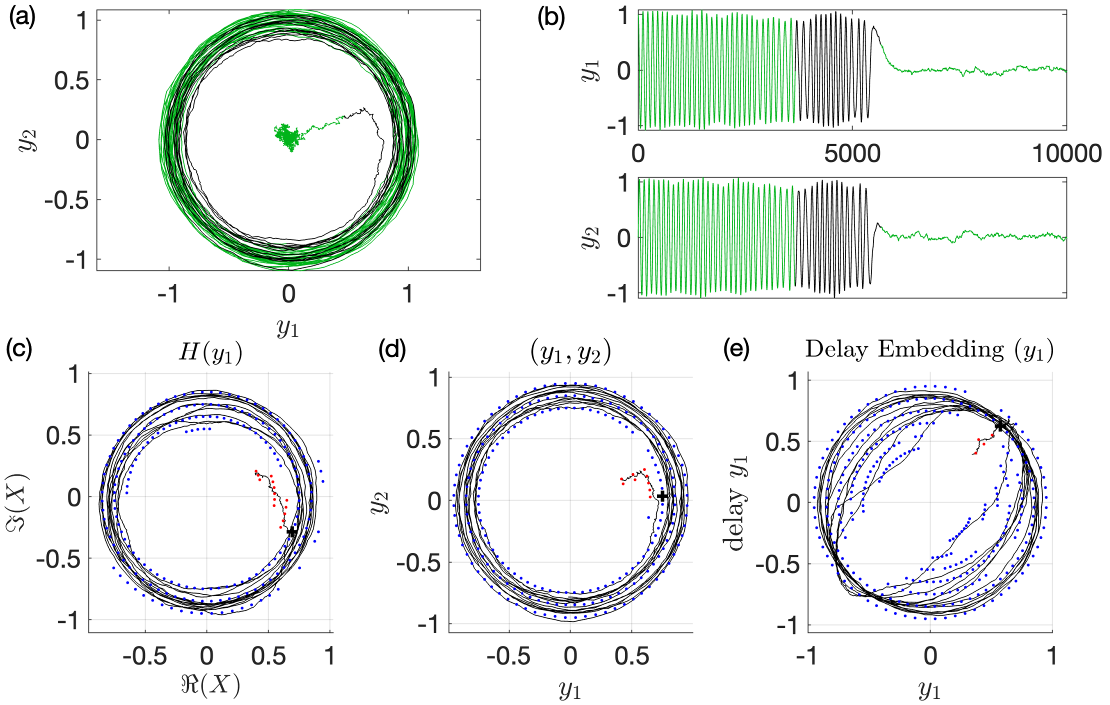


Figure 6.4: This figure generally shows the steps when apply the methodology into one of our measured time series $y_1(t,\Omega_{\text{logi}})$. Panels (a) and (b) show the measured time series in the y_1-y_2 plane and on the respective scalar time profile in green. The black portion represents our selected time interval $[t_{\text{start}}, t_{\text{end}}]$. Panels (c), (d), and (e) present the 2D trajectory when we embed $x(t) = y_1(t,\Omega_{\text{logi}})$ into the unit circle using the Hilbert transform, in the original (y_1,y_2) plane, and using the delay embedding method respectively.

Figure 6.5 presents the probability histogram of $\Psi_{\text{min}}^{\text{hilbert}}(\Omega_{\text{fix}},0^\circ)$, $\Psi_{\text{min}}^{(y_1,y_2)}(\Omega_{\text{fix}},0)$ and $\Psi_{\text{min}}^{\text{delay}}(\Omega_{\text{fix}},0^\circ)$ in Cartesian coordinates (Left) and Polar coordinates (Right) for the complete ensemble of $N = 10\,000$ realisations. For constant frequency Ω_{fix}, the histogram of $\Psi_{\text{min}}^{(y_1,y_2)}(\Omega_{\text{fix}},0^\circ)$ shows approximately uniform distribution which is expected because of the rotational symmetry of (6.1). The projection introduces a breaking of rotational symmetry. So, there will be an effect for both embedding methods ($\text{EM} = \text{hilbert}, \text{delay}$). It shows that only the $\Psi_{\text{min}}^{\text{delay}}(\Omega_{\text{fix}},0^\circ)$, for which $x(t)$ is calculated through delay embedding, manifests two pronounced peaks, shows two maxima (at Ψ around 100° and around 280°) as depicted by the blue bars. This observation indicates that the delay embedding projection destroys the rotational symmetry, an effect that can also be seen in Panel (e) of Figure 6.4, which shows non-circular trajectories in black. We notice that the symmetry breaking by the projection is too small to discern for $N = 10^4$ when using Hilbert transform embedding (EM = hilbert) for constant Ω.

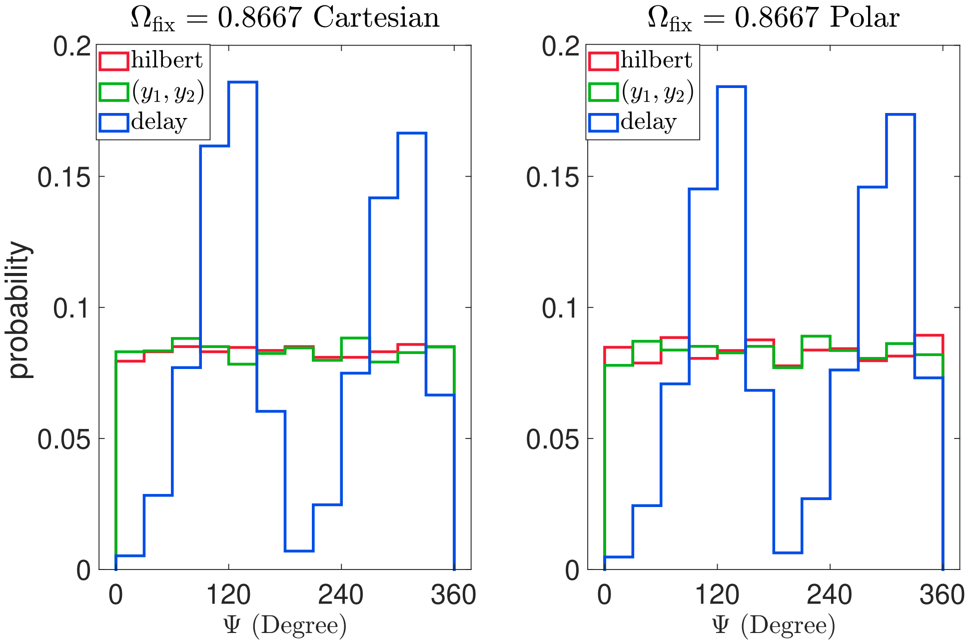


Figure 6.5: Histograms of $\Psi^{\mathrm{EM}}_{\min}(\Omega_{\mathrm{fix}}, 0^{\circ})$ in Cartesian coordinates (Left) and Polar coordinates (Right) for $\Psi^{\mathrm{hilbert}}_{\min}(\Omega_{\mathrm{fix}}, 0^{\circ}), \Psi^{(y_1,y_2)}_{\min}(\Omega_{\mathrm{fix}}, 0^{\circ})$ and $\Psi^{\mathrm{delay}}_{\min}(\Omega_{\mathrm{fix}}, 0^{\circ})$, where $\Omega_{\mathrm{fix}} = 0.8667$, the projection degree θ_r for $x(t) = y_1(t)$ is 0°.

