Nonlinear measures and dynamics from time series data

A thesis

Submitted in partial fulfillment of the requirements

of the degree of

Doctor of Philosophy

By

Sandip Varkey George
(20132013)

To
Amma, who calls everyday
Achan, who keeps me grounded
and
Achu, who makes me laugh.

DECLARATION

I declare that this written submission represents my idea in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Sandip Varkey George
(20132013)

Date:

CERTIFICATE

Certified that the work incorporated in the thesis entitled **"Nonlinear measures and dynamics from time series data"**, submitted by **Sandip V. George** was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other university or institution.

Prof. G. Ambika

Date:

Acknowledgements

The visible parts of this thesis is the research work that was done over nearly six years. However, none of that would have been possible without the invisible support I have received from various people and organisations at a professional and personal level, through the course of my life, especially in the last six years. I am truly grateful to them. I would especially like to express my gratitude to the following people.

To, Prof. Ambika. The work on this thesis began five and a half years ago, in the last week of July 2013. As a new integrated PhD student, I'd walked into Ambika Ma'am's office to request to work with her. From that day to today, she has been a guide to me every sense of the word. I doubt I could have found a better person to work with in IISER. Who would make sure my work was always progressing steadily (refer Figure) without stopping me from going on my one hundred trips, piano classes and practices. Who would teach me how to speak, and how to write and how to teach. Who would make hilarious one liners in the middle of group meetings or dinners. Who would trust me enough to let me work on things I found interesting, even when our interests wouldn't overlap. And yet make sure I didn't wander off into the territory of youthful dreams with no results to show. Thank you for everything.

To, Prof. Ranjeev Misra. I suspect I would have had a significant proportion of this thesis empty if I hadn't met Ranjeev. But the bigger tragedy if I hadn't met him would have been missing out on him as a person! Our long evenings of discussion at IUCAA were one of the most intellectually stimulating parts of my PhD. From the guard stopping me everytime thanks to his envy for my great hair to our coffee breaks to my mild but growing fondness for second hand smoking! But beyond the intellectual, Ranjeev's cerebral humour and amazing dinner table conversations ought to be compelling reasons

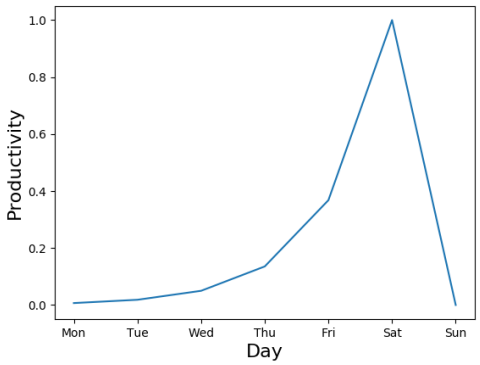


Figure 1: Variation of productivity with day of the week through my PhD. The spike on Saturday can be attributed to the group meeting being scheduled on that day. Productivity has been scaled to [0,1] for clarity.

for him to win an award for the coolest astrophysicist there is. I'm sorry we couldn't finish that ambitious bit in time for the thesis, but I'm on it!

To my RAC members. Prof. Anil Gangal and Prof. Deepak Dhar for making me question and learn to question everything I knew about nonlinear dynamics.

Dr. MS Santhanam who has been there to critique my work in conferences and my seminars. Dr. Prasad Subramanian who gave me insights on astrophysics, jazz and the nature of academia. Dr Rinku Jacob and Dr. K. P. Harikrishnan who've helped out with suggestions and clarifications through the course of my PhD.

To IISER, Pune for my funding, office space, hostel and computational facilities. The director Prof. Jayant Udgaonkar and the former director Prof. K.N. Ganesh, for providing the research environment and facilities, here at IISER Pune. The physics department, the chair and former chair of the program at IISER, Pune, Dr. MS Santhanam and Prof. Sunil Mukhi, for support and specifically for the travel support to various conferences within India. SERB for funding my travel to Dynamics Days Europe in Szeged, Hungary and CSIR for funding my travel to NODYCON-2019 in Rome, Italy. The Infosys foundation for funding my expenses to both these conferences.

To my group at IISER Pune, the Complex Systems Group. From Resmi Chechi who was there when it all began, and never left. Kashyap who walked in with me into this group, and made sure we did everything together including writing our theses together! Snehal who has become many things to me: office mate, teacher, guide, object of tickling (sorry!), and most of all, friend. Comrade Kajari who peppered my years here with music, food and philosophy and introduced me to more people than I could keep track of. Kunal, for all the interesting office discussions we had. Yamini whose entry forced as to start doing things as a group. Sneha, who will be the only remaining member in another month (Have a tomato if that begins to scare you). All the undergrads who came in between, specially Harsh, Amey, Anwesha and Raagini whose names come to my mind. Ending with Saaransh, who was there at the end, whose incessant talking made me younger, wiser and a bit bruised (you really need to stop the punching!).

To my friends in physics, who've helped out in very many ways. Harshini, my Enid Blyton partner who can speak with one half of her face (the number of times we've done that)! Sanku, who completed our "Extended complex systems group". Amrutha who sat next door, for all the chai breaks. Tomin for all those crazy discussions. And Aakanksha, the living reminder in IISER of a time I fondly recollect!

To my supremely amazing integrated PhD batch of 2013. From the minute I walked into HR2 to the day I will walk out of Room 320, you guys have been the most amazing batch I could have asked for. For all the trips and lunches and dinners and birthdays and conversations and stays and dances and kaaravans and everything we've done together! Thank you. Considering the number of times I've had to do this recollection to make sure everyone was invited to things, its only fair I do this one more time and name all of you. Neeladri, Anish, Ron, Bharat, Swati, Shivani, Jay, Mehak, Deepak, Aditi, Kashyap, Dhriti, Amar, Akhila, Divya, Anshul, Adarsh, Charu and Harpreet.

To my Int PhD seniors and juniors. Sukrut (for all those sweets from home), Mukul (for all the hugs), Ankitha (for your smile), Tomin (for so many many memories) and Farhan 'lil Amin!

To Anish, Neeladri and Ron(Its alphabetical :P). The four Indian wisemen. The controlled aggregation (with manager). My family away from family. My confidantes and brothers. You guys have been the centre of my life at IISER. Thank you for reading through my papers, editing my presentations, listening to me whine about work, listening

to me whine about whining, being the only people who know what I work on(if you say gaps one more time, I'll throw you off a building, Ron) and turning up for my talks like my personal cheerleaders. Neeladri, my permanent roommate, rudest and sweetest person on the planet! I suspect about 90% of the reason I'm still single is because everyone is convinced we're together. I'm not sure how I'm going to handle life without you picking up after my mistakes. Anish, my refuge in times of distress. THE nicest and possibly wisest guy I know. I'll miss our swimming, running, home gymming, church going, birthday shopping, movie making, philosophising ... Oh man! (And to think you chose to speak to my mom instead of me.) Ron, my three days younger brother. All those long drawns talks with you about everything under the sun, moon and stars have honestly made me a wiser person(can't say the same for you :P). Thanks for those hundreds of ("quick") morning naps in your room. Being around you three makes me happy. :)

The people I've worked with in various performances through IISER. The people at Sea Minor (2017), the IISER choir, the Independence day play (2016), the Sunday evening choir at St Antony's, theatre workshops, Muziclub, Pune musicale, the Onam play and song (2016), grad students club, march for science and physics day 2018! All the things that kept me sane at the end of all the physics.

To the students and faculty I've worked with in various TA-ships, and the friends I've made in the process. PHY 311(2017), HSS 102 (2017), PHY 221 (2016), PHY 202(2016) and PHY 101(2015).

Friends who've become part of my life in IISER for various reasons. Suman (For those conversations that start about music or work and end up being about everything else), Kabir (nicest celebrity I know), Pushkar("Coolness" should be your middle name, or is it already?), John (For all the music and drama you brought into our lives), Swapna(for all the temple-talk, temple-walk, chai and gossip.), Kitchi (that white tam-bram types who sounded more smug than neccesary) and Abhi (For all the evening chai, random dinners and walks discussing patriarchy, heteronormaivity, casteism, sexism and what not! And yes, the gossip!). I'll miss you all.

Finally, the people who brought me into the world of physics at St. Stephens. Dr. Bikram Phookun and Dr. Abhinav Gupta, who've inspired me in countless ways to learn and understand physics. My friends from Stephens who continue to be my closest companions seven years later. Malayaja, whose questions on nonlinear dynamics, physics and life keeps me re-examining my beliefs on a regular basis. Bijoy, Lijo, Lijo(Wait, did I repeat that?), Stella, Eliza and Elizabeth for being there. Ria, Ranjeet, Gautham and Mary Ma'am for never letting go, all the way from when I was a pre-teen.

My wonderfully huge family. Coming home to all of you for Christmas is one of the happiest parts of my year.

To Valiapachan, Valiyamachi and Ammachi. Who taught me, cared for me and told me stories. I wish you were here.

To everything that went right in these last five and a half years. The serentipity that made sure I stayed on when it seemed hard. The people who turned up when they needed to. The conversations that came when it didn't seem like a great day. To the sweet that always seemed to follow the bitter. And the strength that arrived whenever weakness took a toll. Thank you.

Finally, to Amma, Achan and Achu. And Sreejith who joined us. Thanking you is impossible. And I'm glad I don't have to try.

Thank you, all.

Sandip

Abstract

Our attempts to understand the dynamics of real world complex systems often rely on the observations made on them or average responses from them. The science of extracting information from such time series or observational data forms the area of time series analysis. The dynamics underlying most of the systems in nature is nonlinear and therefore complex. This nonlinearity leaves signatures in the dynamics of the system, captured by the time series of its variables or observables. Nonlinear time series analysis strives to seek and interpret these signatures in order to capture the nature of the underlying dynamics. Such studies are hindered due to various shortcomings of the data itself. Some of the primary problems we encounter in this respect are noise, data gaps and finite size of data sets. This thesis attempts to address these problems in the context of nonlinear time series analysis.

The origins of dynamical systems theory can be traced back to the celestial. Many of the ideas of this theory were developed while trying to address the three body problem. Subsequently, nonlinear dynamics has found applications in many fields of astrophysics. These include stellar pulsations, accretion disc physics, galactic models and so on. The use of time series analysis in astrophysics has been limited in the past due to the absence of long, continuous and high quality datasets. The advent of space telescopes could overcome some of these problems. Subsequently period doubling, chaotic and strange non chaotic behavior etc have been established in a large variety of stars. This thesis is a search for the dynamics of variable stars in more detail, seeking out signatures of nonlinearity and chaos in various scenarios from observational data.

After presenting the basics of nonlinear time series analysis and the physics of variable stars, we proceed to address the issues commonly encountered in the nonlinear time series analysis of real world data. We identify three main issues in this context, namely the presence of data gaps, noise and finite sizes of data.

We start by analyzing the effect that data gaps have on the estimation of the correlation dimension and multifractal spectrum from time series data. This is implemented by introducing gaps into long evenly sampled datasets of standard nonlinear dynamical systems. The frequency and size of gaps introduced are drawn from two Gaussian distributions, with varying means. After the introduction of gaps, the time series is merged, ignoring the gaps and the quantifiers of interest like correlation dimension, multifractal measures etc. are calculated for this gap-affected or unevenly sampled time series. The variation of the value of the quantifier from the evenly sampled value then enables us to identify a region where reliable conclusions can be drawn about the nature of the underlying dynamics of the system. We apply this to calculate the correlation dimension, D_2, of a few pulsating variable stars, from their light curves. We also use the results of the analysis to calculate the multifractal spectrum of multiple ecological and meteorological time-series. In these contexts, we use the method of surrogate data testing to establish that the multifractality arises as a result of deterministic nonlinearity.

We then proceed to address the question how presence of noise can confuse the conclusions while identifying the underlying dynamics of a nonlinear dynamical system. We point out the difficulty in differentiating between limit cycle and chaotic dynamics, in a dynamical system evolving in the presence of noise. We show that the bicoherence, which is a higher order spectrum, proves to be a useful tool in differentiating between these two dynamical states.

We note identifying strange non chaotic behavior from data is a major challenge. Spectral scaling of peaks in the strobed power spectrum is one of the most accepted methods to identify strange non chaotic behavior. We show that noise contaminated quasiperiodicity shows exactly the same scaling behavior as a strange non chaotic time series. The use of a bicoherence based filter while identifying peaks during scaling can differentiate between the two dynamical states. We further use bicoherence based measures to analyze the underlying dynamics of RR Lyrae stars. We show that while RRab Lyrae stars exhibit chaotic behavior, RRc Lyrae stars can be grouped into two subclasses that exhibit strange non chaotic and quasiperiodic dynamics.

We also use the methods of time series analysis to study the Kepler light curves of over contact binary stars. Using the correlation dimension, multifractal spectrum and the bicoherence, we see that most overcontact binaries show chaotic dynamics.

Recurrence networks have been shown to be successful in the analysis of small data sets. We utilize the power of recurrence networks to distinguish between the different classes of RRc Lyrae stars identified using the bicoherence analysis. We also compute the measures of recurrence networks for a large number of binary stars and show how they can help classify close binary stars into semi detached, over-contact and ellipsoidal binaries.

One of the significant results of our study is that we could relate the nonlinear measures computed from time series with relevant astrophysical properties of the stars. Thus the extend of contact between the companion stars in these cases is shown to be significantly correlated with its nonlinear properties, specifically the correlation dimension and bicoherence. Hence the computation of these nonlinear measures gives us an estimate of the extend of contact between these stars.

The thesis is organized as follows.

In Chapters 1 and 2, we introduce the basics of nonlinear time series analysis and the physics of variable stars respectively.

In Chapter 3, we describe the effect that data gaps have on estimated the correlation dimension and multifractal spectrum from time series data and use the results to analyze variable stars light curves and multiple ecological and meteorological time-series.

In Chapter 4, we use the bicoherence function to distinguish between periodic states contaminated with noise and chaotic and strange non chaotic states. In Chapter 5 we present the measures correlation dimension, multifractal spectrum and the bicoherence computed from data and discuss the dynamics of overcontact binary stars derived from these measures.

In Chapter 6, we present recurrence networks constructed form observational data of RRc Lyrae stars and contact binary stars and use the characteristics of the recurrence networks to classify them into various sub categories.

In the concluding chapter, we summarize the importance of the results presented in the thesis and provide future directions for research in this area. Identifying noisy quasiperiodicity and distinguishing it from strange non chaotic behavior is a difficult task, which we could successfully address in our study. We could relate measures from nonlinear

time series analysis to astrophysical properties of the stars studied. This could lead to classification of different types of astrophysical systems just from their observational data alone.

The work presented in this thesis has appeared in the following publications and conference presentations:

Research publications:

[1] **George, S. V.**, Ambika, G., & Misra, R. (2015). *Effect of data gaps on correlation dimension computed from light curves of variable stars*, Astrophysics and Space Science, 360(1), 5.