Figure 6.6 displays the histograms for the preferred transition phase for amplitude-dependent rotation frequency (Ω_{logi} given in (6.5)) for all embedding methods and the benchmark, $\Psi^{\mathrm{hilbert}}_{\min}(\Omega_{\mathrm{logi}}, 0^{\circ})$ in red, $\Psi^{(y_1,y_2)}_{\min}(\Omega_{\mathrm{logi}}, 0^{\circ})$ in green and $\Psi^{\mathrm{delay}}_{\min}(\Omega_{\mathrm{logi}}, 0^{\circ})$ in blue, using the same approach as for constant rotation frequency Ω_{fix} in Figure 6.5. We observe two symmetric peaks in the histogram of $\Psi^{\mathrm{hilbert}}_{\min}(\Omega_{\mathrm{logi}}, 0^{\circ})$, where $x(t)$ is calculated by the Hilbert Transform of $y_1(t)$, i.e $H(y_1(t))$ at angles 180° apart. Despite many unknowns regarding the Hilbert Transform, we will explore the effect of projection

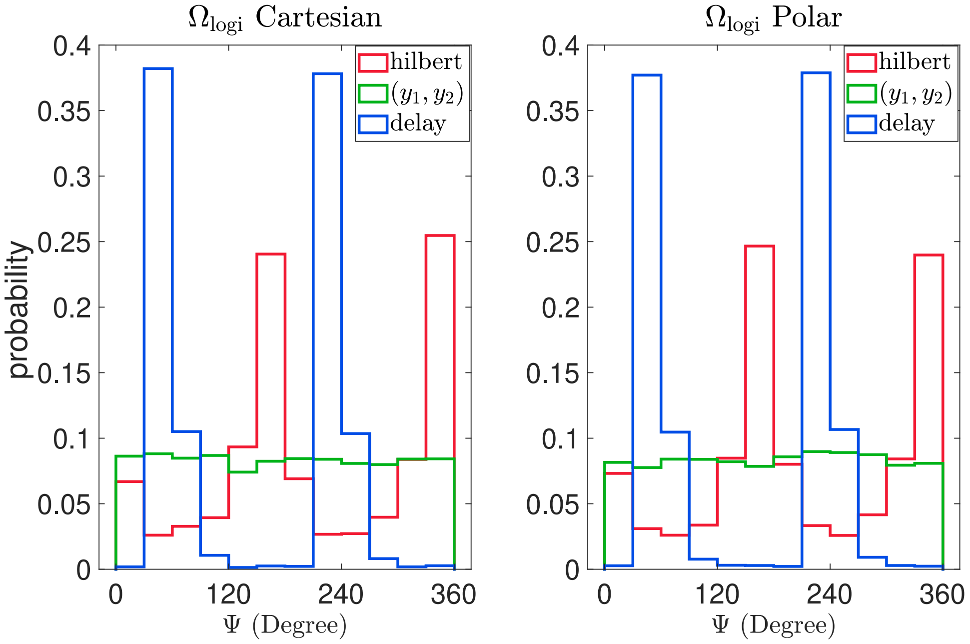


Figure 6.6: Histograms of $\Psi^{\text{EM}}_{\text{min}}(\Omega_{\text{logi}}, 0^{\circ})$ in Cartesian coordinates (Left) and Polar coordinates (Right) for $\Psi^{\text{hilbert}}_{\text{min}}(\Omega_{\text{logi}}, 0^{\circ})$,$\Psi^{(y_1,y_2)}_{\text{min}}(\Omega_{\text{logi}}, 0^{\circ})$, and $\Psi^{\text{delay}}_{\text{min}}(\Omega_{\text{logi}}, 0^{\circ})$. Ω_{logi} (6.5) is a logistic function without bias, the projection degree θ_r for $x(t) = y_1(t)$ is 0°.

in the next section, or more importantly, determine whether this phenomenon follows any discernable pattern or rule.

6.2.1 Effect of Projection Angle θ_r

Our experimental method reconstructs an oscillator from a scalar time series, which is assumed to be a projection of the full phase space variables of the underlying oscillator. In our synthetic data set we artificially perform a similar projection (6.7). The projection will affect the resulting escape angle Ψ_{min}. This section analyses how the effect plays out. Since the underlying oscillator is rotationally symmetric we expect the bias observed in Figure 6.6 to depend on the projection angle θ_r in a rotationally invariant manner: a projection to angle θ_r should shift the distribution of escape angles Ψ_{min} by θ_r. Figure 6.7 Panel(a) shows a heatmap of the two-dimensional frequency distribution of $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{logi}}, 0^{\circ})$ and $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{logi}}, 0^{\circ})$, x-axis is Hilbert embedding after projection with angles $0, 30, 60$ degree, y-axis is the benchmark. The heatmap grid is arranged in a matrix of 12×12 cells, with each cell corresponding to a 30-degree by 30-degree region in the plane. The color in each cell represents the count of data points falling within the corresponding 30-degree region. Darker colors indicate a higher frequency while lighter colors indicate a lower frequency. The x-axis of panel(b) and Panel(c) are the frequency distribution of $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{logi}}, 30^{\circ})$ and $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{logi}}, 60^{\circ})$ respectively. The Hilbert Transform of $y_{\text{new}}[\theta_r]$ which has been rotated counterclockwise by θ_r degrees, causes the original $x - y$ axis to also rotate counterclockwise by θ_r degrees. As a result, the calculated $\Psi_{\min}$ in the new coordinate system should be θ_r degrees smaller than in the original coordinate system. Each panel and horizontal and vertical axis used the same noise realisations, just applied different projection before the embedding. Therefore, in the heatmap, each original escape should consistently shift down by θ_r degrees. As can be seen from the Figure 6.7, the high value boxes have a similar distribution across the x direction and move down 30° and 60° respectively, which aligns with our expectations. The phenomenon shows that the Hilbert transform also breaks symmetry and introduces a bias in the reported transition phase. It shows phases close to 0° and 180° (for $\theta_r = 0$) and correspondingly rotated phases for other θ_r as more likely transition phases. This is a bias artificially introduced by projecting and embedding. However, the bias is still reflection symmetric (with two preferred angles 180 degree apart. When we applied the method to the patient data we did not see clusters 180 degrees apart for any patient. This suggests that our approach may be able to detect biases, however, only as a bias between two entire half circles of phases.

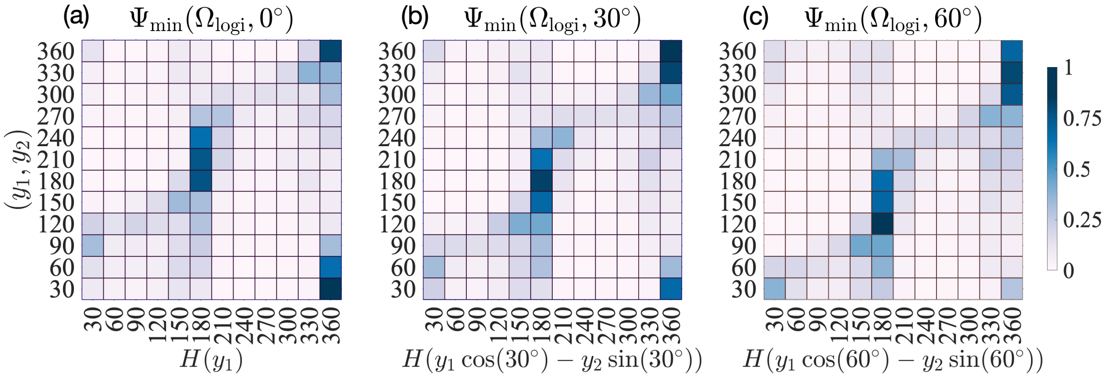


Figure 6.7: Panel(a): Heatmap which represents a two-dimensional frequency distribution of $\Psi^{(y_1,y_2)}_{\text{min}}(\Omega_{\text{logi}}, 0^{\circ})$ and $\Psi^{\text{hilbert}}_{\text{min}}(\Omega_{\text{logi}}, 0^{\circ})$. Panel(b): Heatmap which represents a two-dimensional frequency distribution of $\Psi^{(y_1,y_2)}_{\text{min}}(\Omega_{\text{logi}}, 0^{\circ})$ and $\Psi^{\text{hilbert}}_{\text{min}}(\Omega_{\text{logi}}, 30^{\circ})$. Panel(c): Heatmap which represents a two-dimensional frequency distribution of $\Psi^{(y_1,y_2)}_{\text{min}}(\Omega_{\text{logi}}, 0^{\circ})$ and $\Psi^{\text{hilbert}}_{\text{min}}(\Omega_{\text{logi}}, 60^{\circ})$.

6.2.2 Distribution of threshold crossings after Hilbert Embedding

To further examine the bias in transition phases caused by projection and Hilbert Embedding, we can perform the following analysis: for each trajectory of synthetic data $y(t, \Omega_{\text{fix}})$, $y(t, \Omega_{\text{logi}})$, $y(t, \Omega_{\text{bias}}) \in \mathbb{R}^2$ that makes a transition to the origin, we can identify the point where the trajectory first intersects $R = R_{\text{last}}$ (we will investigate several $R_{\text{last}} \in [0.5, 1]$). We denote this point as $Y^{\text{EM}}_{\text{last}}(\Omega)[R_{\text{last}}]$ for each respective embedding method EM. While phase at these threshold crossings may not be good measures for the phase at which transition occurs (since the phase will strongly depend on R_{last}) the threshold crossings are a simple trajectory feature that shows how the Hilbert transform of a scalar projection distorts a two-dimensional trajectory near its transition. Biases that occur for all threshold crossings will also be present in any algorithm determining a transition phase using a scalar time series and Hilbert embedding.