[2] **George, S. V.**, Ambika, G., & Misra, R. (2017). *Detecting dynamical states from noisy time series using bicoherence.* Nonlinear Dynamics, 89(1), 465-479.

[3] **George, S. V.**, & Ambika, G. (2017). *Nonlinearity in data with gaps: Application to ecological and meteorological datasets.* Indian Academy of Sciences Conference Series(2017) 1:1 DOI: 10.29195/iascs.01.01.0002

[4] **George, S. V.**, Misra, R., & Ambika, G. (2019). *Fractal Measures and Nonlinear Dynamics of Over-contact Binaries.* Communications in Nonlinear Science and Numerical Simulation (2019): 104988

[5] **George, S. V.**, Misra, R., & Ambika, G. (2019) *Classification of Close Binary Stars Using Recurrence Networks.* arXiv preprint arXiv:1907.10602 (2019).

[6] **George, S. V.**, Misra, R., & Ambika, G. (2019) *Nonlinear Dynamics of RRc Lyrae Stars* Accepted for publication in NODYCON 2019 Proceedings (2019).

Papers/posters presented in conferences:

[1] *Nonlinear dynamics of RRc Lyrae stars* (NODYCON-2019, Rome)

[2] *Strange Correlations From the Nonlinear Dynamics of Over-contact Binary Stars* (CNSD-2018, Delhi)

[3] *Detecting Dynamical States From Real World Datasets* (DDE-2017, Szeged)

[4] *Detecting Dynamical States Using Bicoherence Function* (CNSD-2016, Kolkata)

[5] *Bispectral Analysis for Strange Non-chaotic Time Series* (CSAS-2016, Chennai)

[6] *Effect of uneven sampling on correlation dimension computed from time series data* (DDAP-2014, Chennai)

Contents

Chapter 1

Introduction: Nonlinear Time Series Analysis

Our understanding of the physical world has been largely shaped through interpretation of observations. A very important kind of observation is of phenomena that vary over time. A series of such observations, ordered in time is called a time series. The techniques used to analyze these time series are what constitute the field of time series analysis.

Depending on the nature of the underlying system from which the time series is derived, we may broadly classify the time series as linear and nonlinear time series. The dynamics of real world systems is generally nonlinear in nature. To develop an understanding of these systems, we rely mostly on observations from any one of its variables or an average response. Modeling these systems based on the observations of just one of their variables alone can be challenging. Nonlinear time series analysis aims to provide possible directions which aid in this modeling process. A second way to classify time series is based on whether the underlying process is primarily stochastic or deterministic. Natural systems may have a combination of deterministic or stochastic dynamics or may even switch between these two types of dynamics. Distinguishing between deterministic and stochastic phenomena based on the time series is an important question that intrigues scientists. One of the objectives of time series analysis is to classify time series into these subcategories by determining properties of the underlying system from which the data is derived.

1.1 Nonlinear Dynamical Systems

Dynamical systems are systems whose states change in time. Depending on whether this change is mapped continuously in time, using differential equations or discretely in time using difference equations, we have continuous and discrete dynamical systems. In this thesis we will concentrate mostly on continuous dynamical systems. Discrete dynamical systems can be generically represented by

$$x_{n+1} = F(x_n, r) \tag{1.1}$$

where x_n is the state of the system at time step n, r is the list of parameters and F is a nonlinear function of x_n.

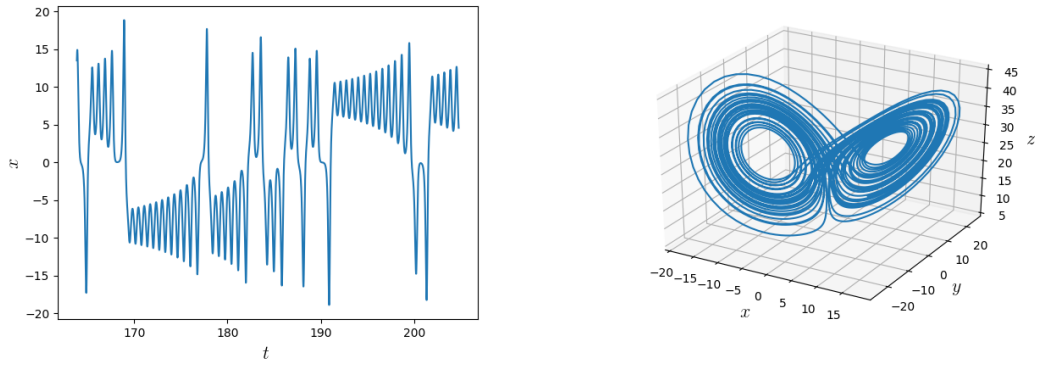


Figure 1.1: Time series and state space of the Lorenz system. Parameters used are $\sigma = 10$, $\rho = 28$ and $\beta = \frac{8}{3}$

Continuous dynamical systems can be generically represented by

$$\frac{d\vec{x}}{dt} = f(\vec{x}, a) \tag{1.2}$$

In general for an n dimensional system, $\vec{x} \in \mathbb{R}^n$, a is the list of parameters and f is a nonlinear function of $\vec{x}$. The nature of dynamics may either be stochastic or deterministic. We will proceed to discuss various deterministic dynamical systems and the dynamical states they exhibit.

Let us consider a few examples of nonlinear deterministic dynamical systems. The systems we consider here will be used in subsequent chapters as test systems to illustrate our results.

Lorenz system

The Lorenz system was first derived by meteorologist E. N Lorenz to describe unpredictable changes associated with the weather. The equations were derived through mode truncation of a partial differential equation describing fluid convection in a fluid cell [1,2]. The equations are

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= x(\rho - z) - y \\ \dot{z} &= xy - \beta z \end{aligned} \tag{1.3}$$

Here the x variable is related to the rate of convective overturning, y to the horizontal temperature variation and z to the vertical temperature variation. σ is proportional to the Prandtl number, ρ to the Rayleigh number and β to the dimensions of the region under consideration [2]. The time series of the x variable and the state space of the system is shown in Figure 1.1.

Rössler system

The Rössler equations were constructed by Otto Rössler in an attempt to design chaotic flow in a system with only one nonlinearity [3]. While the equations were constructed with no immediate physical interpretation in mind, it was subsequently found to be useful

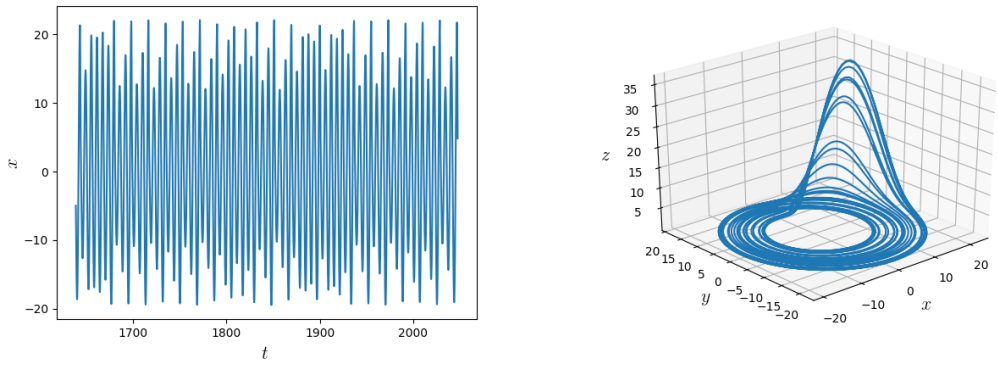


Figure 1.2: Time series and state space of the Rössler system. The parameters used are $a = 0.1, b = 0.1, c = 14$.

in describing chaos in certain chemical reactions [4]. The form of the equations are as follows.

$$\begin{aligned} \dot{x} &= -y - z \\ \dot{y} &= x + ay \\ \dot{z} &= b + z(x - c) \end{aligned} \tag{1.4}$$

The time series from the x-variable and the state space of the attractor is shown in Figure 1.2

Pendulum

A simple pendulum is an idealization of a real bob oscillating on a string. The air drag, mass and extensibility of the string, dimensions of the bob etc are ignored for convenience in the simplest model, whose equation is given as

$$\ddot{\theta} = \sqrt{\frac{g}{l}} sin(\theta) \tag{1.5}$$

Here g is the acceleration due to gravity and l is the length of the string. The small angle approximation to this is the simple harmonic oscillator. A variation of the simple pendulum is the driven damped pendulum, which incorporates friction and a driving force. The equation is then given by,

$$\ddot{\theta} + \gamma\dot{x} + \sqrt{\frac{g}{l}} sin(\theta) = Acos(\omega t) \tag{1.6}$$

The second term here describes friction and the $R.H.S$ defines the driving term. Time series and state space plots for equation 1.6 is shown in Figure 1.3.

1.1.1 Dynamical States

One of the important pre-requisites for understanding dynamical states is the concept of a state space. The state space is the space of all the variables in a system. A point in the state space gives a possible state of a system. The time evolution of a system leads to trajectories in state space. A certain dynamical system in a particular state may give rise to different trajectories in state space depending on the parameter values and initial

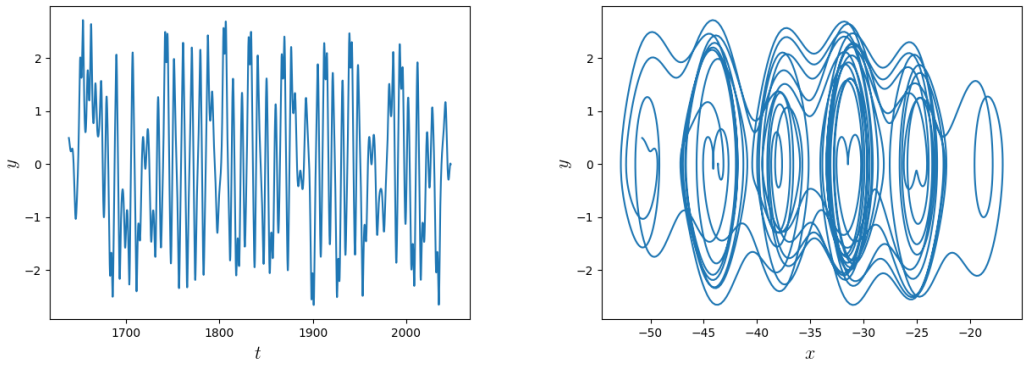


Figure 1.3: Time series and state space of the driven damped pendulum. The parameters used are $\gamma = 0.05, \omega = 0.7, A = 0.6, l = g$.

values for the dynamical variables. We explore the different dynamical states that can be exhibited by nonlinear dynamical systems. In order to analyze these different dynamical states let us consider a system given by the general equation

$$\dot{\vec{X}} = f(\vec{X}) \tag{1.7}$$

Fixed point

A fixed point is the state where the system settles to a state of no dynamics i.e.

$$\dot{\vec{X}} = 0 \tag{1.8}$$

For simplicity we can assume a two dimensional state space. $\vec{X} = x\hat{i} + y\hat{j}$. Equation 1.7 can be rewritten as

$$\begin{aligned} \dot{x} &= f(x, y) \\ \dot{y} &= g(x, y) \end{aligned} \tag{1.9}$$

The fixed points are then found as the points where

$$\begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0 \end{aligned} \tag{1.10}$$

The stability of the fixed point, i.e. how trajectories that are near the fixed point behave, is given by the derivatives of $f(x, y)$ and $g(x, y)$ evaluated at the fixed point. The matrix of the derivatives is called the Jacobian matrix J.

$$[J] = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x} & \frac{\partial f(x,y)}{\partial y} \\ \frac{\partial g(x,y)}{\partial x} & \frac{\partial g(x,y)}{\partial y} \end{bmatrix} \tag{1.11}$$

The determinant, Δ and trace, Tr of the matrix determine the stability of the fixed point. The time series and state space of the Rössler system approaching a fixed point is shown in Figure 1.4. As the parameters of a system change, fixed points may change stability which could lead to other types of dynamics like limit cycles in these systems.

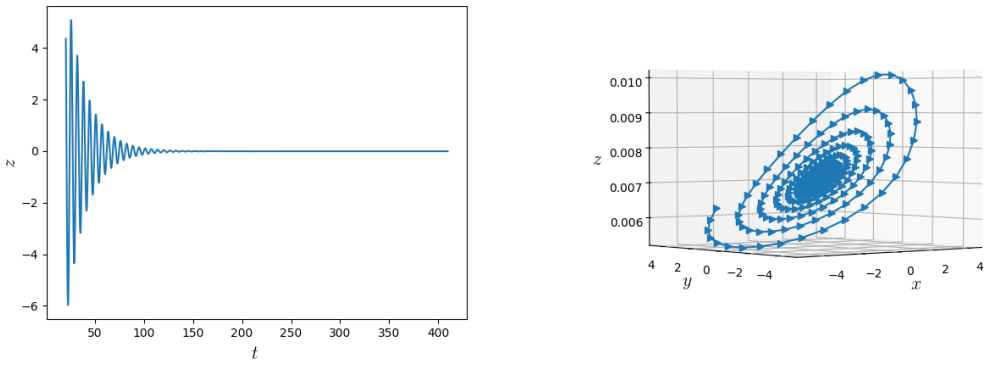


Figure 1.4: Rössler system approaching an attracting fixed point for parameter values $a = -0.1, b = 0.1, c = 14.0$

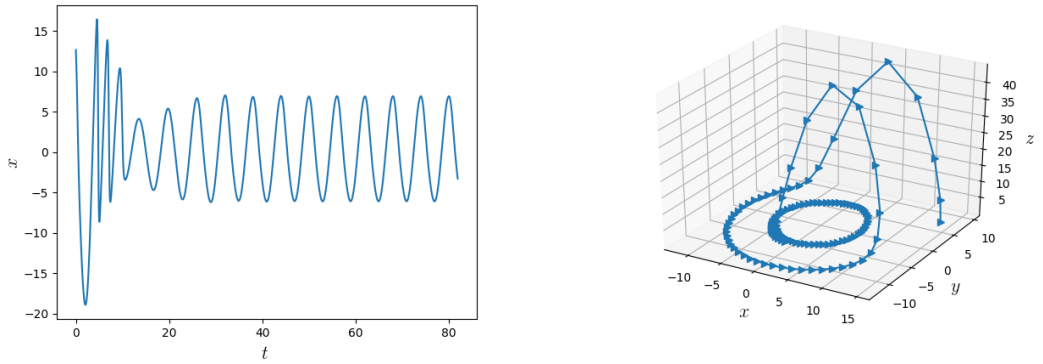


Figure 1.5: Rössler system approaching an attracting limit cycle for parameter values $a = 0.1, b = 0.1, c = 4.0$

Limit cycles

For state spaces with dimension greater than 1, limit cycles are also a possible dynamical state. In state space, these would be represented by a closed loop. It is a special type of periodic orbit where, in a neighbourhood of the limit cycle, trajectories may be either attracted to or repelled away from it, depending on its stability [5]. This is different from periodic orbits in a harmonic oscillator,for instance, where there is no attracting or repelling set. An attracting limit cycle is often called an $\alpha-$limit cycle and a repelling one as an $\omega-$limit cycle [6]. The time series and state space plot of a limit cycle of the Rössler system is shown in Figure 1.5.

Quasiperiodicity

Quasiperiodicity is a dynamical state that arises in systems where two or more competing frequencies, that are in an irrational ratio with each other, exist. This may occur either due to the natural frequency of the system being incommensurate with one or more driving frequencies, or due to the development of two or more natural frequencies in the system itself [5]. If n independent incommensurate frequencies are present in a system, the quasiperiodic motion takes place on an n-torus. A quasiperiodic orbit is made up of linear combinations of the different frequencies in the system [7]. We illustrate this point

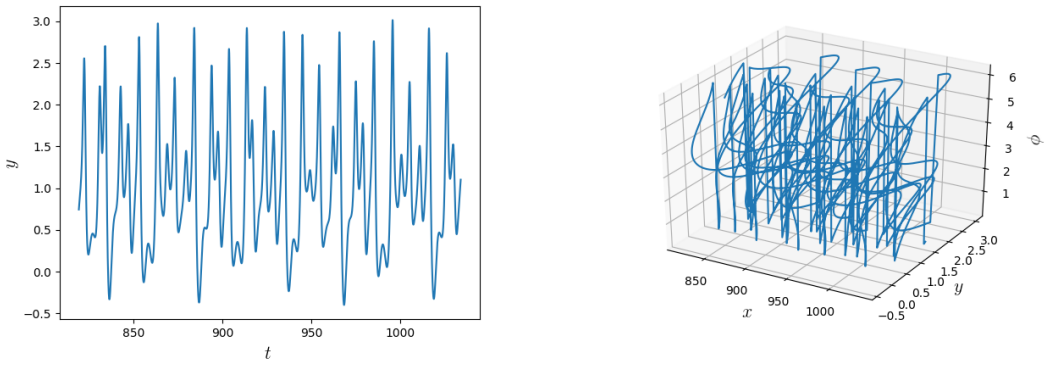


Figure 1.6: Doubly driven pendulum showing quasiperiodicity for parameter values $K = 1.34, V = 0.55, p = 3.0, \omega_1 = \frac{\sqrt{5}-1}{2}$ and $\omega_2 = 1.0$

further when we discuss power spectra in Section 1.2.

As an example system where quasiperiodicity exists, we consider a doubly forced pendulum [8].

$$\frac{1}{p}\frac{d^2\theta}{dt^2} + \frac{d\theta}{dt} - cos\theta = f(t) \tag{1.12}$$

where

$$f(t) = K + V[cos(\omega_1 t) + cos(\omega_2 t)]$$

The time series and state space trajectory of the doubly driven pendulum in the quasiperiodic state is shown in Figure 1.6.

Chaos

One of the most fascinating dynamical states that can be observed in nonlinear dynamical systems with state space dimension greater than 2, is chaos. A chaotic trajectory is bounded and irregular, composed of all possible frequencies, making it difficult to distinguish from randomness. In general, as the values of the parameters change, a dynamical system may approach chaos in various ways. These are called routes to chaos. One route that is of particular interest to us is the period doubling route. In this route the dynamical system undergoes repeated period doublings i.e. a period 1 limit cycle changes character to a period 2 limit cycle and so on to reach chaos [9]. The time series of the Rössler system in period-1 and period-2 states are shown in Figure 1.7. A dynamical system may also show other routes to chaos such as the Ruelle-Takens-Newhouse scenario that approaches chaos via a quasiperiodic state and the intermittency route [5]. We will now define two quantifiers to describe two aspects of the state space of chaotic systems.

Lyapunov exponent

The Lyapunov exponent characterizes the rate of separation between adjacent trajectories in state space. A rudimentary definition of the Lyapunov exponent can be achieved by considering with two trajectories that are initially separated by a distance d_0. Lets call the initial time t_0. The separation as a function of time, t, goes as

$$d(t) = d_0 e^{\lambda t} \tag{1.13}$$

λ is called the Lyapunov exponent. In general for an n dimensional system we have n such Lyapunov exponents. If one of the $\lambda > 0$, we have exponential divergence of nearby

points, which leads to unpredictability that is typical of chaotic states[1]. All λs that are ≤ 0 correspond to periodic or fixed point states. The sum of all the Lyapunov exponents give the expansion or contraction of the state space. Hence dissipative systems would have a negative sum [10]. A fixed point has all Lyapunov exponents negative. A limit cycle has one zero Lyapunov exponent and the remaining negative. A quasiperiodic attractor has two zero Lyapunov exponents and one negative and a chaotic attractor has one positive, one negative and one zero Lyapunov exponent.

Fractal dimension

The complex nature of the chaotic trajectories in state space can be captured geometrically through fractal dimensions. We will first talk about the simplest kind of dimension, the box counting dimension . To understand the box counting dimension, let us consider an object that needs two variables to describe it (we will stop short of calling this object 2 dimensional in order to avoid confusion between the topological and fractal dimensions). Now, we cover this object with a grid of squares of size R_0, and count the number of occupied squares, $n(R_0)$. As we reduce the size of the squares, and count the number of occupied squares each time, the number of squares changes as

$$n(R) = AR^{d_0} \tag{1.14}$$

d_0 is the box counting dimension. The concept is easily extended to higher dimensions. One easily sees that for a solid object in 2 or 3 dimensions, the box counting dimension reduces to the usual idea of dimension. For the complicated geometry of a chaotic system, d_0 will not in general be an integer [5, 11]. Hence it is called a fractal.

The two quantifiers described above characterize two of the primary characteristics of chaotic systems. The exponential divergence of nearby trajectories in state space is characterized by a positive Lyapunov exponent . We illustrate this divergence in Figure 1.8. If the system under consideration is dissipative, the motion is restricted to a set with zero n-volume(volume generalized to n-dimensions, where n is the topological dimension of the system) [5]. This set is called a strange attractor, because of its fractal nature. The fractal structure of this state space is characterized using fractal dimensions like the box counting dimension. The trajectories shown in Figures 1.1 and 1.2 are examples of chaotic and strange attractors.

Strange non chaotic behavior

Another interesting dynamical state that occurs in deterministic nonlinear dynamical systems is the strange non chaotic state. Strange non chaotic trajectories are also bounded irregular trajectories, similar to chaotic trajectories. The attracting set also has fractal geometry. However nearby trajectories do not show exponential divergence, which is characteristic of chaos. This combination of strangeness (fractal geometry) and lack of exponential divergence is why this dynamical state is called "strange non chaotic". The combination of the two factors give rise to a non integer fractal dimension and a zero largest Lyapunov exponent. The doubly driven pendulum given in equation 1.12 shows strange non chaotic dynamics for appropriate values of parameters. The time series and state space structure for this system in the strange non chaotic state is shown in Figure 1.9.

[1]Systems with more than one positive Lyapunov exponent are collectively called hyperchaotic systems.

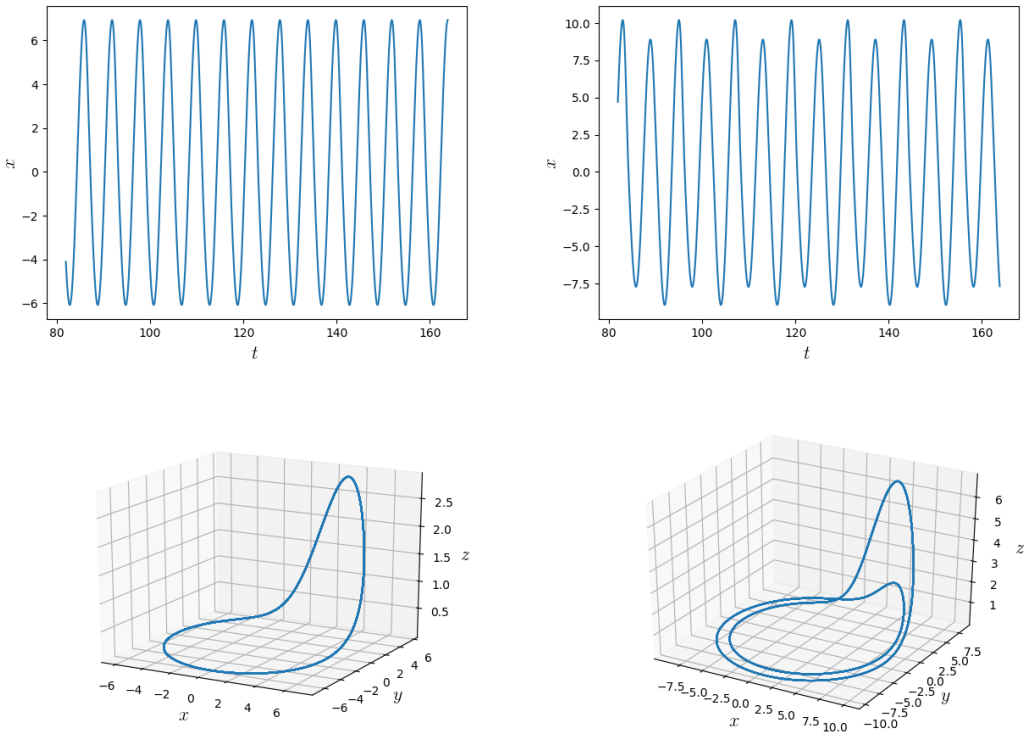


Figure 1.7: Time series for the Rössler system in the period-1 limit cycle and period-2 limit cycle. The lower panels show the corresponding state space plots.

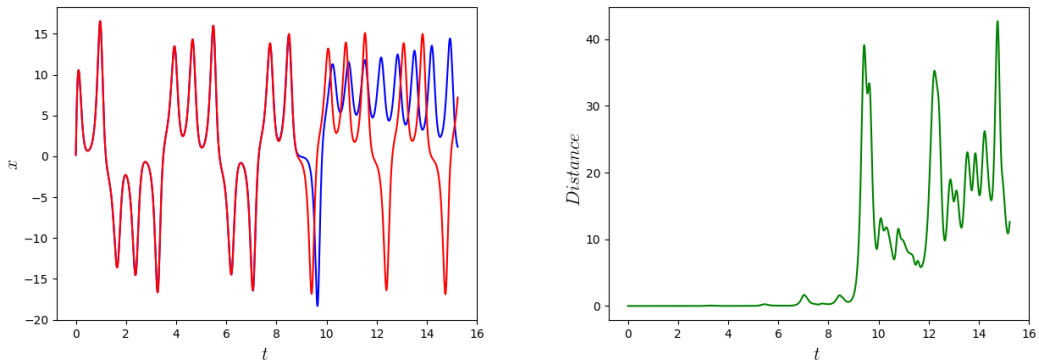


Figure 1.8: Divergence of nearby trajectories in the Lorenz system. The trajectories were initially separated by .001 in the x variable, and evolved. The second graph shows the distance between the trajectories at each time step.

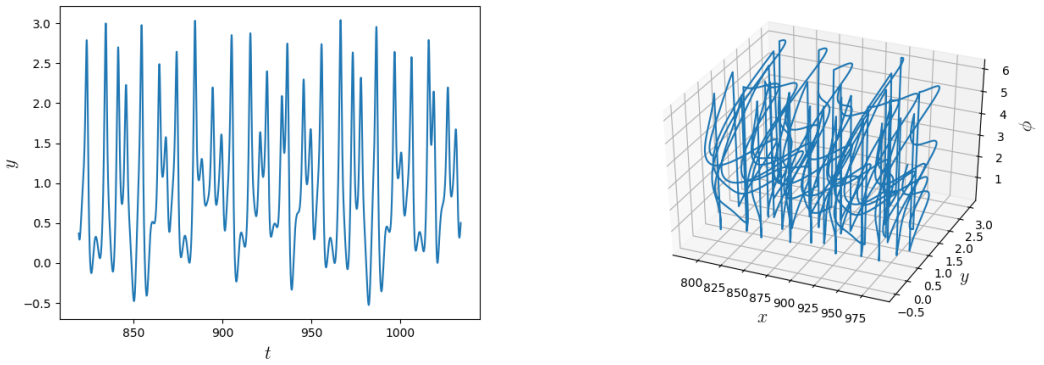


Figure 1.9: Doubly driven pendulum showing strange non chaotic behavior for parameter values $K = 1.33, V = 0.55, p = 3.0, \omega_1 = \frac{\sqrt{5}-1}{2}$ and $\omega_2 = 1.0$

1.2 Time Series Analysis

In this section we introduce the tools of time series analysis we use along the course of this thesis. The different tools we develop, operate on time series and give us information about them. This would help us understand the nature of the underlying dynamics from which the time series is derived, or relationships between multiple time series or so on. For instance it may give clues about whether the underlying system is stochastic or deterministic, linear or nonlinear, nature of the dynamical state of the system etc.

1.2.1 Time Domain

In this subsection we deal with time series analysis tools that act directly in the time domain. Essentially, the tools we discuss in this section, act directly on the time series, and not on a mapping of it. The preliminary statistical properties are the mean(μ), variance(σ^2) and standard deviation(σ) of a time series $x(t)$, that are defined as

$$\begin{aligned} \mu = E[x] &= \frac{1}{N}\sum_{i=1}^{N} x_i \\ \sigma^2 &= \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu)^2 \\ \sigma &= \sqrt{\sigma^2} \end{aligned} \tag{1.15}$$

These would be useful as we proceed to define correlations.