Figure 6.8 Panel (a) shows scatter plots of $Y^{(y_1,y_2)}_{\text{last}}(\Omega_{\text{logi}})[0.5]$ in green stars and $Y^{\text{hilbert}}_{\text{last}}(\Omega_{\text{logi}})[0.5]$ in red circles. Panel(b) and Panel(c) display the datasets in the polar coordinate system using the same color scheme, where all $R_{\text{last}} = 0.7$ in Panel(b) and

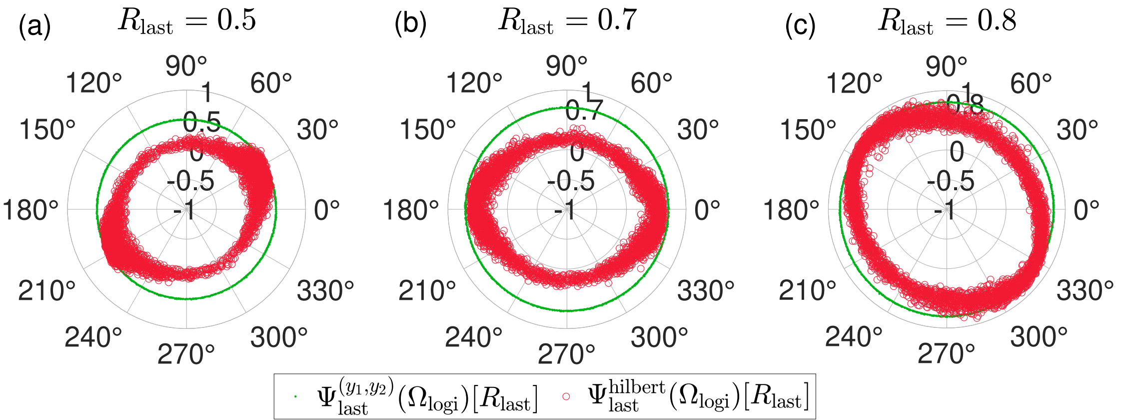


Figure 6.8: Panel(a)(b)(c) shows $Y_{\text{last}}^{(y_1,y_2)}(\Omega_{\text{logi}})[R_{\text{last}}]$ in green stars and $Y_{\text{last}}^{\text{hilbert}}(\Omega_{\text{logi}})[R_{\text{last}}]$ in red circles where R_{last} is 0.5, 0.7 and 0.8 respectively.

all $R_{\text{last}} = 0.8$ in Panel(c). The trajectories are compressed towards the center from two symmetric points. Thus, the Hilbert embedding from a scalar projection reduces rotation symmetry to reflection symmetry of the transition phase. The transition threshold should be close to the radius of the unstable limit cycle for the deterministic dynamical system at around 0.7. As can be seen from Figure 6.8 Panel(b), more points are projected to around 180°and 360°, which is consistent with the distribution of Ψ_{min} shown in Figure 6.6 and Figure 6.7. Appendix D shows the distribution of $Y_{\text{last}}^{\text{EM}}(\Omega)[R_{\text{last}}]$ when we add bias to the stochastic differential equation.

6.2.3 Ψ_{min} after adding bias

After analyzing the bias caused Hilbert embedding of a projection of the dynamical system, we will now demonstrate if Ψ_{min} identifies preferred transition phases by introducing a bias through making the rotation frequency different in different half-planes (see expression for Ω_{bias} in (6.6)). Figure 6.9 Panels (a) and (b) display the distribution of $\Psi_{\text{min}}^{\text{EM}}(\Omega_{\text{bias}}^{y_1}, 0^{\circ})$ and $\Psi_{\text{min}}^{\text{EM}}(\Omega_{\text{bias}}^{y_2}, 0^{\circ})$ with Ω calculated using (6.6) under the y_1-y_2 plane and Hilbert complex plane respectively. In Panel (a), we observe a distinct peak for $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{bias}}^{y_1},0^{\circ})$ and $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{bias}}^{y_2},0^{\circ})$ in comparison to $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{logi}},0^{\circ})$ which has no added bias (shown in green bins). This indicates that (6.6) can produce an expected preferred angle leading to escape. From Panel (b), it is evident that the previously symmetrical two peaks are no longer present due to the introduced bias, where $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{logi}},0^{\circ})$ is shown in red bins. Panel (b) shows that phase preferences caused by bias in component 2 are far easier to detect than those from biases in component 1. So, the breaking of rotation symmetry by projection and Hilbert transform has only the effect of showing an additional (potentially artificial) preferred transition phase at a secondary angle. (see orange histograms). Appendix E shows the results for $y_{1\text{bias}} = 0.1, 0.2$ and $y_{2\text{bias}} = 0.1, 0.2$.

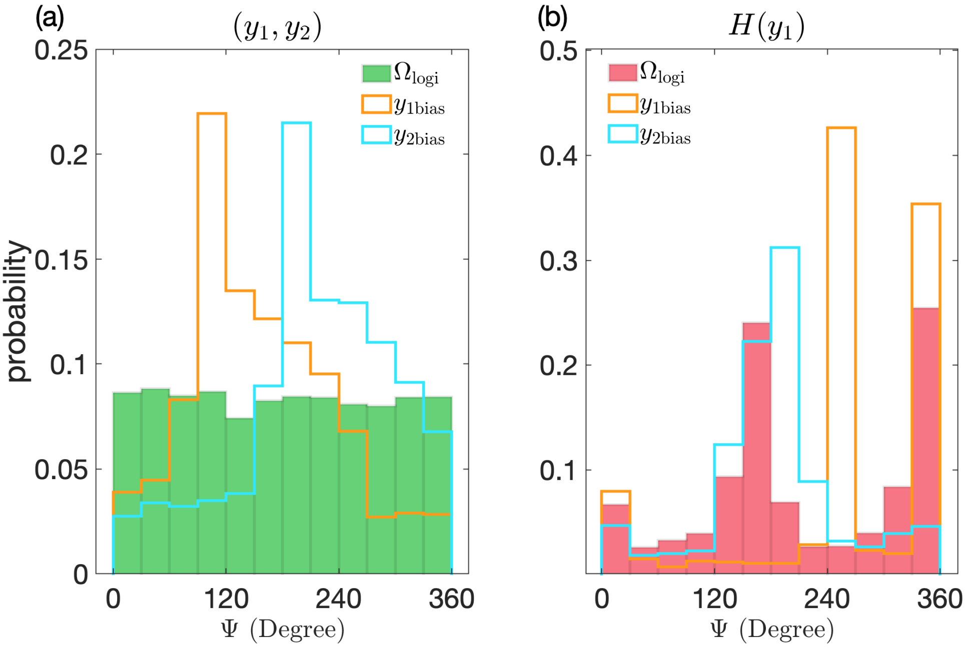


Figure 6.9: Panel(a) shows the histogram of $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{logi}},0°)$ (green bins), $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{bias}}^{y_1},0°)$ where $y_{1\text{bias}} = 0.3$ (orange stairs) and $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{bias}}^{y_2},0°)$ where $y_{2\text{bias}} = 0.3$ (cyan stairs); Panel(b) shows the histogram of $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{logi}},0°)$ (red bins), $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{bias}}^{y_1},0°)$ where $y_{1\text{bias}} = 0.3$ (orange stairs) and $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{bias}}^{y_2},0°)$ where $y_{2\text{bias}} = 0.3$ (cyan stairs).

6.3 Conclusion

To study how the preferred transition phase detected by the Markov chain model designed in Section 5.2.7 behaves for a rotationally symmetric nullmodel, we tested it on synthetic data $y(t,\Omega) \in \mathbb{R}^2$, produced by stochastic differential equations (6.1), which are a stochastic perturbation of the generalised Hopf normal form model introduced in Chapter 3 Section 3.2. We modified the stochastic differential equation by making the rotation frequency a function (Ω) that depends on the state $y \in \mathbb{R}^2$. Multiple parameters, including the large-amplitude oscillation frequency, the Euler-Maruyama step size, and the logistic growth rate, are carefully chosen to mimic the behavior observed in experimental data.

To closely replicate the observed data, two different classes of the frequency function (Ω_{fix} and Ω_{logi}) are tested, one is a constant frequency and the other is a logistic function of the amplitude $R = \sqrt{y_1^2 + y_2^2}$ (keeping the system rotationally symmetric). The latter aims to mimic more closely the absence of gradually decaying oscillations in transitions from stepping to freezing seen in the experimental data. Additionally, two further modifications of the frequency function ($\Omega_{\text{bias}}^{y_1}$ and $\Omega_{\text{bias}}^{y_2}$) have been tested. These modification introduce a preferred transition phase such that we can test if this preference is visible in artificial data compared to distortions and breaking of rotational symmetry by projection and embedding.

To analyze the synthetic data, a transition interval is defined by the first point hitting a radius of 0.5, and the 1999th point before that as the start of the interval. Only time series where an escape from oscillatory attractor to the steady state exists are considered. After filtering and defining the transition interval, the methodology outlined in Section 4 is applied to calculate the escape phase for each dataset.