Autocorrelation function

The autocorrelation function finds the extent of correlation between a time series and a lagged version of itself [12]. It helps find the periodicities hidden in a time series. For a time series x_n, the autocorrelation function, for a lag τ is defined as [13]

$$C(\tau) = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i - \mu)(x_{i+\tau} - \mu)}{\sigma^2} \tag{1.16}$$

Here μ and σ^2 are as defined in Equation 1.15. Stochastic processes are known to have decaying autocorrelation functions in time. Time series from chaotic systems too are

thought to decay exponentially but much slower than noisy systems. However, this is not a conclusive signature, and cannot differentiate between chaotic and random signals by themselves [12]. The autocorrelation function for a white noise signal and the x variable of the Lorenz attractor is shown in Figure1.10

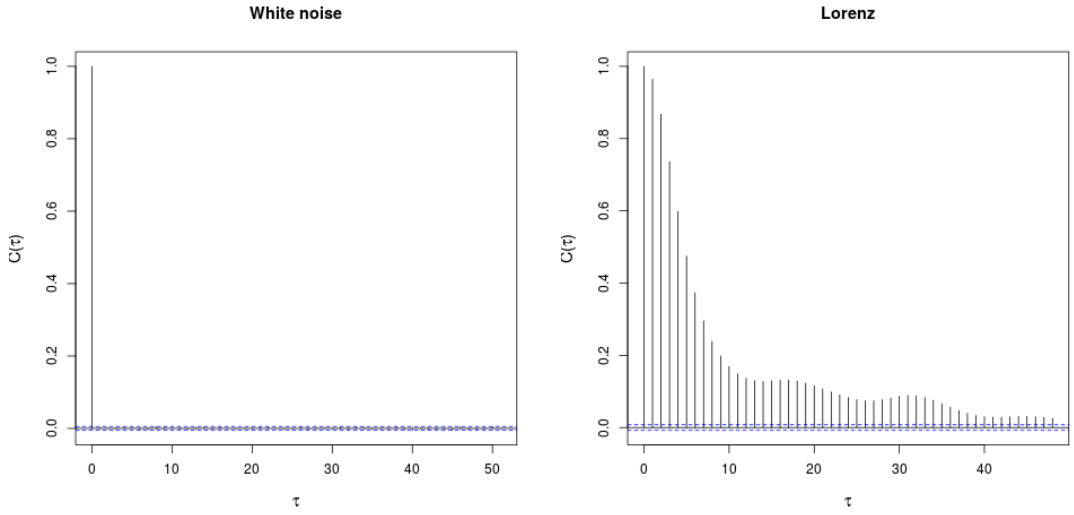


Figure 1.10: Autocorrelation function for a white noise signal and for the x variable of the chaotic Lorenz system. Unlike for white noise, the correlation falls off more slowly in the Lorenz system.

Crosscorrelation function

A cross correlation function is an extension of the autocorrelation function, where we consider the extent of linear correlation between one time series and a lagged version of another. If x_n and y_n are time series derived from two processes, the cross correlation for a delay τ is given as,

$$C_{xy}(\tau) = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i - \mu_x)(y_{i+\tau} - \mu_y)}{\sigma_x \sigma_y} \tag{1.17}$$

Here the subscripts refer to the statistical measure corresponding to the time series x and y. We show the cross correlation between two white noise processes and between the x and y variables of the Lorenz systems in Figure 1.11

Triple correlation function

The autocorrelation function explores linear correlations in a time series. It gives no information about the nonlinear correlations. Higher order correlations give us this information. The simplest instance of this is the triple autocorrelation function, which checks for correlations in a time series at two delays, τ_1 and τ_2 [14, 15].

$$T(\tau_1, \tau_2) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)(x_{i+\tau_1} - \mu)(x_{i+\tau_2} - \mu) \tag{1.18}$$

The triple autocorrelation function has the following symmetry [16]

$$T(\tau_1, \tau_2) = T(\tau_2, \tau_1) = T(\tau_1 - \tau_2, -\tau_2) = T(\tau_2 - \tau_1, -\tau_1) = T(-\tau_2, \tau_1 - \tau_2). \tag{1.19}$$

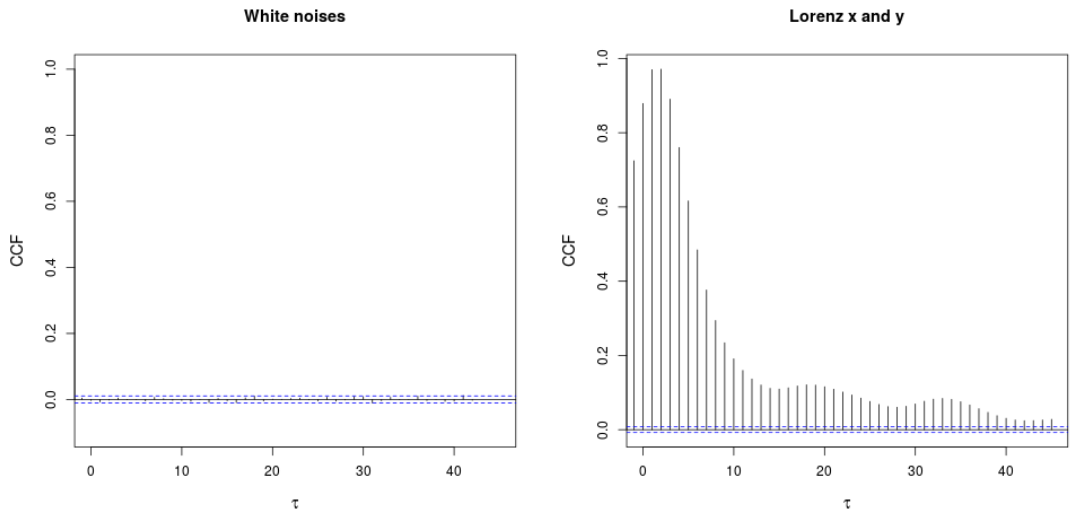


Figure 1.11: Cross correlation function between two white noise signals and between the x and y variables of the chaotic Lorenz system. We observe a lack of correlation between the white noise signals, while the correlation between the x and y variables of the Lorenz system falls off much slower.

1.2.2 Frequency Domain

It is often more instructive to quantify the time series in frequency space instead of time space. This transformation is achieved using the Fourier transform. The Fourier transform is a specific transform belonging to the family to integral transforms, generally described by

$$f'(\beta) = \int_a^b f(\alpha)K(\alpha,\beta)d\alpha \tag{1.20}$$

where $K(\alpha,\beta)$ is called the kernel. For the Fourier transform the kernel function is given by $e^{i\alpha\beta}$. A time series $x(t)$ can be expressed in frequency space $x'(f)$ using the Fourier transform and the reverse is achieved using the inverse transform [17]

$$\begin{aligned} x'(f) &= \int_{-\infty}^{\infty} e^{2\pi ift}x(t)dt \\ x(t) &= \int_{-\infty}^{\infty} e^{-2\pi ift}x(f)df \end{aligned} \tag{1.21}$$

The adaptation of the Fourier transform to deal with discretely sampled time series is the discrete Fourier transform (DFT). Before defining the discrete Fourier transform, we need to consider the information that is preserved after sampling a continuous process discretely. If Δt is the sampling time, the time series is bandwidth limited to frequencies upto $f_c = \frac{1}{2\Delta t}$, by a theorem called the sampling theorem. The frequency, f_c is called the Nyquist critical frequency [13].The power associated with any frequency component lying outside the $(-f_c, f_c)$ range is transferred to this range, through a phenomenon called aliasing. Further, since the time span of the time series is finite, the function that is being Fourier transformed is the product $R(t)x(t)$, where $R(t)$ is the rectangle function defined as,

$$R(t) = \begin{cases} 0 & t < 0 \\ 1 & 0 < t < t_0 \\ 0 & t > t_0 \end{cases} \tag{1.22}$$

Here, t_0 is the length of the time series. The Fourier transform of a product of functions is given by the convolution of the Fourier transforms.

$$\mathcal{F}(R.x) = \mathcal{F}(R) * \mathcal{F}(x) \tag{1.23}$$

where

$$R * x = \int_{-\infty}^{\infty} R(\tau)x(t-\tau)d\tau \tag{1.24}$$

The Fourier transform of $R(t)$ gives rise to a sinc function[2], giving rise to spurious power appearing in the transform.

For a time series of length N, discretely sampled at time step Δt, the frequencies will be sampled as

$$f_i = \frac{i}{N\Delta t}, i = -\frac{N}{2}, ..., \frac{N}{2} \tag{1.25}$$

The discrete Fourier transform(DFT) and inverse transform are then given by

$$\begin{aligned} X'_n &= \Delta t \sum_{k=0}^{N-1} x_k e^{2\pi ikn}/N \\ x_k &= \frac{1}{N} \sum_{k=0}^{N-1} X_n e^{-2\pi ikn}/N \end{aligned} \tag{1.26}$$

An important theorem called Parseval's theorem states that the total power in both spaces is preserved. In discrete space this is written as

$$\sum_{k=0}^{N-1} |x'_k|^2 = \frac{1}{N} \sum_{n=0}^{N-1} |H_n|^2 \tag{1.27}$$

A remarkable speed up can be obtained computationally[3], using the fast Fourier transform (FFT) algorithm proposed by Cooler and Tukey. It is a divide and conquer algorithm that reduces the problem to finding the DFT of the time series split into smaller segments. A detailed description is given in [13].

Power spectrum

The power spectrum can be roughly defined as the modulus square of the discrete Fourier transform. It gives the amount of power between frequency f_i and f_{i+1}(Equation 1.25). If the function x_k is real, $X'_{-n} = X'^*_n$, where the $*$ denotes conjugate. If A_k is the FFT of the time series at f_k, one can then write the appropriately normalized power spectrum as

$$\begin{aligned} P(0) = P(f_0) &= \frac{1}{N^2}|A_0|^2 \\ P(f_k) = \frac{1}{N^2}[|C_k|^2 + |C_{N-k}|^2], \quad k &= 1, 2, ..., (\frac{N}{2} - 1) \\ P(f_c) &= P(f_{\frac{N}{2}})|C_{\frac{N}{2}}|^2 \end{aligned} \tag{1.28}$$

Time series with continuous spectra, either arising from chaos or stochasticity, tend to have a large variance associated with the calculated power spectrum. One popular way

[2] $sinc(x) = \frac{sin(\pi x)}{\pi x}$

[3] An $O(N^2)$ process can be reduced to a $O(Nlog_2(N))$ process.

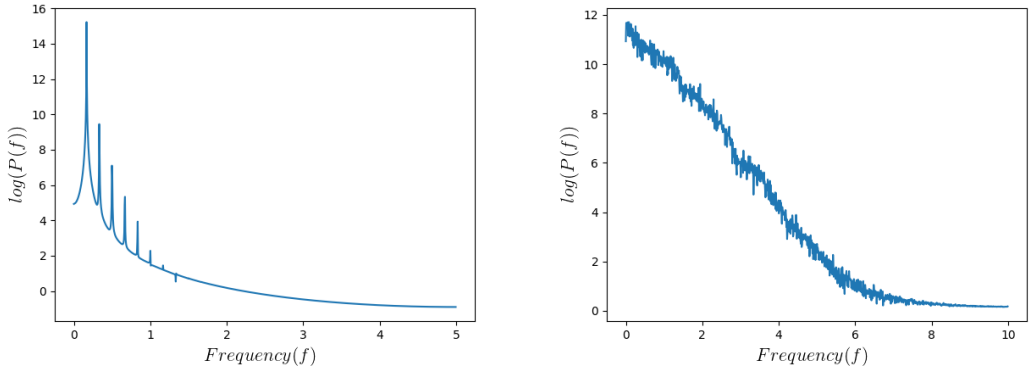


Figure 1.12: Power spectra for the Rössler system in the limit cycle regime and the chaotic Lorenz system. The power spectrum of the limit cycle shows peaks only at the primary and its harmonics, whereas the chaotic power spectrum shows continuous power at all the frequencies.

to reduce the error on the power spectrum, which we will use throughout power spectrum estimation in this work, is through Bartlett's method of averaged periodograms [18]. In this method, individual periodograms calculated from k independent segments of the time series are averaged to find a smoothed estimate of the power spectrum. This can be shown to reduce the variance of the power spectrum by a factor $\frac{1}{k}$.
The power spectrum is related to the autocorrelation. The power spectrum is the Fourier transform of the non-normalised autocorrelation. The theorem that relates the two is called the Wiener Khinchin theorem.

$$A(f) = \sum_{\tau=-\infty}^{\infty} c(\tau) e^{-2\pi f \tau} \tag{1.29}$$

The power spectrum for the Rössler system in the limit cycle regime and the Lorenz system in the chaotic regime is shown in 1.12.

Noise

Before moving forward to higher order spectra, we define what we mean by noise. In section 1.1 we considered the dynamics when the system is governed by a set of deterministic differential equations. In this section we consider the case when the dynamics is essentially random. A dataset that is gathered from random dynamics too would show irregular behavior. One of the fundamental questions of nonlinear time series analysis is to differentiate between the two kinds of irregularity. An even more challenging problem is noise contamination. Observations from real world systems inevitably come with noise contamination, even when the system is completely deterministic as we considered in section 1.1. Noise may be present in the measurement or during the evolution of the system. The noise may also be additive or multiplicative. Understanding the nature of the contaminating noise and the its effects on time series quantifiers is hence an important step in the modelling process.

White Noise

The simplest example of a noisy process is a white noise process. This is a stationary stochastic process where adjacent data points are independent of each other. In terms of the power spectrum, which gives the distribution of power across various frequencies

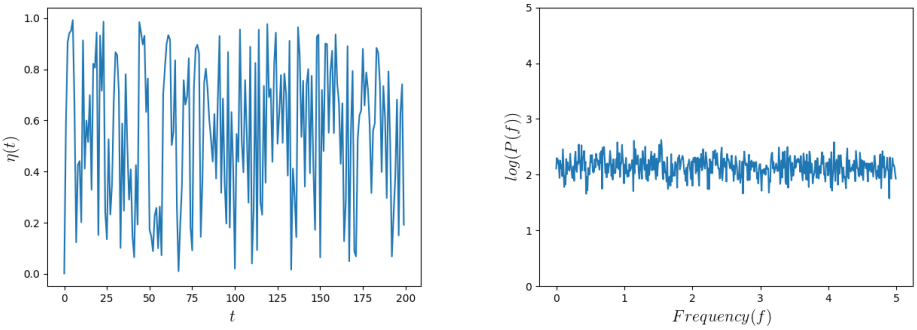


Figure 1.13: Time series and power spectrum of a pure white noise process. The power spectrum of a white noise process shows equal power at all frequencies.

in a signal, a white noise process has equal power at every frequency. If the amplitude distribution of the white noise is Gaussian, it is termed Gaussian white noise. White noise is present in many different natural systems. An example is the Johnson-Nyquist noise that is present in electronic circuits. The time series and power spectrum of a white noise process is shown in Figure 1.13.

Colored Noise

Correlations in stochastic processes give rise to colored noise. Depending on the shape of the power spectrum the color of the noise is called pink, red etc. The power spectra of colored noise processes all have a power law behavior in their power spectrum. In general if the slope of the power spectrum on a log-log graph is (a) -2, the noise is called red noise (b) -1 is called pink noise and (c) 0 is called white noise. Colored noise, especially with spectral indices between -1 and -2, is ubiquitous in nature [19–21]. Many attempts have been made to explain the reason for ubiquity with some success [22–24]. We will describe red noise in some detail.

A red or Brownian noise or a Wiener process is a random process whose power spectrum falls with an index -2. It can be obtained as the integral of a white noise process. It is a non stationary[4] Gaussian Markov process and can be expressed as

$$\dot{x} = \eta(t) \tag{1.30}$$

where $\eta(t)$ is a white Gaussian white noise process. The Wiener process is the continuum limit of a random walk [25]. Red noise processes appear as models in a large number of fields [26–29]. We plot an instance of a red noise process and its power spectrum in Figure 1.14.

Cross spectral density

The cross spectral density is defined as the Fourier transform of the non normalised cross correlation function. In terms of Fourier components we can define it as

$$S_{xy}(f) = A_x(f)B_y^*(f) \tag{1.31}$$

where A_x and B_y^* are the Fourier transforms of $x(t)$ and $y(t)$ respectively. Essentially the power spectral density then becomes a special case of the cross spectral density when $x(t)$ and $y(t)$ are the same.

[4]For a process to be stationary the probability distributions of its variables must remain time independent.

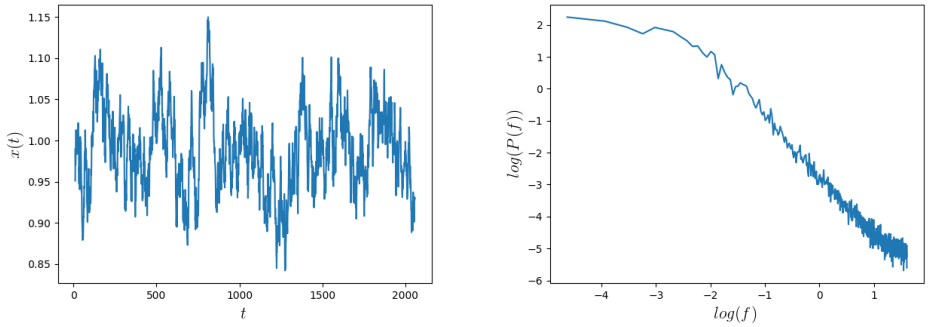


Figure 1.14: Time series and power spectrum of a red noise process. The power spectrum shows a power law decay with an exponent ≈ -2.

Bicoherence

The bispectrum is the Fourier transform of the three point correlation function defined in Equation 1.18. It can be estimated in terms of the Fourier transform of a time series as

$$B(f_1, f_2) = \sum_{i=1}^{k} A_i(f_1) A_i(f_2) A_i^*(f_1 + f_2) \tag{1.32}$$

Here, $A(f)$ is the Fourier transform at f and k is the number of segments. The magnitude of the bispectrum is called the bimagnitude and the phase is called the biphase. The bicoherence is a normalised version of the bimagnitude defined as

$$b(f_1, f_2) = \frac{|\sum_{i=1}^{k} A_i(f_1) A_i(f_2) A_i^*(f_1 + f_2)|}{\sum_{i=1}^{k} |A_i(f_1) A_i(f_2) A_i^*(f_1 + f_2)|} \tag{1.33}$$

One can immediately see that the sum over k segments is similar to a random walk in $2 - D$, which falls as $\sqrt{\frac{1}{N}}$ [16]. The bicoherence finds the extent of quadratic coupling between different frequencies in the time series. The significance levels for various values of bicoherence can be obtained using χ^2 distributions. For 99% significance, using k segments of time series, the required value for bicoherence is $\sqrt{\frac{9.2}{2k}}$ [30]. The bicoherence for the Rössler system in limit cycle and chaotic regimes is plotted in Figure 1.15. Only the region defined by

$$\begin{aligned} f_2 &> 0 \\ f_1 &\geq f_2 \\ f_1 + f_2 &\leq f_{max} \end{aligned} \tag{1.34}$$

is plotted. This is because the bicoherence function is completely defined in this region due to its inherent symmetries.

Tricoherence

The trispectrum is defined as the Fourier transform of the 3^{rd} order cumulant [16]. In terms of the Fourier transform, we can write it as

$$T(f_1, f_2, f_3) = \sum_{i=1}^{k} A_i(f_1) A_i(f_2) A_i(f_3) A_i^*(f_1 + f_2 + f_3) \tag{1.35}$$

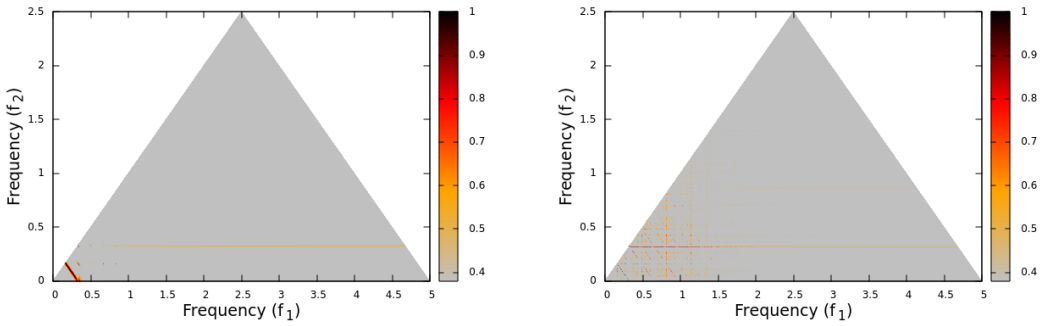


Figure 1.15: Full bicoherence plots for the x time series of the Rössler system in the limit cycle and chaotic regimes.

The tricoherence is given as,

$$t(f_1, f_2, f_3) = \frac{|\sum_{i=1}^{k} A_i(f_1)A_i(f_2)A_i(f_3)A_i^*(f_1+f_2+f_3)|}{\sum_{i=1}^{k} |A_i(f_1)A_i(f_2)A_i(f_3)A_i^*(f_1+f_2+f_3)|} \tag{1.36}$$

The tricoherence gives the extent of cubic coupling between frequencies in the time series. The statistics of significance for the tricoherence can be shown to be similar to the bicoherence function [31]. A table of Fourier transform pairs in time and frequency domain is shown in Table 1.1

Table 1.1: Quantifiers in the time and frequency domains.

Time domain	Frequency domain
Auto-correlation	Power spectrum
Cross-correlation	Cross spectral density
Triple correlation function	Bispectrum
4^{th} order cumulant	Trispectrum

1.2.3 State Space Quantifiers

Upto now we talked about methods of time series analysis that works on either time series data directly or on their Fourier transforms. This section will deal with quantifiers that will quantify the state space directly. Though it seems that state space quantification would need time series from all the variables of the system, a remarkable theorem by Floris Takens showed that you can reconstruct the dynamics of the entire state space using just one of the variables. The method employed is called the method of delay embedding. Before we describe the method, we briefly state the statement of the theorem, which in turn is an extension of Whitneys embedding theorem.

Taken's Embedding Theorem: *Let M be a compact manifold of dimension m. For pairs (ϕ, y), where $\phi : M \longrightarrow M$ is a smooth diffeomorphism*[5] *and $y : M \longrightarrow \mathbb{R}$ a smooth function, it is a generic property that the $(2m+1)$-delay observation map $\Phi_{(\phi,y)}$:*

[5]A diffeomorphism is a bijective differentiable map between two manifolds, which also has a differentiable inverse.

$M \longrightarrow \mathbb{R}^{2m+1}$ *given by* $\Phi_{(\phi,y)}(x) = (y(x), y \circ \phi(x), ..., y \circ \phi^{2m}(x))$ *is an embedding* [32,33]. Essentially this tells us that given a time series of one of the variables of a dynamical

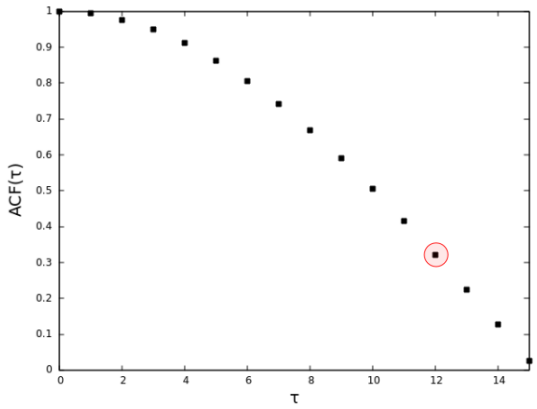


Figure 1.16: Autocorrelation function as a function of τ of the Rössler system. The circled point shows the point where the autocorrelation falls to $\frac{1}{e}$

system,$x_1, x_2..., x_n$, one can construct a delay vector in m dimensions as

$$\vec{v} = (x_i, x_{i+\tau}, x_{i+2\tau}, ..., x_{i+(m-1)\tau}) \tag{1.37}$$

A space of these vectors will be topologically equivalent to the original state space of the dynamical system, provided $m > 2d_e$ [12]. The first step in this process is the estimation of the delay time τ. In practice τ has to be defined such that the components of the reconstructed vectors are independent of each other. We resort to the autocorrelation function to quantify this. The delay time is chosen as the time where the autocorrelation function falls to $\frac{1}{e}$(Figure 1.16). To illustrate this, we show the state space of the Rössler system and the state space reconstructed from the dynamics of its x variable in Figure 1.17. When the time delay used is too small, the components end up being highly correlated and lying along the diagonal, whereas when the delay is too large, the components have no relation to each other and end up being space filling, like white noise. We illustrate these two scenarios in Figure 1.18.

The Lyapunov exponent is one of the quantifiers that can be determined from this reconstructed space. Many popular algorithms exist to calculate Lyapunov exponents

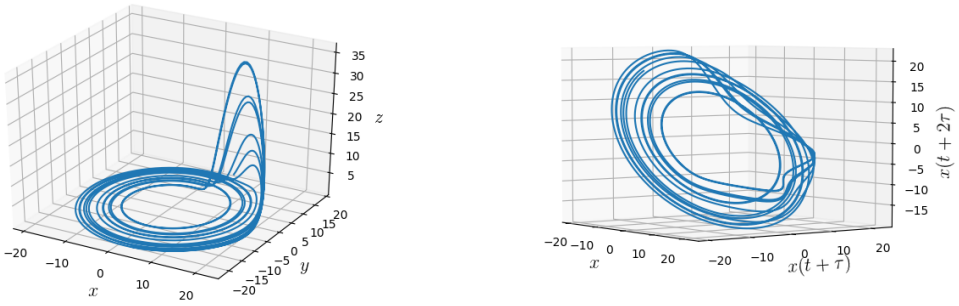


Figure 1.17: State space of the Rössler system without reconstruction and after reconstruction using the x variable time series.

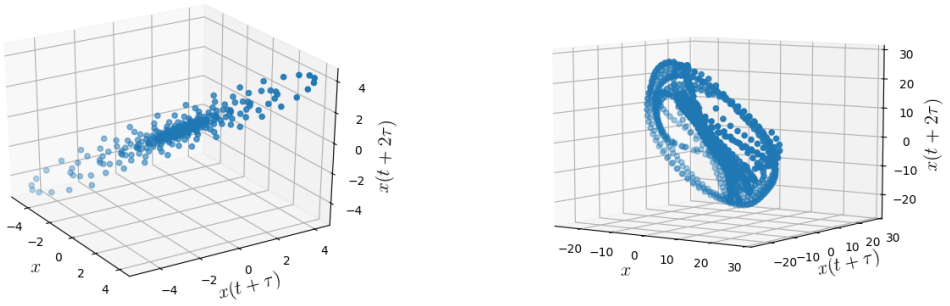


Figure 1.18: State space of the Rössler system after reconstruction with time delays that are (a) much smaller than $\frac{1}{e}$ and (b)much larger than $\frac{1}{e}$, using the x variable time series.

from time series, like the Wolf algorithm [34]. However the Lyapunov exponent is very susceptible to issues like noise which makes it a bad quantifier when dealing with real time series [35, 36]. Further, one may find spurious instances of positive Lyapunov exponents, if the trajectories are not followed for long enough, especially in the vicinity of a saddle point [5]. For these reasons, we do not rely on Lyapunov exponents for the analysis that will be considered in subsequent chapters.

Correlation dimension

One of the features of the state space structure of dissipative chaotic dynamical systems is its fractal nature or strangeness. One finds that the state space attractors of chaotic systems are fractals with a non integer box counting dimension. The boxcounting dimension described in section 1.1.1 is one of the fractal dimensions that can be calculated.

The correlation dimension is an alternative measure of the fractal dimension, which we will use extensively in the following chapters. Hence we describe how to calculate it algorithmically from a time series. Let us first consider a scalar time series x_i, sampled at every Δt. As described above we reconstruct the state space of the underlying dynamical system, using the method of delay embedding. This gives us a series of N vectors as described in 1.37. Lets now take every vector, i, in this space, and find the relative number of vectors with a distance R from this vector. If N_i is the number of points within a distance R,

$$p_i(R) = N_i/N \tag{1.38}$$

The correlation sum, $C(R)$ is then defined as

$$C(R) = \frac{1}{N}\sum_{i=1}^{N} p_i(R) \tag{1.39}$$

The correlation dimension D_2 is then related to $C(R)$ as [5]

$$C(R) = \lim_{R\to 0} kR^{D_2} \tag{1.40}$$

Taking log we have

$$D_2 = \lim_{R\to 0} \frac{logC(R)}{logR} \tag{1.41}$$

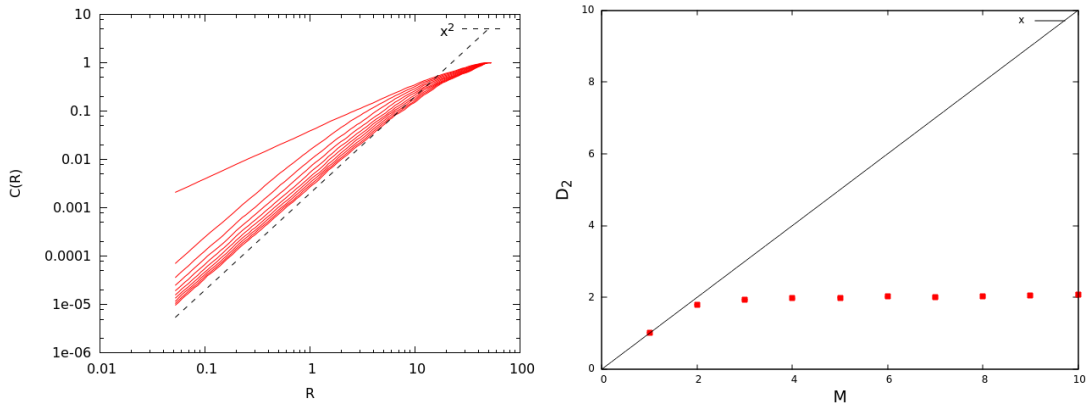


Figure 1.19: $C(R)$ vs R and D_2 vs M plots for the chaotic Rössler system. The actual value of D_2 is 1.99 ± 0.08 [37].