We test three embedding methods (EM) for further analysis:

- Hilbert Transform: where the measured scalar time profile $x(t) = y_1(t)$ is embedded into the complex plane.
- Cartesian Coordinate (y_1, y_2): where both y_1 and y_2 variables are used in the unit circle, taking full advantage of the information contained in the synthetic data.
- Delay: which offers a different perspective on the embedding process by introducing a 2-D delay embedding approach.

Probability histograms of the escape phase Ψ_{min} are presented for each embedding method, both for constant (Ω_{fix}) and logistic (Ω_{logi}) frequency function. The histograms indicate that the delay embedding method may be able to recover rotational symmetry from a scalar projection of the signal. Furthermore, we observed two symmetric peaks in the histogram of the Hilbert Transform using the logistic frequency function Ω_{logi}. To further investigate this phenomenon, we examined the effect of the projection angle θ_r on the resulting escape angle. By rotating the measured time series $y_1(t)$ counterclockwise by different degrees, we observed that the Hilbert Transform of the rotated time series resulted in a consistent shift in the preferred transition angle, leading to a breaking of the rotational symmetry into a reflection symmetry, with two artificial modes of the transition phase distribution 180° apart from each other.

7. Summary and Conclusion

This thesis presented an in-depth exploration of the Freezing of Gait (FOG) phenomenon, a debilitating symptom experienced by individuals with Parkinson's Disease (PD). The investigation was carried out using time series analysis, mathematical modelling, and data collected from stepping-in-place experiments.

The study treated the dynamics underlying the transition from regular stepping to freezing as stochastically driven. The stochastic nature was modelled initially with a generalised Hopf normal form, perturbed by additive noise to a stochastic differential equation. The deterministic part of the model has two stable states: a limit cycle (modelling the stepping) and an equilibrium at the origin (modelling freezing). The transition from stepping to freezing is a noise-induced escape from the region near the deterministic limit cycle across an unstable limit cycle to the origin. We investigated the escape problem, quantifying the distribution of escape times and comparing simulation results to theoretical predictions for the mean survival time before transition. Furthermore, our analysis showed that the results are sensitive to (e.g.) stepsize of the numerical methodology employed on the accuracy of the model's predictions and simulations.

Through the combination of nonlinear time series analysis and mathematical modelling, we developed a methodology based on communicating classes in Markov chains for timing the transition from stepping to freezing and locating this transition in a reconstructed phase space, especially its angle (phase). Our patient-specific approach has potential applicability in the context of personalized interventions in the future, particularly for early warning signals for freezing episodes that could be incorporated into wearable devices or pressure-sensing shoe insoles.

In order to check how our Markov-chain based approach behaves on a nullmodel that had originally no preferred phase, we generated synthetic data using a stochastic differential equation to mimic the observed experimental stepping data, but was rotationally symmetric (ensuring that all transition phases should have equal probability) . Our analysis of the synthetic data tested different embedding methods for scalar projections of the nullmodel trajectories. We found that the delay embedding broke the rotational symmetry of the system (and the resulting uniform transition probability for all phases) severely even for the standard generalised Hopf normal form model with constant rotation frequency Ω_{fix}. Embedding based on Hilbert transforms breaks the rotation symmetry into a reflection symmetry if the rotation frequency depends strongly on the amplitude with two artificial modes in the resulting distribution of transition phases 180° apart. Despite these symmetry-breaking effects genuine preferences in the transition phase should still be detectable as histograms in Figure 6.9 and Appendix E show.

Overall, the research conducted in this thesis contributes to a more profound understanding of the complex dynamics involved in the FOG phenomenon, thereby providing novel insights that may inform future intervention strategies for PD patients. Furthermore, our results indicate promising directions for future research, particularly the implementation of our methodology in wearable technologies for real-time FOG prediction and early warning systems, ultimately aiming to improve the quality of life of PD patients.

Outlook Several hypothesis underlying the FOG phenomenon based on mechanistic studies have been proposed in the literature [46]. These involve: a threshold mechanism [47]; an inference mechanism [48]; a cognitive mechanism [49]; and a decoupling mechanism [50]. In Chapter 3 Section 3.2 presenting a nullmodel, we hypothesise that the transitions associated with spontaneous involuntary freezing episodes that are not triggered by any apparent external stimuli (as was the case in the experiments producing the data set we used) can be described as noise-induced escape in a bistable oscillator setting. In this sense, our modelling approach is phenomenological and could be most closely associated with the threshold model [47] mentioned above. This model does not account for mechanisms driving freezing but rather the dynamical (geometric) properties of the transition in to freezing. In fact, the analysis we have carried out and properties we have defined could be equally applied to both, patients and healthy individuals, should there be available data.

The quantity $\psi_{\min}$ extracted from the data by our method identifies for each freezing episode a unique threshold-independent time and stepping phase after which the subject is committed to freezing in this particular episode. The definition of such a characteristic quantity permits classification and clustering of freezing episodes and subjects according to in-person-between-episode and between-person mean and variation of $\psi_{\min}$. A non-uniform and possibly patient-specific distribution of preferred escape phases $\psi_{\min}$ is still compatible with the hypothesis of noise-induced escape from a limit cycle as proposed as an underlying mechanism in Chapter 3 Section 3.2. A non-uniform distribution would merely provide evidence against the rotation invariance present in nullmodel (3.11). In this case one can explore whether this quantity $\psi_{\min}$ is correlated with other characteristics of freezing or the disease (e.g., severity or frequency of freezing events, progression of the disease or effectiveness of therapies). If it is the case that there may be a preferred phase at which an individual patient enters into freezing is another interesting result in terms of future model development that accounts for this transition. It could lead to the development of intervention strategies such as early warnings of FE based on patient-specific phase information. Therefore, a potentially very important direction for future work is to systematically determine preferred transition phases by analysing transitions into FE not just in stepping but also in the case of walking and freezing. In principle the method permits one to construct an empirical transition matrix from several time series by merging the box counts in (5.8). Then we can not necessarily expect a single (unique) entrance state i_{tr}. There will still be a unique ψ_{min}, however, several local minima may make the transferred transition angle more uncertain.

Data availability

Full data sets and processing scripts are available at the following link `https://figshare.com/s/a14be7360925639736ba`.

8. Appendix

A Dependence on discretization parameters and length of transition intervals

Figure 1, Figure 2 and Figure 3 present a sensitivity analysis of the results with respect to various method parameters. Let us denote by $\psi_{\min}$ the phase (or angle) that corresponds

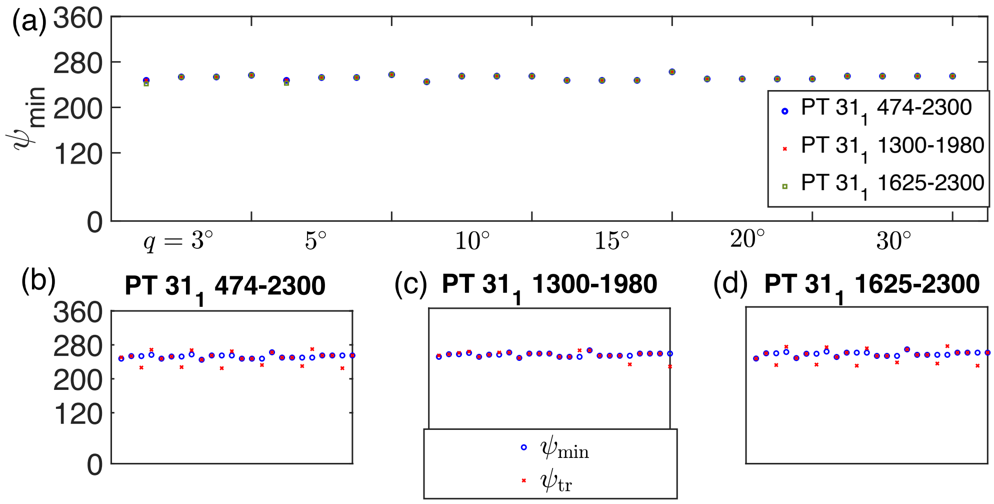


Figure 1: Panel (a): phase $\psi_{\min}$(in degree), where MET is minimal for different box sises and transition intervals. The label on the x-axis shows box sizes (p,q) for $p \in \{0.05, 0.1, 0.15, 0.2\}, q \in \{3°, 5°, 10°, 15°, 20°, 30°\}$. For each p, q increases within its range. Blue circles, red crosses and green square markers correspond to transition interval lengths, as indicated in the legend. Panels (b)–(d) show dependence of $\psi_{\min}$ (blue circles) and ψ_{tr} (red crosses) on same box sizes (using same x-axis as panel (a)) and transition interval.