The subscript 2 is usually used for correlation dimension since it is one of the generalized dimensions D_q, given in the next section, with $q = 2$ An important point to notice here is that for low dimensional chaos, D_2(or any other fractal dimension) will saturate as we increase M to beyond the dimension of the system. This is different from say, white noise, which will continue to grow as we embed it in larger dimensions. Graphs of $C(R)$ vs R for different M, and D_2 vs M is shown in Figure 1.19. In our work, we make use of a slightly modified version of this algorithm as described in [37]. Initially, the amplitude distribution of the time series is converted to a uniform distribution. This transformation stretches the reconstructed space without any change to any of the dynamical invariants. Multiple advantages of using this transformation in the context of different nonlinear quantifiers is described in [37–39]. Another important change in the algorithm is the use of the maximum norm instead of the Euclidean norm. This means we find the number of vectors that lie in a box of size R instead of a sphere of radius R, in equation 1.38 and 1.39. This change to the maximum norm makes it convenient to account for edge effects correctly. We do this by imposing that the M dimensional cube has to lie within the embedding space.

Multifractal spectrum

Our discussion of fractal dimensions in the previous subsection considered an average over all state space points. In general, different parts of the attractor may contribute differently to this average. Hence, the full complexity is not captured by using just one of the measures of the fractal dimension we considered. We hence introduce the idea of generalized dimensions to characterize the complexity of the attractor, which is related to the idea of moments in statistics which is used to characterize a probability distribution[6].

Let us cover the state space attractor again with boxes of size R and define the relative number of points in the i^{th} cell as $p_i = N_i/N$. We can then define the generalized dimension as

$$D_q = \lim_{R \to 0} \frac{1}{q-1} \frac{ln \sum_{i=1}^{N} N(R) p_i^q}{lnR} \tag{1.42}$$

We could also alternatively define a generalized correlation sum, analogous to the corre-

[6]The q^{th} moment of a distribution is given as, $M_q = \frac{1}{N} \sum_{i=1}^{N} (x_i - \langle x \rangle)^q$

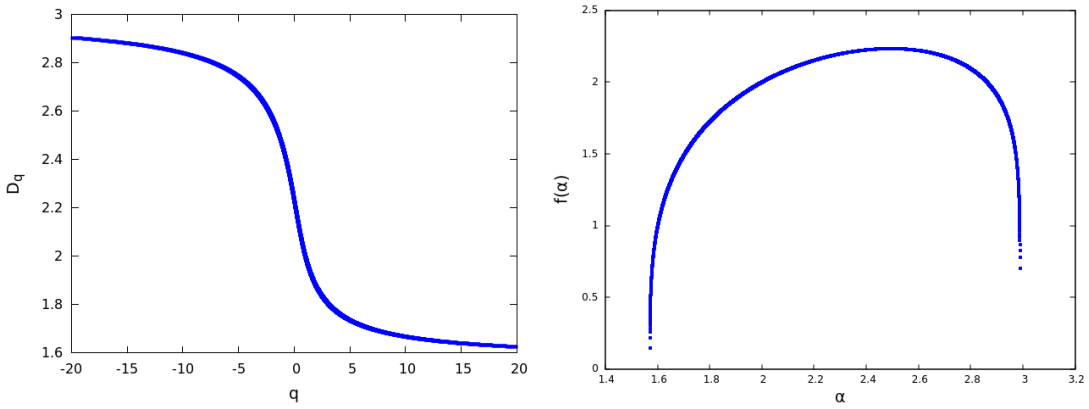


Figure 1.20: D_q vs q and $f(\alpha)$ vs α plots for the chaotic Lorenz system.

lation dimension, as [5]

$$C_q(R) = \sum_{j=1}^{N} p_j^q \tag{1.43}$$

This gives the generalized dimension as

$$D_q = \lim_{R \to 0} \frac{1}{q-1} \frac{lnC_q(R)}{lnR} \tag{1.44}$$

The dimensions we considered, like the box counting dimension, D_0 and the correlation dimension, D_2, are specific cases of the generalized dimension for the values of $q = 0, 2$. There is an analogous formulation, where we consider how p_i changes with R. p_i is assumed to scale as

$$p_i(R) = R^{\alpha_i(R)} \tag{1.45}$$

We then counts the number of boxes with α between α and $\alpha + \Delta\alpha$ [40]. This is related to the size of the box R as,

$$n(\alpha, R) \propto R^{-f(\alpha)} \tag{1.46}$$

D_q and $f(\alpha)$ are related through Legendre transformations.

$$(q-1)D_q = q\alpha - f(\alpha) \tag{1.47}$$

The D_q vs q and $f(\alpha)$ vs α plots for the Lorenz system are shown in Figure 1.20.

The $f(\alpha)$ curve can be characterized using the following function fit [41].

$$f(\alpha) = A(\alpha - \alpha_{min})^{\gamma_1}(\alpha_{max} - \alpha)^{\gamma_2} \tag{1.48}$$

This quantification using α_{min}, α_{max}, γ_1 and γ_2 of the $f(\alpha)$ curve has been put to extensive use to multiple fields like black hole characterization, ECG characterization etc [42, 43].

1.2.4 Recurrence Analysis

Recurrence plots

A recurrence plot is a concept that helps us quantify the recurrences and underlying patterns in a state space trajectory. This method basically takes a structure embedded in M dimensional space, and converts it into a binary matrix, by quantifying recurrences

in the structure. In our case this structure is an attractor in state space. First a point in the attractor i is chosen, and all points with a distance ϵ is found. If a point j lies within ϵ distance of i, the matrix element (i, j) is set to 1. Otherwise it is set to 0. That is

$$R_{i,j} = \Theta(\epsilon - ||\vec{x_i} - \vec{x_j}||) \tag{1.49}$$

Here Θ is the Heaviside step function. The recurrence plot is the visual representation of this matrix.

Once we have the recurrence plot, a variety of quantifiers are used to quantify the patterns it shows. This study of recurrence plot patterns is called recurrence quantification analysis(RQA). We will mainly be looking at the recurrence rate, determinism and laminarity. The three measures we choose is dependent on three different aspects of the recurrence plot, namely the recurrence density, the diagonal structures and the vertical structures.

Recurrence rate: This is the simplest measure derived from the recurrence plot. In the limit of large N it describes the probability of recurrence of a state to its ϵ neighborhood. It's given as

$$RR(\epsilon) = \sum_{i,j=1}^{N} R_{i,j}(\epsilon) \tag{1.50}$$

Determinism: This is linked to the diagonal structures on the recurrence plot. We can define the distribution of diagonal structures $P(\epsilon, l)$ as

$$P(\epsilon, l) = \sum_{i,j=1}^{N} (1 - R_{i-1,j-1}(\epsilon))(1 - R_{i+l,j+l}(\epsilon)) \prod_{k=0}^{l-1} R_{i+k,j+k}(\epsilon) \tag{1.51}$$

Weakly correlated processes cause very small or no diagonal structures whereas deterministic correlated processes cause long diagonal structures. The determinism is then defined as the ratio of recurrence points that form diagonal structures of minimum length l_{min} to the total number of recurrence points(for constant ϵ).

$$DET = \frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)} \tag{1.52}$$

Laminarity: The vertical structures in a recurrence plot is often associated with the tangential motion of the state space trajectory. Though this is not always the case. We can find the distribution of vertical structures $P'(v)$ as

$$P(v) = \sum_{i,j=1}^{N} (1 - R_{i,j})(1 - R_{i,j+v}) \prod_{k=0}^{v-1} R_{i,j+k} \tag{1.53}$$

Then the Laminarity is defined as [44]

$$LAM = \frac{\sum_{v=v_{min}}^{N} vP(v)}{\sum_{v=1}^{N} vP(v)} \tag{1.54}$$

The recurrence plots for the Rössler system in the chaotic regime and white noise in $M = 3$ is shown in Figure 1.21.

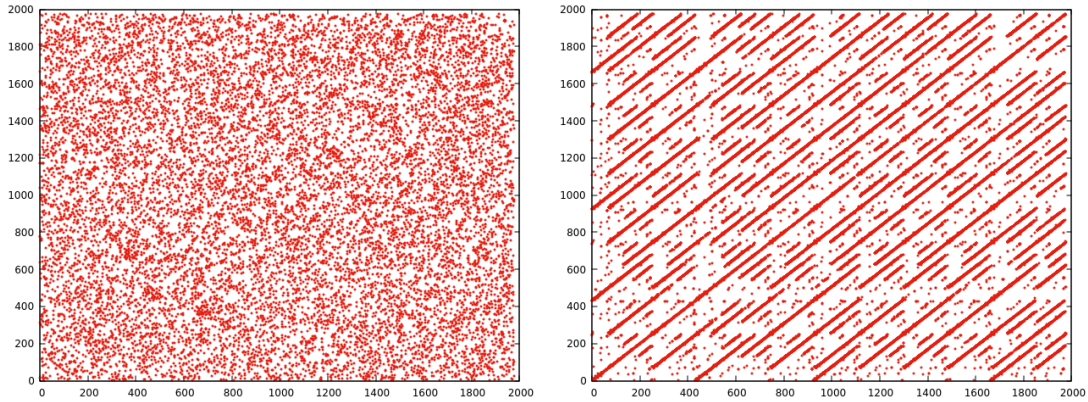


Figure 1.21: Recurrence plots for a white noise signal and the Rössler system in the chaotic regime.

Recurrence networks

Complex networks have been one of the most intriguing additions to the field of complex systems in recent times. The tools of complex networks have been put to use in a wide variety of fields, including divisibility patterns in natural numbers, heteronormativity in society and predicting complexes in protein interaction networks [45–47]. One of the major developments in nonlinear time series analysis in the last two decades has been the use of complex networks to study the properties of time series. This is achieved using transformations that convert a time series into a network. Many algorithms exist to do this, the popular among which are the visibility algorithm [48], neighbourhood algorithm [49], coarse graining method [50], recurrence networks [51] etc. See [52] for a detailed review on the different approaches. In this section and throughout this thesis, we will concentrate mostly on the last method, the ϵ-recurrence networks. In this method, we use the binary matrix given by equation 1.49 with the diagonals replaced by 0s, as the adjacency matrix to generate a network. One of the primary questions that arises is the choice of ϵ or recurrence threshold that should be used to generate the network. We follow the method laid out in [39] for the work presented in this thesis. In this method, ϵ is chosen based on two main conditions. First, the quantifiers of the recurrence network are kept as being distinctly different from the network constructed from noise. Secondly, 95% of the nodes fall into a single component. Keeping these two criteria, ranges of ϵs can be constructed for systems of different embedding dimensions [39]. Once the recurrence networks are constructed, we can use network quantifiers like the average degree, average path length and clustering coefficients to characterize the networks. Example recurrence networks constructed from the Rössler system in the chaotic regime and from white noise is shown in Figure 1.22.

Average degree: Once we have the recurrence network, we can plot the distribution of degrees of all the individual nodes. The average of this distribution is the average degree of the node. If N is the total number of nodes and $A_{i,j}$ is the adjacency matrix, the average degree is given by [53]

$$d_{avg} = \frac{1}{2N} \sum_{i=1}^{N} A_{i,j} \tag{1.55}$$

Average path length: Another important measure of the network is the efficiency of information transport on it. The average or characteristic path length provides the

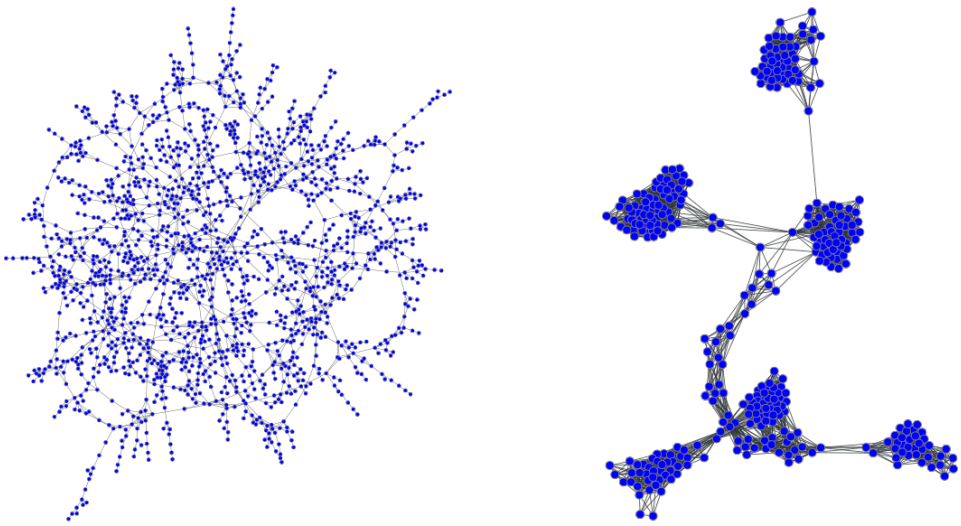


Figure 1.22: Recurrence networks for a white noise signal and the Rössler system in the chaotic regime.

average number of steps required to traverse from one part of the network to another.

$$C_{PL} = \frac{1}{N \cdot (N-1)} \cdot \sum_{i \neq j} d(v_i, v_j) \tag{1.56}$$

Here v_i and v_j are nodes in the network and $d(v_i, v_j)$ is the shortest distance between them [53].

Clustering coefficient: The clustering among nodes in the network is another important measure. It is given by the clustering coefficient. The global clustering coefficient measures the fraction of closed triplets in the network.

$$CC_{gl} = \frac{3 \times \text{number of triangles}}{\text{number of all triplets}} \tag{1.57}$$

The other measure of clustering defines a local clustering for each node in the network. It measures how close a node is to being a clique or a complete graph. Let k_i be the degree of a node v_i and let $e_{i,j}$ be the link between two nodes v_i and v_j. If N is the set of all nodes and E is the set of all links, the local clustering for a node i is given by

$$C_i = \frac{2|\{e_{jk} : v_j, v_k \in N_i, e_{jk} \in E\}|}{(k_i)(k_i - 1)} \tag{1.58}$$

The average clustering for the whole network is [53]

$$CC_{avg} = \frac{1}{n} \sum_{i=1}^{n} C_i \tag{1.59}$$

One of the main advantages of the CPL and CC has been the ability of the CPL-CC plane to distinguish between deterministic and noisy dynamics [54]. We show how the values of the CPL and CC of the Lorenz system lies in a distinct part of the CPL-CC plane, away from the red and white noise in Figure 1.23.

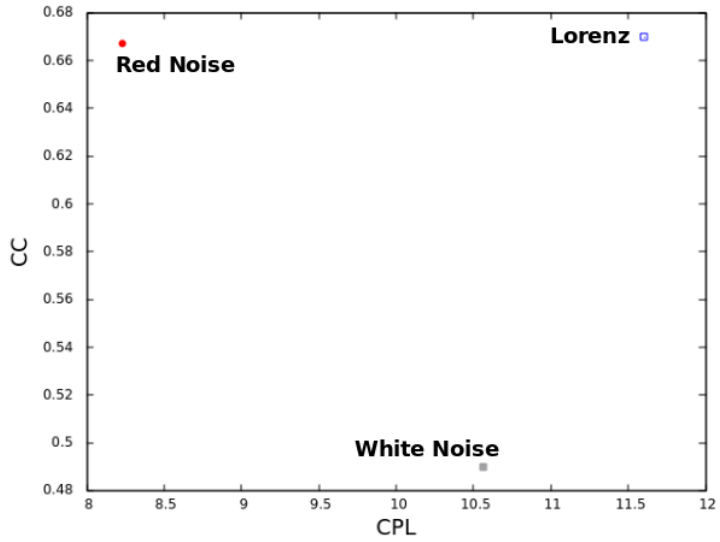


Figure 1.23: CPL-CC plane for showing positions of white noise, red noise and the Lorenz system. We find that the deterministic Lorenz system occupies a position away from white and red noise in the CPL CC plane.