to the box with minimal survival time MET_i in transition set F (derived by (5.27)). Figure 1(a) shows $\psi_{\min}$ for different discretization box sizes q, p, as introduced in Figure 5.1.4. The x-axis in Figure 1(a) shows all combinations of p and q while the color and marker type encode different transition interval lengths (blue circles, red crosses and green squares, respectively). The transition intervals tested are: $[474, 1980]$ (of length 15.06 s, containing more than four large-amplitude oscillations); $[1300, 1980]$ (of length 6.8 s, containing approximately four large-amplitude oscillations); and $[1625, 1980]$ (of length 3.55 s, containing approximately two oscillations) respectively. Panels (b–d) in Figure 1 compare for the same range of box sizes (p, q) and transition intervals the sensitivity of $\psi_{\min}$ (blue dots, the phase determined by minimal MET_i in the transition set F) to the sensitvity of ψ_{tr} (red crosses, the phase determined by the first state reachable in the absorbing set E). We observe that the phase $\psi_{\min}$ is less sensitive to box sizes and transition interval length than ψ_{tr}. Figure 2 summarizes the sensitivity for

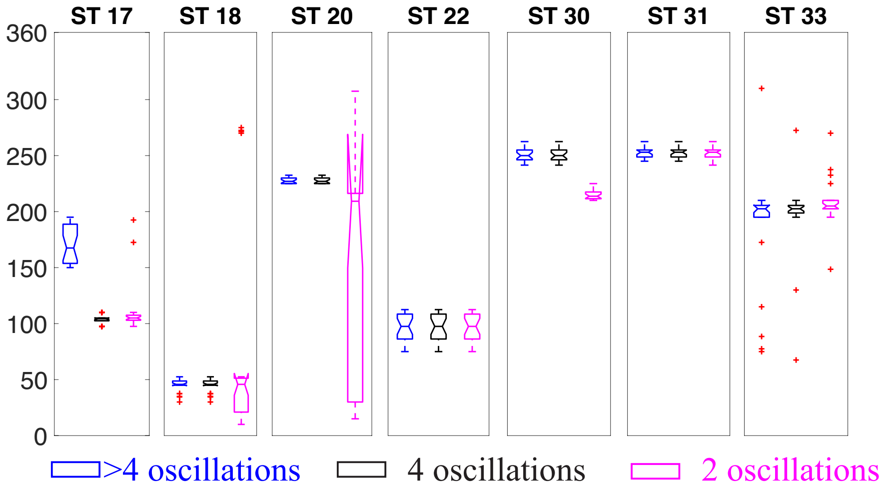


Figure 2: Boxplot for phase $\psi_{\min}$ for same range of box sizes as in Figure 1 and interval lengths for different subjects. The left boxplot (blue), middle boxplot (black) and right boxplot (magenta) of each patient represent 3 different choices of *transition interval*: $[t_{\mathrm{start}}, t_{\mathrm{end}}]$ in the same time series that correspond to the maximal possible transition interval, approximately four oscillations and approximately two oscillations, respectively.

further freezing episodes and different subjects as box plots for the phase $\psi_{\min}$, varying discretization box sizes over the same range as Figure 1 within each box plot. The left, middle and the right box plot of each subject are for 3 different transition interval lengths $[t_{\text{start}}, t_{\text{end}}]$: the transition interval for the left box plot contains more than 4 stepping periods (large-amplitude oscillations), the transition interval for the middle box plot contains approximately 4 oscillations, and the transition interval for the right box plot contains 2 oscillations, respectively. The distribution of extracted $\psi_{\min}$ supports our general observation that, if the stepping class contains sufficiently many oscillations (approximately 4), the resulting transition phase $\psi_{\min}$ is largely insensitive to our method parameters.

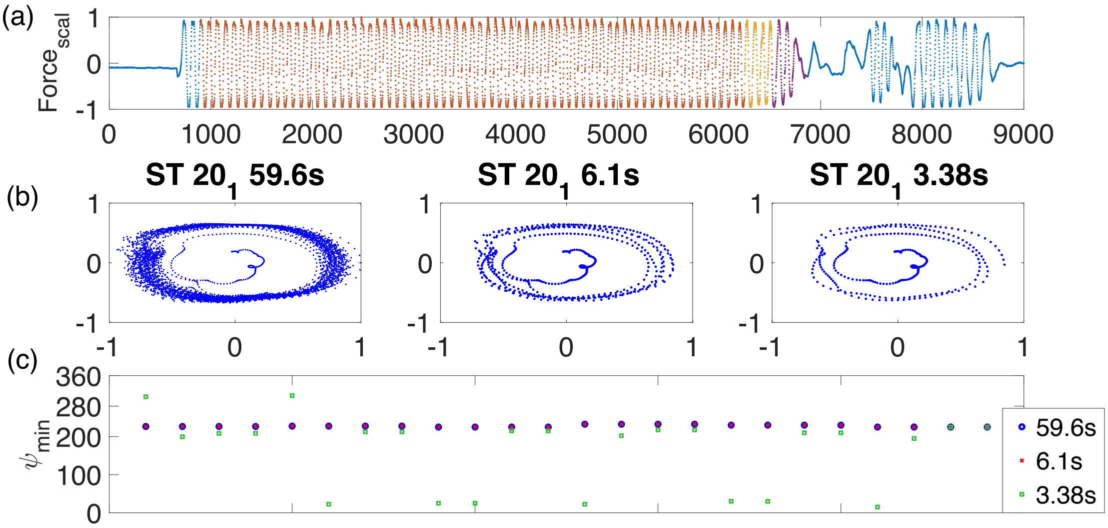


Figure 3: Panel (a) shows time series of freezing episode for subject ST20. Panel (b) shows the embedding trajectories for different transition interval lengths (left: many oscillations, middle: approximately 4 oscillations, right: approximately 2 oscillations) corresponding to the force data shown in Panel (a). Panel (c) shows the transition phase $\psi_{\min}$ for different box sizes p and q (see Figure 1).

Figure 3 shows in more detail the transition phases $\psi_{\min}$ for subject ST20 (Figure 2). In panel (b), in the case of 2 oscillations, the data in the stepping class is not sufficient to divide the state space into classes (transition set F and absorbing set E), as they are too sensitive to the transition interval length in this example. The above sensitivity analysis indicates that ≈ 4 oscillations preceding a freezing episode in the transition interval are recommended in order to obtain robust phase estimate in this data set. The Markov chain also allows us to generate "surrogates": time series $X_{\mathrm{mc}}(t)$ in the complex plane that share properties with the embedded time series $X(t)$ of the data shown in Figure 5.3(a) for each freezing event and each subject.

B Examples of *transition intervals* for subject ST31

Figure 4 shows two data sets for subject ST31, containing 8 different freezing episodes. The transition intervals selected for Markov chain analysis are highlighted in red. The time profiles from these highlighted sections have been used to determine in the phase $\psi_{\min}$ with minimal mean escape time from transition set F in Figure 5.12.

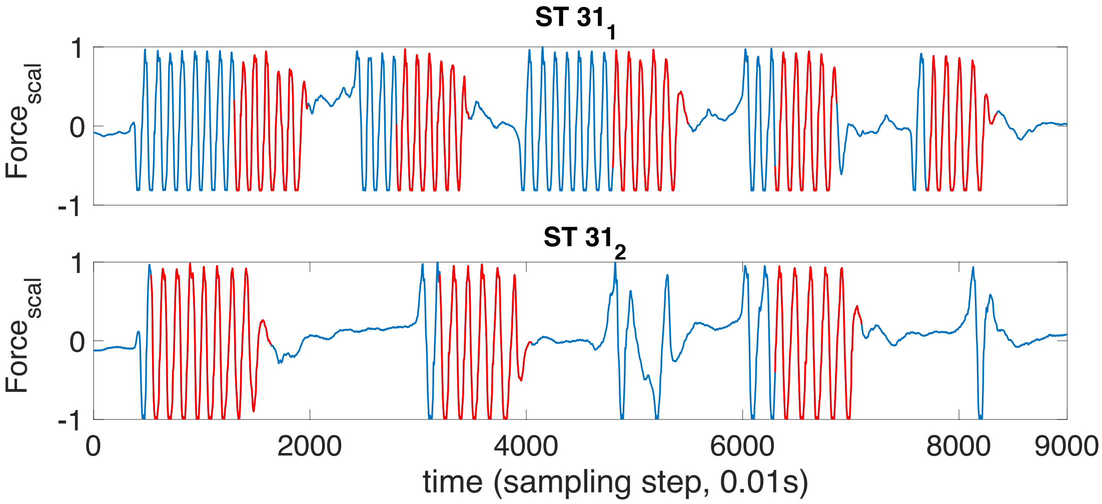


Figure 4: 8 *transition intervals* shown in red (in the form of time series) for ST31 with time step $\delta t = 0.01$.