1.3 Summary and Discussion

Now we are in a position to present the studies on the techniques of time series analysis to understand the dynamics of real world systems. In the next chapter, we will proceed to introduce the physics of variable stars. We will then go on to address the important problem of datagaps in real world time series, and analyze how they affect the quantifiers we considered in this chapter. In chapter 4 we consider the problem of noise and how the bicoherence function in various forms can help tackle this question and recover the underlying dynamical state. We also analyze the nonlinear dynamics of RR Lyrae stars. In chapter 5 we use these quantifiers to study the dynamics of contact binary stars. In chapter 6, we use recurrence based analysis to classify RRc Lyrae stars into two subgroups. We also use recurrence networks to differentiate between different kinds of close binaries. We present the summary of our results and scope for future work in the concluding chapter.

Chapter 2

Introduction: Stellar Variability

Since the dawn of civilisation, stars have fascinated the human race. Multiple legends and myths have been spun around them, with many stars and constellations of stars being worshipped as Gods and demigods. They were probably the first objects to be methodically observed across civilizations. Over time the fascination with stars and systems of stars and systems of systems of stars has not diminished, but has possibly increased. Stars are now known to be laboratories where various branches of physics from quantum mechanics to general relativity interplay with each other to reveal thoroughly fascinating phenomena!

One of the earliest observations made about stars has been the light variability in some of them. The earliest known record of stars whose light varied over time was of Algol, an eclipsing binary noted as exhibiting light variability by the ancient Egyptians. In this chapter we will study sources of variability in stars and the nature of their variation. This will help us, as we proceed to subsequent chapters and explore the nonlinear dynamics of variable star light curves.

2.1 Basics

Before we proceed further, we will define two main terms that will be fundamental to understanding stellar physics. The first among these is the effective temperature, T_{eff}. To define the effective temperature, lets recollect the form of Plank's law , given by

$$B_{\lambda b} = \frac{2hc^2}{\lambda^5} \frac{1}{exp(hc/\lambda kT) - 1} d\lambda \tag{2.1}$$

The T_{eff} of a source is defined as the temperature of a black body which would radiate the same amount of radiation as the source under consideration [55]. The second of these is the luminosity, which is defined as the total amount of energy emitted by the star in unit time. The knowledge of the T_{eff} and the radius of the star completely determines its luminosity [56]. One way to measure the luminosity is using the bolometric or absolute magnitude. To define this, we initially define the apparent magnitude, which is the brightness of a star as measured from the earth. The apparent magnitude of the star Vega at any specified wavelength, is defined as the zero magnitude. The apparent magnitudes of other stars are compared in relation to Vega. The apparent magnitude uses a logarithmic scale and is negative for objects brighter than Vega and positive for objects that are dimmer. The absolute magnitude is defined as the apparent magnitude

an object would possess if placed at a distance of 10 parsecs or 32.6 light-years[1] [57]. The difference in bolometric magnitude (total energy across wavelengths) is related to the bolometric luminosity ratio as

$$M_* - M_\odot = -2.5 log \frac{L_*}{L_\odot} \tag{2.2}$$

where $\odot$ represents the value of the measure for the sun and $*$ is the value of the measure for the object in question.

2.1.1 The H-R Diagram

The Hertzsprung-Russel diagram is a plot of luminosity versus effective temperature for all the observable stars. An example of this H-R diagram, for some sample stars from the Hipparcos catalog, is plotted in Figure 2.1 [58]. Immediately we see that there is a strip along which most of the stars lie[2]. This strip is called the main sequence of stars. The hotter and heavier stars live towards the top left of the main sequence, whereas the cooler and lighter stars live on the bottom right. The stars that are off the main sequence form what are called the horizontal branches and the instability strip. Many of the variable stars, we will talk about fall into this instability strip. This diagram forms the central tenet of stellar astrophysics, and gives insight regarding the structure and evolution of stars. Alternative to luminosity and T_{eff} we can also use the absolute magnitude versus the color or spectral type to plot another variation of the HR diagram. The spectral type is related to T_{eff} and arranges stars (primarily) based on the strength of their Hydrogen absorption lines[3] [59].

Curves on the HR diagram that tracks how stars develop over time are called evolutionary tracks[4]. A star starts its life by giving out light by burning hydrogen in its core to give helium. These are the stars that lie on the main sequence. Once the core runs out of hydrogen, a star moves horizontally out of the main sequence and stars growing in size and decreasing in effective temperature. These stars start burning helium in their cores and hydrogen in a shell around the core. Along the HR diagram, one finds a narrow region or strip of stars which are unstable. This strip is called the instability strip. All the stars that fall into this strip are pulsating stars. The region where this strip intersects the horizontal branch is populated by RR Lyrae stars. Super giant stars lie above the main sequence, again in an almost horizontal line. The instability strip intersects this super giant strip to give rise to a region populated by Cepheid variables.

One of the other important characteristics of a star is its metallicity. It measures the metal content in a star, and is often captured by the concentration of Fe to H in a star as compared to the sun.

$$[Fe/H] = log_{10}([Fe/H]_{star}) - log_{10}([Fe/H]_\odot) \tag{2.3}$$

Generally, based on the metallicity, stars are divided into Population I and Population II stars. Population I stars are younger stars which are formed in a metal rich environment

[1] Assuming no extinction of light

[2] In this diagram, the color temperature measured by B-V is used instead of T_{eff}. It is measured by finding the difference in magnitude when using a B filter ($\approx 440nm$) and a V filter ($\approx 550nm$) [57].

[3] In the order of decreasing T_{eff}, the spectral classes are ordered **OBAFGKM** that can be easily remembered by the non-sexist mnemonic "**O**h**B**e**AF**ierce**G**irl,**K**arate-kick**M**e".

[4] Pre main sequence stars follow a nearly vertical path called the Hayashi track, soon after which it follows a horizontal track called the Henyey track into the main sequence.

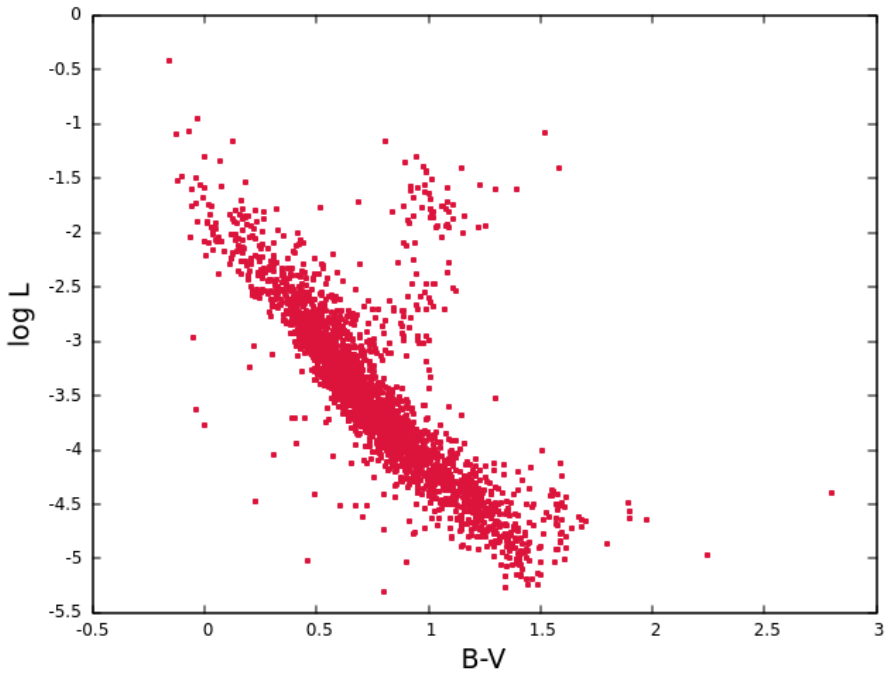


Figure 2.1: HR diagram of logL(in units of solar luminosity) vs. B-V for 2719 stars from the Hipparcos catalog [60].

and consequently have higher metallicity. Population II stars are older and were formed in a metal poor environment. Hence they have lower metal content [57].

2.2 Stellar Variability

Through the course of this thesis, we will explore the dynamics of stars that exhibit variability in light intensities. Broadly the variation may be divided into intrinsic and extrinsic variation. When the source of variation is in the star itself, like in pulsating stars, the star is called an intrinsic variable. Alternately the star may also vary in luminosity due to a external event like an eclipse or due to rotations. This results in a perceived variation of light as observed from the earth. Such variables are called extrinsic variables.

Based on the source of variation, the stars may be broadly divided into the following categories.

- **Pulsating stars** In these stars the light variation takes place due to the swelling and shrinking of the star itself. They fall into the instability strip on the H-R diagram, that we discussed in the last section. The intersection of the instability strip with (a)the super giant strip is populated by Cepheids and RV Tauri stars, (b)the horizontal branch is populated by RR Lyrae stars and (c)with the main sequence is populated by δ Scuti star.

- **Rotating stars** The light variation here is primarily due to rotations of the star. The changing faces of the star as it rotates gives differing amounts of light due to presence of starspots on the surface or due to a distortion of the stellar surface due to gravitation or rotation. An important example of the former are pulsars which are rotating neutron stars. Two examples of the latter are the ellipsoidal variables

which are close binary stars and FK Comae Berenices variables which are distorted due to rapid rotation of a single star [61].

- **Eruptive stars** These are variable stars where the variability is due to sudden bursts of energy and subsequent responses. These may be sudden short lived events like supernovae or flares. They may also be a long term variation like mass transfer in cataclysmic stars. When the mass transfer is unstable it may lead to an eruption like a supernova or a coalescence like a red nova [62]. Another kind of eruptive variable is the R Coronae Borealis type stars, where fadings are thought to be due carbon condensation leading to a reduction in visible light [63].

- **Eclipsing stars** Eclipsing stars are stars where the primary source of light variation is due to an eclipsing phenomena, either another star or a planet. If the eclipsing object is a second star, we can have subdivisions based on how close the stars are to each other (detached, semi-detached and contact binary stars.) [64].

An important aspect to pay attention to is that these categories are not mutually exclusive. In fact semi detached binary stars would form an instance of cataclysmic stars. Variations in close binary stars is not solely because of eclipses alone. Hence while these classifications may be used for increasing one's understanding of variable stars, they by no means constitute mutually exclusive, independent categories.

2.3 Stellar pulsations

In this section we will consider stellar pulsations in greater detail. We mentioned earlier that the reason for stellar pulsations is the swelling and shrinking of the star. This results in light intensity variations.

One of the most popular theories for stellar pulsations is the Eddington valve mechanism or κ mechanism named after Arthur Eddington [65]. It takes place in the following steps.

- A layer of the star collapses against gravity and falls.
- This layer then compresses (heats up) and hence becomes more opaque to radiation
- The opacity leads to slower diffusion of radiation, which moves more slowly through the layer. This leads to a build up of heat beneath.
- This leads to an increase in pressure and causes the layer to be pushed
- The layer then expands and becomes more transparent to radiation, and the process repeats.

Stellar pulsations may lead to rather regular pulsations like in the case of Type-I Cepheids or rather irregular pulsations like in RV Tauri variables. We will explore now, in greater detail some specific kinds of pulsating stars.

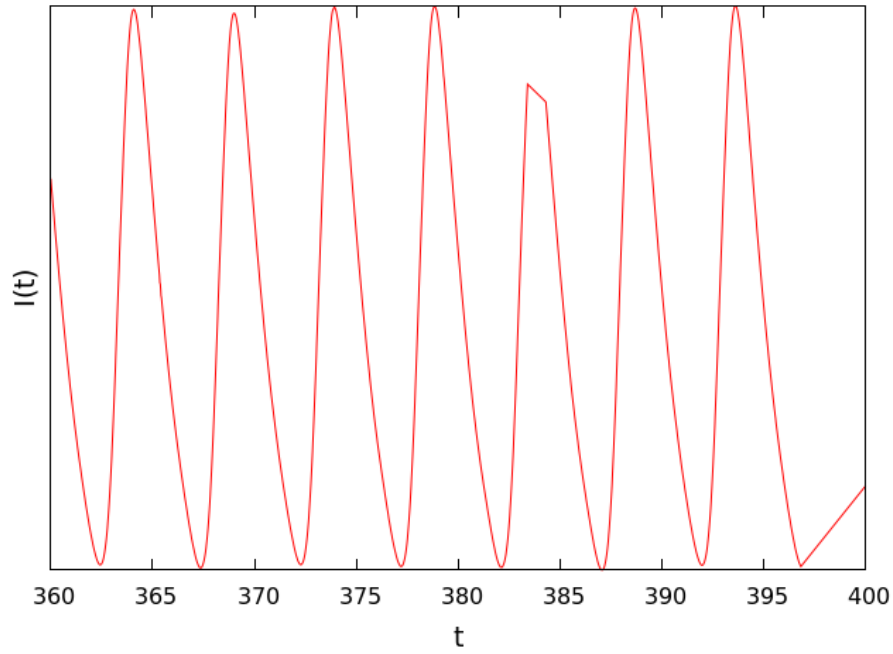


Figure 2.2: Section of the light curve of Cepheid variable V1154 Cyg (KIC 7548061). We see very regular variations in the light curve.

2.3.1 Cepheids

As we mentioned earlier, Cepheid variables are found at the intersection of the instability strip with the supergiants. Type I Cepheids undergo almost perfectly regular pulsations. One of the most important properties of Type I Cepheid variables is a period-luminosity relationship they show. This meant that a straight line could be drawn between the log of the period and the magnitude. This implied that the knowledge of the period would imply knowledge of the luminosity, which in turn implied that the distance to that Cepheid could be calculated exactly[5].

Type I Cepheids are young Population I giant stars. Type II Cepheids on the other hand, are older and less massive Population II stars. They are also called W Virginis stars after their prototype. Initially, the two populations of Cepheids were difficult to distinguish, and lead to a massive miscalibration of the period-luminosity relation. The pulsations of W Virginis stars are more irregular than type I Cepheids. Models of W Virginis stars have been analysed and seem to show period doubling chaos [66, 67]. The light variation of Type 1 Cepheid variable V1154 (KIC 7548061) is plotted in Figure 2.2. The x-axis is the modified Julian day number subtracted by a constant. The Julian day number is an integer that is assigned to every day starting from noon of 1 January 4713 BC. The Julian day number for a particular day is the number of days that has elapsed since 12 noon of 1 January 4713 BC. For artificial Earth satellites like the Kepler, time is expressed in Modified Julian Date Numbers. In this case the zero point is 17 November 1858 [55]. Hence, Modified Julian Date = Julian Date - 2,400,000Âů5 days.