C Simple example to show Relationship between MET_i and the mean escape time from transition set F

Now we present two simple examples, the Markov transition matrix A_1 and A_2 are given by:

$$A_1 = \begin{bmatrix} 2/3 & 1/3 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 0 & 1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 2/3 & 1/3 & 0 & 0 \\ 1/4 & 1/2 & 1/8 & 1/8 \\ 0 & 0 & 4/5 & 1/5 \\ 0 & 0 & 3/10 & 7/10 \end{bmatrix}$$

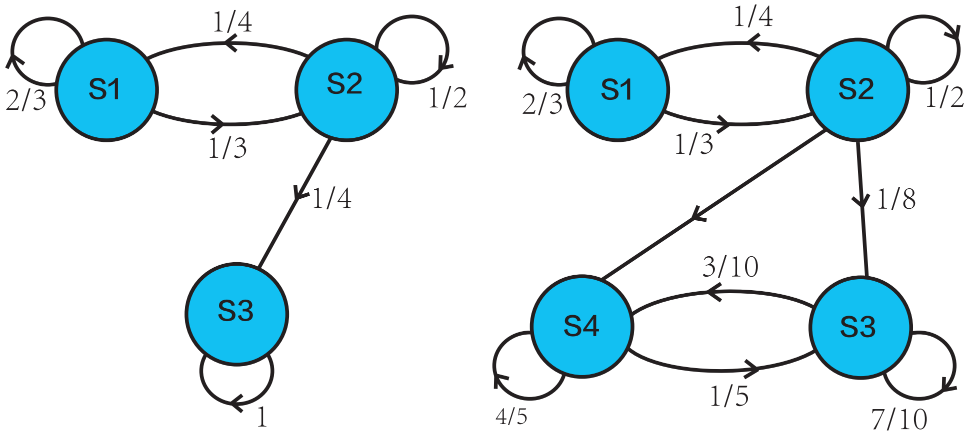


Figure 5: left: transition matrix A_1, state 1 and 2 are in the Transition set F, when the Markov chain goes into state 3, it will never back to state 1 or 2, thus state 3 is the absorbing state also the Absorbing set E. right: transition matrix A_2, state 1 and 2 are in the Transition set F, we define state 3 and 4 are the absorbing states since after t steps, state 3 and 4 are in the Absorbing set E. For $t \to \infty$, the Markov chain will only oscillate between state 3 and state 4.

Using Matlab to calculate the MET, we show the results in table 1. From row 1 and row 2 (or row 3 and row 4), MET calculated by the definition and the simplified equation are consistent, supporting the validity of our equation. Since the starting state in the Markov chain is a random event, so each MET calculated from a specific transition state should be equally considered:

$$\mathrm{MET} = \frac{1}{m}\sum_i \mathrm{MET}(\mathbf{s_i}) \quad i \in m \text{ transition states} \tag{1}$$

Row 5 shows the MET calculated by equation(1), which considered the evenly weight each of the starting state. The MET calculated for A_1 and A_2 are both 8.5s. Row 6 shows MET calculated by equation (5.41), which is close to the MET calculated by equation (1) and the expected MET. From this example and the equation(5.27), we can easily conclude the MET only related to the matrix with states corresponding to the Transition set F

		A_1	A_2
$\mathbf{s_1} = (1,0,\mathbf{0})$	MET(eqn(5.15))	10s	10s
$\mathbf{s_1} = (1,0,\mathbf{0})$	MET($\mathbf{s_1}$)	10s	10s
$\mathbf{s_2} = (0,1,\mathbf{0})$	MET(eqn(5.15))	7s	7s
$\mathbf{s_2} = (0,1,\mathbf{0})$	MET($\mathbf{s_2}$)	7s	7s
mean($\mathbf{s_1}+\mathbf{s_2}$))	MET(eqn(1))	8.5s	8.5s
	MET(λ_1) = $\frac{1}{1-\lambda_1}$	8.6056s	8.6056s

Table 1: MET for A_1 and A_2. The first row and the third row show the MET starting from state 1 and state 2 respectively, calculated by equation (5.15). The second row and the 4th row show the MET starting from state 1 and state 2 respectively, calculated by equation (5.27). The 5th row shows the mean value of the MET starting from each transition states. The 6th row shows the MET calculated by equation (5.41).

D Distribution of $Y^{\text{EM}}_{\text{last}}(\Omega)[R_{\text{last}}]$ with bias

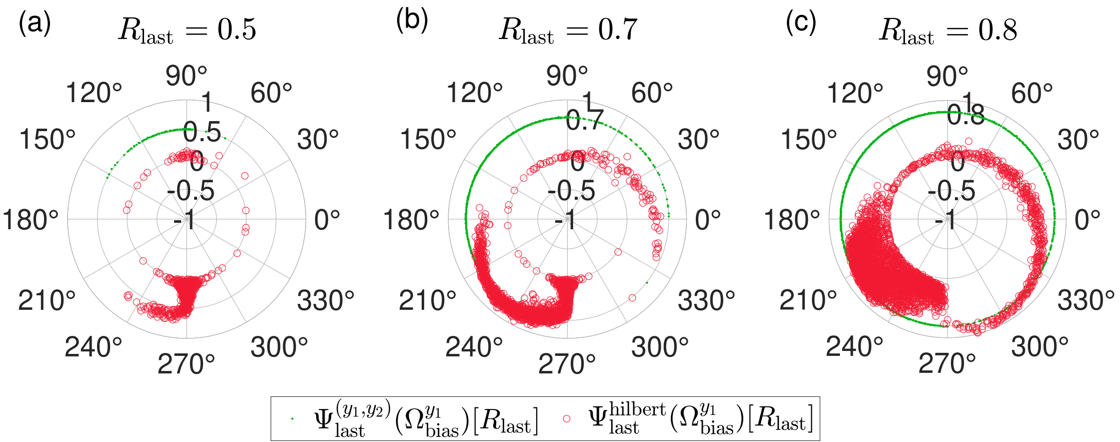


Figure 6: Panel(a)(b)(c) shows $Y^{(y_1,y_2)}_{\text{last}}(\Omega^{y_1}_{\text{bias}})[R_{\text{last}}]$ in green stars and $Y^{\text{hilbert}}_{\text{last}}(\Omega_{\text{logi}})[R_{\text{last}}]$ in red circles where R_{last} is 0.5, 0.7 and 0.8 respectively.

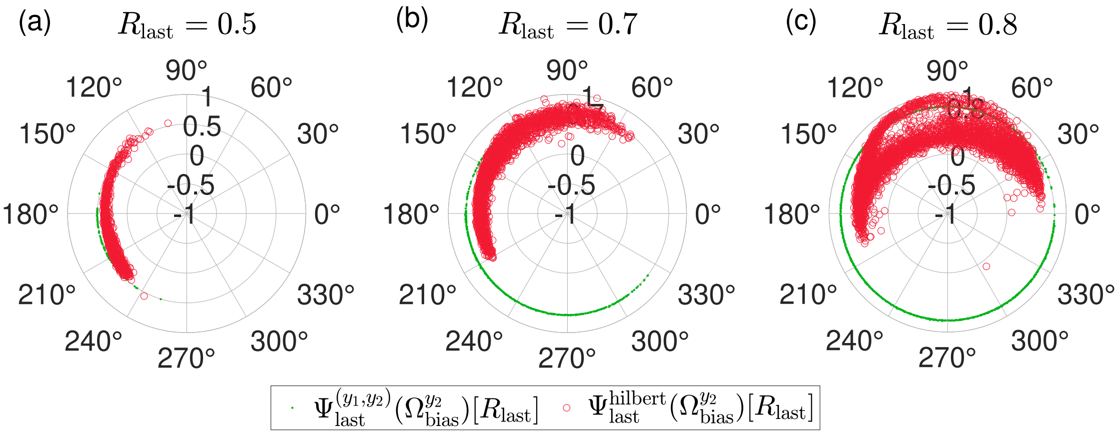


Figure 7: Panel(a)(b)(c) shows $Y^{(y_1,y_2)}_{\text{last}}(\Omega^{y_2}_{\text{bias}})[R_{\text{last}}]$ in green stars and $Y^{\text{hilbert}}_{\text{last}}(\Omega_{\text{logi}})[R_{\text{last}}]$ in red circles where R_{last} is 0.5, 0.7 and 0.8 respectively.

E Demonstration of probability distribution of Ψ_{min} when $y_{1\text{bias}} = 0.1, 0.2$ and $y_{2\text{bias}} = 0.1, 0.2$

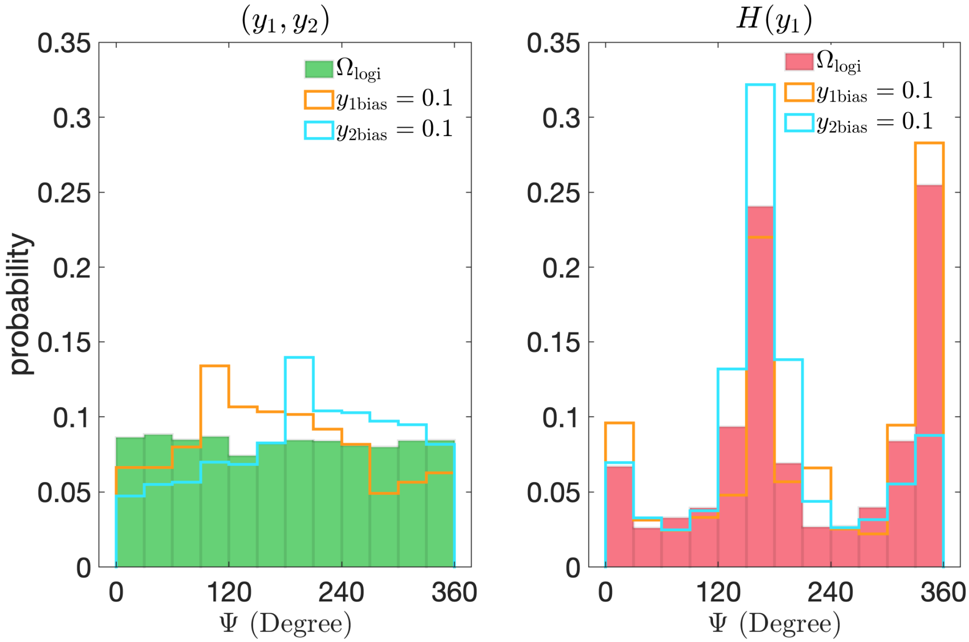


Figure 8: Panel(a) shows the histogram of $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{logi}}, 0°)$ (green bins), $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{bias}}^{y_1}, 0°)$ where $y_{1\text{bias}} = 0.1$ (orange stairs) and $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{bias}}^{y_2}, 0°)$ where $y_{2\text{bias}} = 0.1$ (cyan stairs); Panel(b) shows the histogram of $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{logi}}, 0°)$ (red bins), $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{bias}}^{y_1}, 0°)$ where $y_{1\text{bias}} = 0.1$ (orange stairs) and $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{bias}}^{y_2}, 0°)$ where $y_{2\text{bias}} = 0.1$ (cyan stairs).

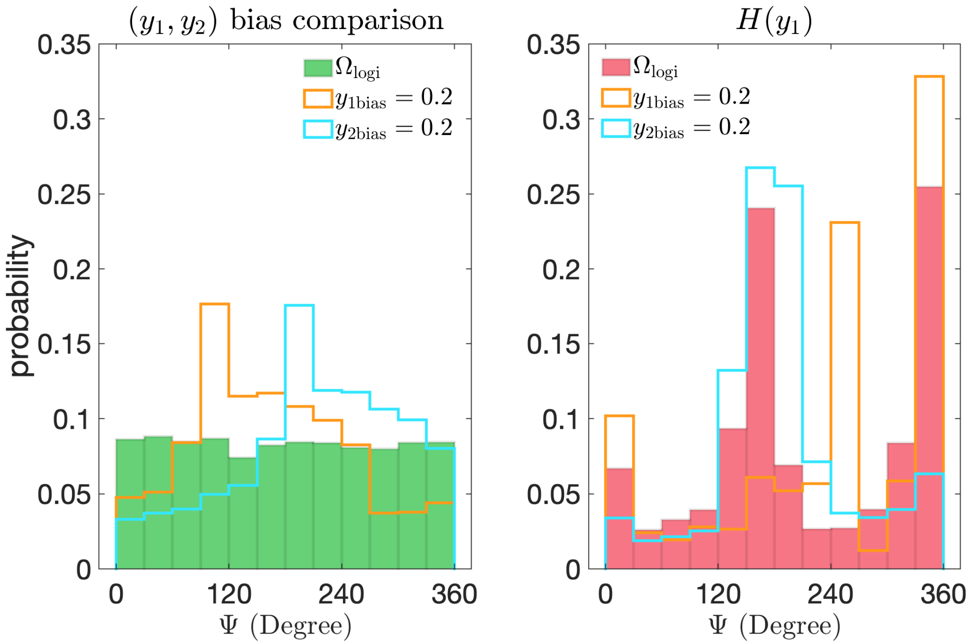


Figure 9: Panel(a) shows the histogram of $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{logi}}, 0°)$ (green bins), $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{bias}}^{y_1}, 0°)$ where $y_{1\text{bias}} = 0.2$ (orange stairs) and $\Psi_{\min}^{(y_1,y_2)}(\Omega_{\text{bias}}^{y_2}, 0°)$ where $y_{2\text{bias}} = 0.2$ (cyan stairs); Panel(b) shows the histogram of $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{logi}}, 0°)$ (red bins), $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{bias}}^{y_1}, 0°)$ where $y_{1\text{bias}} = 0.2$ (orange stairs) and $\Psi_{\min}^{\text{hilbert}}(\Omega_{\text{bias}}^{y_2}, 0°)$ where $y_{2\text{bias}} = 0.2$ (cyan stairs).

References

[1] Michael Feldman. Hilbert transform in vibration analysis. *Mechanical Systems and Signal Processing*, 25(3):735–802, 2011. doi:10.1016/j.ymssp.2010.07.018.

[2] Yuri A Kuznetsov Springer. Elements of Applied Bifurcation Theory, Second Edition. Technical report.

[3] JM Hausdorff, ME Cudkowicz, R Firtion, JY Wei, and AL Goldberger. Gait variability and basal ganglia disorders: stride-to-stride variations of gait cycle timing in parkinson's disease and huntington's disease. *Movement disorders : official journal of the Movement Disorder Society*, 13(3):428—437, May 1998. doi:10.1002/mds.870130310.

[4] Jeffrey M Hausdorff. Gait variability: methods, modeling and meaning. *Journal of neuroengineering and rehabilitation*, 2(1):1–9, 2005. doi:10.1186/1743-0003-2-19.

[5] N Giladi, TA Treves, ES Simon, H Shabtai, Y Orlov, B Kandinov, D Paleacu, and AD Korczyn. Freezing of gait in patients with advanced parkinson's disease. *Journal of neural transmission (Vienna, Austria : 1996)*, 108(1):53—61, 2001. doi:10.1007/s007020170096.

[6] Q J Almeida and C A Lebold. Freezing of gait in parkinson's disease: a perceptual cause for a motor impairment? *Journal of Neurology, Neurosurgery & Psychiatry*, 81(5):513–518, 2010. `doi:10.1136/jnnp.2008.160580`.

[7] Julie Nantel, Camille de Solages, and Helen Bronte-Stewart. Repetitive stepping in place identifies and measures freezing episodes in subjects with Parkinson's disease. *Gait Posture*, 34(3):329–333, jul 2011. `doi:10.1016/j.gaitpost.2011.05.020`.

[8] Meir Plotnik, Nir Giladi, and Jeffrey M. Hausdorff. Bilateral coordination of walking and freezing of gait in parkinson's disease. *European Journal of Neuroscience*, 27(8):1999–2006, 2008. `doi:10.1111/j.1460-9568.2008.06167.x`.

[9] Yashar Sarbaz and Hakimeh Pourakbari. A review of presented mathematical models in parkinson's disease: black-and gray-box models. *Medical & biological engineering & computing*, 54:855–868, 2016. `doi:10.1007/s11517-015-1401-9`.

[10] Midhun Parakkal Unni, Prathyush P Menon, Lorenzo Livi, Mark R Wilson, William R Young, Helen M Bronte-Stewart, and Krasimira Tsaneva-Atanasova. Data-driven prediction of freezing of gait events from stepping data. *Frontiers in Medical Technology*, 2:581264, 2020. URL: `https://doi.org/10.3389/fmedt.2020.581264`.

[11] Scott Pardoel, Jonathan Kofman, Julie Nantel, and Edward D Lemaire. Wearable-sensor-based detection and prediction of freezing of gait in parkinson's disease: a review. *Sensors*, 19(23):5141, 2019. URL: `https://doi.org/10.3390/s19235141`.

[12] Scott Pardoel, Gaurav Shalin, Edward D. Lemaire, Jonathan Kofman, and Julie Nantel. Grouping successive freezing of gait episodes has neutral to detrimental effect on freeze detection and prediction in parkinson's disease. *PLOS ONE*, 16(10):1–19, 10 2021. doi:10.1371/journal.pone.0258544.

[13] Gaurav Shalin, Scott Pardoel, Edward D Lemaire, Julie Nantel, and Jonathan Kofman. Prediction and detection of freezing of gait in parkinson's disease from plantar pressure data using long short-term memory neural-networks. *Journal of neuroengineering and rehabilitation*, 18(1):1–15, 2021. doi:10.1186/s12984-021-00958-5.

[14] Yosef Ashkenazy, Jeffrey M Hausdorff, Plamen Ch Ivanov, and H Eugene Stanley. A stochastic model of human gait dynamics. *Physica A: Statistical Mechanics and its Applications*, 316(1-4):662–670, 2002. URL: https://doi.org/10.1016/S0378-4371(02)01453-X.

[15] Matthew B. Kennel, Reggie Brown, and Henry D. I. Abarbanel. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A*, 45:3403–3411, Mar 1992. doi:10.1103/PhysRevA.45.3403.

[16] Floris Takens. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence, Warwick 1980: proceedings of a symposium held at the University of Warwick 1979/80*, pages 366–381. Springer, 2006.

[17] Alan Davies. Linear integral equations, by rainer kress. pp 299. dm78. 1989. isbn 0-387-50616-0 (springer). *The Mathematical Gazette*, 74(470):405–406, 1990. doi:10.2307/3618171.

[18] AV Bitsadze. Boundary value problems of analytic function theory. *Hazewinkel, Michiel, Encyclopedia of Mathematics, Springer*, 2001.

[19] BV Khvedelidze. Hilbert transform. *Encyclopaedia of mathematics. New York: Springer*, 2001.

[20] David Hilbert. Grundzüge einer allgemeinen theorie der linearen integralgleichungen. In *Integralgleichungen und Gleichungen mit unendlich vielen Unbekannten*, pages 8–171. Springer, 1989.

[21] Michael Rosenblum, Arkady Pikovsky, Jurgen Kurths, Carsten Schäfer, and Peter A Tass. Phase synchronization: from theory to data analysis. In *Handbook of biological physics*, volume 4, pages 279–321. Elsevier, 2001.

[22] Arkady Pikovsky, Jurgen Kurths, Michael Rosenblum, and Jürgen Kurths. *Synchronization: a universal concept in nonlinear sciences*. Cambridge university press, 2003.

[23] R.P. Kanwal. *Linear Integral Equations*. Birkhäuser Boston, 1996. URL: `https://books.google.com.tw/books?id=-bV9Qn8NpCYC`.

[24] Holger Kantz and Thomas Schreiber. *Nonlinear time series analysis*, volume 7. Cambridge university press, 2004.

[25] Liangyue Cao. Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D: Nonlinear Phenomena*, 110(1):43 – 50, 1997. `doi:https://doi.org/10.1016/S0167-2789(97)00118-8`.

[26] Andrey Andreyevich Markov. Extension of the limit theorems of probability theory to a sum of variables connected in a chain. *Dynamic probabilistic systems*, 1:552–577, 1971.

[27] Peter Watts Jones and Peter Smith. *Stochastic processes: an introduction*. CRC Press, 2017.

[28] Shaodi Qian and Chun-An Chou. A koopman-operator-theoretical approach for anomaly recognition and detection of multi-variate eeg system. *Biomedical Signal Processing and Control*, 69:102911, 2021.

[29] Natasza Marrouch, Joanna Slawinska, Dimitrios Giannakis, and Heather L Read. Data-driven koopman operator approach for computational neuroscience. *Annals of Mathematics and Artificial Intelligence*, 88(11-12):1155–1173, 2020.

[30] B. O. Koopman. Hamiltonian Systems and Transformation in Hilbert Space. *Proc. Natl. Acad. Sci.*, 17(5):315–318, 1931. `doi:10.1073/pnas.17.5.315`.

[31] I Mezic. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dyn.*, 41:309–325, 2005. URL: `http://www.springerlink.com/index/P3731751XX0V798N.pdf`.

[32] Igor Mezić and Andrzej Banaszuk. Comparison of systems with complex behavior. *Phys. D Nonlinear Phenom.*, 197(1-2):101–133, 2004. `doi:10.1016/j.physd.2004.06.015`.

[33] Igor Mezić. Analysis of Fluid Flows via Spectral Properties of the Koopman Operator. *Annu. Rev. Fluid Mech.*, 45(1):357–378, 2013. `doi:10.1146/annurev-fluid-011212-140652`.

[34] Steven L. Brunton, Bingni W. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control. *PLoS One*, 11(2):1–19, 2016. doi:10.1371/journal.pone.0150171.

[35] Matthew B Kennel, Reggie Brown, and Henry DI Abarbanel. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical review A*, 45(6):3403, 1992. URL: https://doi.org/10.1103/PhysRevA.45.3403.

[36] J. Guckenheimer and Y. A. Kuznetsov. Bautin bifurcation. *Scholarpedia*, 2(5):1853, 2007. revision #91035. doi:10.4249/scholarpedia.1853.

[37] B. Dybiec, E. Gudowska-Nowak, and P. Hänggi. Escape driven by α -stable white noises. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, 75(2):1–8, 2007. doi:10.1103/PhysRevE.75.021109.

[38] Nils Berglund. Kramers' law: Validity, derivations and generalisations. jun 2011. URL: http://arxiv.org/abs/1106.5799.

[39] Jennifer Creaser, Krasimira Tsaneva-Atanasova, and Peter Ashwin. Sequential noise-induced escapes for oscillatory network dynamics. *SIAM J. Appl. Dyn. Syst.*, 17(1):500–525, 2018. arXiv:1705.08462, doi:10.1137/17M1126412.

[40] C. W. Gardiner. *Handbook of stochastic methods for physics, chemistry and the natural sciences*, volume 13 of *Springer Series in Synergetics*. Springer-Verlag, Berlin, third edition, 2004.

[41] L. N. Trefethen, A. Birkisson, and T. A. Driscoll. *Exploring ODEs*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2018.

[42] Günter Rudolph. *Convergence properties of evolutionary algorithms*. Verlag Dr. Kovač, 1997.

[43] Fletcher Fan, Bowen Yi, David Rye, Guodong Shi, and Ian R Manchester. Learning stable koopman embeddings. In *2022 American Control Conference (ACC)*, pages 2742–2747. IEEE, 2022. `doi:doi:10.23919/ACC53348.2022.9867865`.

[44] Simon Lewis, Stewart Factor, Nir Giladi, Alice Nieuwboer, John Nutt, and Mark Hallett. Stepping up to meet the challenge of freezing of gait in parkinson's disease. *Translational Neurodegeneration*, 11(1):23, 2022. `doi:10.1186/s40035-022-00298-x`.

[45] Peter Eris Kloeden, Eckhard Platen, and Henri Schurz. *Numerical solution of SDE through computer experiments*. Springer Science & Business Media, 2002.

[46] Chao Gao, Jun Liu, Yuyan Tan, and Shengdi Chen. Freezing of gait in parkinson's disease: pathophysiology, risk factors and treatments. *Translational neurodegeneration*, 9:1–22, 2020. `doi:10.1186/s40035-020-00191-5`.

[47] Meir Plotnik, Nir Giladi, Jeffrey M Hausdorff, et al. Is freezing of gait in parkinson's disease a result of multiple gait impairments? implications for treatment. *Parkinson's disease*, 2012, 2012. `doi:10.1155/2012/459321`.

[48] Simon JG Lewis and Roger A Barker. A pathophysiological model of freezing of gait in parkinson's disease. *Parkinsonism & related disorders*, 15(5):333–338, 2009. `doi:10.1016/j.parkreldis.2008.08.006`.

[49] Jochen Vandenbossche, Natacha Deroost, Eric Soetens, Daphné Coomans, Joke Spildooren, Sarah Vercruysse, Alice Nieuwboer, and Eric Kerckhofs. Freezing of gait in parkinson's disease: disturbances in automaticity and control. *Frontiers in human neuroscience*, 6:356, 2013. `doi:10.3389/fnhum.2012.00356`.

[50] Jesse V Jacobs, John G Nutt, Patricia Carlson-Kuhta, Marilee Stephens, and Fay B Horak. Knee trembling during freezing of gait represents multiple anticipatory postural adjustments. *Experimental neurology*, 215(2):334–341, 2009. `doi:10.1016/j.expneurol.2008.10.019`.