[5]This relationship was first discovered by Henrietta Leavitt in the early 1900-s, and became one of the most important results in understanding distances in the universe. Leavitt observed that among the Cepheids in the large Magellanic cloud (which were all roughly at the same distance), the brightest Cepheids had the longest periods.

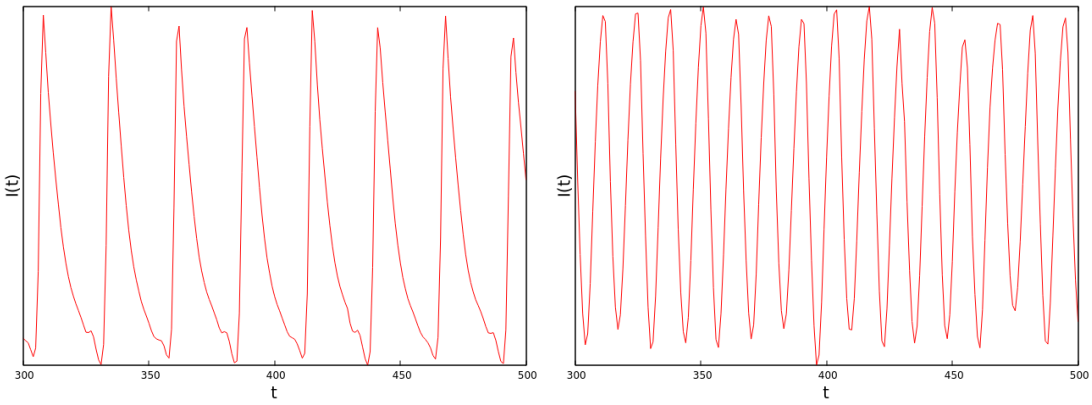


Figure 2.3: Section of the light curves for (a)RRab type variable KIC 4484128 and (b)RRc type variable KIC 5520878.

2.3.2 RR Lyrae

RR Lyrae stars are another important class of pulsating variable stars which lie in the intersection between the instability strip and the horizontal branch. These stars have been long thought to be perfectly periodic pulsating stars and have been used as standard candles to measure distances. They have pulsation periods that range from 0.1 to 1 day, and fall in the spectral types A5 to F5. RR Lyrae stars have been subject to modelling using dynamical equations [68, 69]. Hydrodynamic simulations of RR Lyrae stars have suggested that they may undergo period doubled chaos [70].

One of the most curious phenomena in RR Lyrae light curves is a modulation that can be observed in them. This modulation has been a paradox that has not been completely explained to this day and is called the Blazhko effect. Recent work suggests that the Blazhko effect may be a result of the 9^{th} overtone that destabilizes the fundamental mode (in a 9:2 resonance) [71]. Alternate hypothesis include a cyclical weakening of the turbulent convection in the hydrogen and helium ionization zones [72].

RR Lyrae stars are subdivided into RRabcd Lyrae stars depending on the mode in which they pulsate. RRab Lyrae stars pulsate in the radial fundamental mode, whereas the RRc Lyrae stars pulsate in the first overtone. RRd variables pulsate simultaneously in both [73]. Further, RRab stars have slightly longer periods than RRc Lyrae variables. RRab stars have pulsation periods that range between 0.3 to 1 days. RRab Lyrae stars were thought be be perfectly periodic until the advent of the Kepler space telescope. One of the first results regarding these stars was the discovery of the period doubling phenomena which was missed until the advent of Kepler [74]. This discovery of period doubling coupled with the model based predictions of chaotic behavior in these stars suggests that these stars may be undergoing richer dynamics. We explore this further in chapter 4.

RRc Lyrae stars are overtone pulsators with periods between 0.2 and 0.5 days. One of the main observations about RRc Lyrae variables is the presence of a secondary pulsation mode that is about 0.6 times the fundamental in many of them [75]. The fact that this ratio is very close to the golden ratio has evoked considerable excitement about them. In fact the nonlinear analysis of RRc Lyrae variable stars showed that they may be exhibiting strange non chaotic behavior [76]. We show sections of light curves from RRab and RRc Lyrae stars in Figure 2.3.

2.4 Eclipsing binaries

We mentioned earlier that eclipsing binary stars are stars whose primary variation is due to an eclipse. Binary stars are thought to form more than 60% of stellar systems in the universe. When the components are inclined such that one passes in front of the other from our line of sight, we call them eclipsing binary stars.

Observationally eclipsing binaries were classified according to certain prototype stars. Broadly they are classified into Algol type, β Lyrae stars and W UMa type stars. These are also called EA, EB and EW type variables. Irrespective of whether the stars eclipse each other they can be morphologically classified depending on how close the components are to each other. In order to do so, we first define the concept of Roche lobes , which is the area around the star where the orbiting material is gravitationally bound. A schematic for the same is shown in Figure 2.4. If both the stars, like in Figure2.4a, does not fill their Roche lobes, they are called detached binaries. If only one star does, as depicted in the Figure2.4b, it is called a semi detached binary. If both stars fill their Roche lobes as in Figure2.4c, we have what is called a contact binary star. The latter two classes of stars are together called close binary stars, since there is a possibility of mass and/or energy exchange in them. Algol type or EA stars are the semi detached type. WUMa or EW stars are the overcontact type stars. EB type stars form what are called ellipsoidal type of stars, which can be detached, semi detached or contact binary stars that are tidally distorted [64].

The evolution of binary stars forms a very active branch of stellar astrophysics. One of the popular theories for binary star evolution assumes that these stars start off as detached binary stars. Due to angular momentum loss because of stellar winds combined with the evolution of the stars, one of the stars fills its Roche lobe. This gives rise to mass loss from the bigger companion to the smaller one through the Lagrangian point, L_1. This mass flow in turn could accelerate the evolution of the companion, which could grow to a common envelope encompassing both stars [77, 78].

This mass and energy transfer that is characteristic of close binary stars leads to light variation that does not stem from the eclipse. These effects have been long observed. When the light curves show unequal maxima it is known as the O'Connell effect and when the predicted time of a minimum does not match with the expected time, it is known as an eclipse time variation [79, 80]. Overcontact binary stars are especially interesting in this context, as mass transfer and loss of angular momentum in these stars are thought be be responsible for period changes. This change in the period leads to an eventual coalescence in the form of a red nova outburst [62]. Hunts for possible merger candidates have been of special interest in the last few years [81, 82].

We conclude this subsection by mentioning two important quantifiers that help quantify overcontact binary stars. These are the fill-out factor (ff) and mass ratio (q). The former quantifies the extent of contact between the components while the latter measures the ratio of masses. The fill out factor is given by the formula [83],

$$ff = \frac{\Omega^I - \Omega}{\Omega^I - \Omega^O} \tag{2.4}$$

where Ω is the potential at the surface of the common envelope, Ω^I is the potential at the inner Lagrangian surface and Ω^O is the potential at the outer Lagrangian surface.

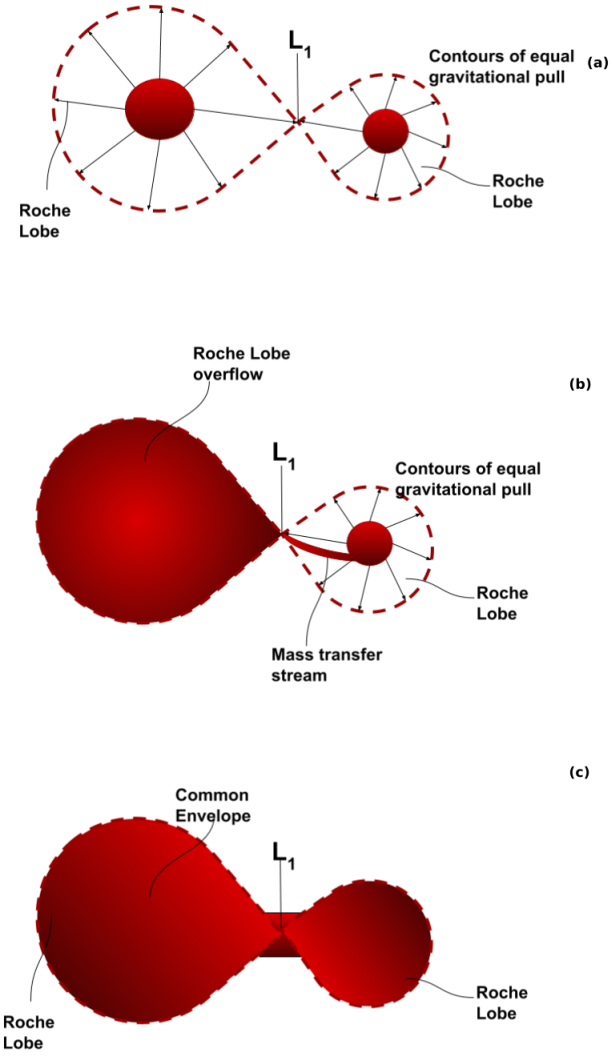


Figure 2.4: Schemes of binary star for (a)detached (b)semi detached and (c)contact binary stars.(Created using Google drawings.)

2.5 Summary and Discussion

In this chapter we present a basic understanding on the origins of variability in stars. Often, more than one source is involved in light variation in variable stars. Broadly we divide variability into intrinsic and extrinsic variables, based on whether the source of the variation is internal or external to the source[6]. We discuss specifically the case of pulsating stars and eclipsing stars which we will analyse extensively in the subsequent chapters.

We also study eclipsing binary stars where the primary source of variation is thought to be an eclipse of one star by the other. However, when the component stars are close to each other, a possibility of mass and energy exchange exists. This results in variation of light in the light curve, from sources other than stellar eclipses. Depending on how close the stars are to each other they may be detached, semi detached or contact binary stars. The variation from strict periodicic behavior is studied generally as small changes to the periodicity and are called the O'Connell effect or eclipse time variations.

Astrophysics has historically been the cradle of dynamical systems theory. Studies into the three body problem was probably the first chaotic problem to be addressed. Many important tools of dynamical systems theory like Poincare sections were first developed in the context of the three body problem. However as the area of nonlinear dynamics grew into an independent field with extensive applications, the use of the tools of nonlinear dynamics to answer questions in astrophysics became significantly rarer. Part of the reason has been the lack of high quality data that is needed for nonlinear time series analysis that is often absent in astrophysics. Two areas where chaos theory has been very successful in addressing questions posed by astrophysics has been in the theory of stellar pulsations, and the theory of matter accretion. The former has been significantly linked with nonlinear dynamics with dynamical systems models existing from the 1970-s. In this chapter we saw the use of similar models to predict chaos in RR Lyrae stars. Even earlier, nonlinear time series analysis seemed to suggest low dimensional chaotic behavior in RV Tauri variables like R Scuti [84]. The second situation where nonlinear time series analysis showed promise has been in the study of accretion phenomena. A striking example of this has been in the black hole system GRS 1915+105, which was shown to switch between random and chaotic states [85, 86]. Similar analysis for the neutron star system, Sco X-1, showed low dimensional chaotic behavior [87]. However exploration for chaos using nonlinear time series analysis of the light curve of the AGN source W2R 1946+42 was shown to be inconclusive [88].

In this thesis, we will attempt to address some of the problems we face in dealing with astrophysical data like the presence of datagaps and noise, and the lack of long datasets. We will also attempt to explore some areas of astrophysics like close binary physics where the tools of nonlinear time series analysis has not been used before.

[6]Often an extrinsic variable may have secondary sources of variability as in the case of close binary stars, or an intrinsic source may have an extrinsic component as is the case for a planetary eclipse for a pulsating star. Hence these classes are broad and do not represent independent groups.

Chapter 3

Effect of Datagaps on Nonlinear Measures

The techniques of time series analysis are mostly developed to deal with observations that are continuous and evenly spaced in time. In actual practice however, observations are missed intermittently for a variety of reasons many of which are difficult to avoid. Dealing with missing data or unevenly sampled data is a real challenge in data analysis. What is required is a detailed analysis on how datagaps affect time series quantifiers.

In this chapter we will address this question in the context of dealing with nonlinear time series analysis. We will consider many of the time series analysis measures we have discussed in chapter 1 and analyse their resilience to datagaps. This will help us gain an understanding of how reliable time series analysis measures are in the presence of datagaps.

3.1 Datagaps

Datagaps are ubiquitous in observational time series. They appear due to number of reasons like instrument failure, imperfect sensors, environmental factors, human errors etc. Datagaps are present in time series across fields including time series of packet traffic data, astrophysical time series, ecological time series, radars and so on [89].

Gaps in data or missing data is a specific kind of uneven sampling of data. Data is said to be unevenly sampled when the time series points are not evenly spaced in time. This can happen for instance when there is no underlying sampling time in the data. This kind of uneven sampling takes place for instance in time series of medical or psychological self report studies, variable stars reported by amateur astronomers, paleoclimate data etc.[1]

Techniques like the Lomb-Scargle periodogram, have been developed to compute power spectra and auto correlations for time series that are unevenly sampled [90, 91]. However, the most usual technique that is employed to deal with datagaps and uneven sampling in general is interpolation. Specifically, when the time series is affected by datagaps, the gaps are interpolated over. One major drawback of this method in the

[1] An example to illustrate the difference between datagaps and pure uneven sampling would be two ways of recording temperatures over a day. First consider that the temperature was measured at every hour from 0800 hrs to 2000 hrs, but two observations were missed. Second consider that the temperature was measured at 10 points in time, randomly between 0800 to 2000 hrs. The former would be an example of datagaps with a distinct sampling time of one hour, while the second is an example of pure uneven sampling with no obvious underlying sampling time.

context of nonlinear time series quantifiers is spurious signatures of nonlinearity that the smoothing can generate. Historically this has lead to controversy regarding detection of a low dimensional attractor in climate data [92–94]. In the following section we will aim to quantify the effect that datagaps have on nonlinear time series quantifiers in the absence of interpolation.

3.2 Synthetic Data with Gaps

In this section we will describe the process of how we quantify the effect datagaps have on different nonlinear measures. We initially start from a time series of an evenly sampled standard nonlinear dynamical system. For this, we choose the Rössler and Lorenz systems described in chapter 1. To ensure that the attractor is covered uniformly, the Rössler system is sampled at 0.1 and the Lorenz system at 0.01. Gaps are then introduced in these evenly sampled time series. The size of gaps and frequency of gaps are drawn from two Gaussian distributions, independent of each other[2]. Since the causes for gaps are thought to arise from multiple independent sources, it is reasonable to assume a Gaussian distribution for them. The Gaussian distributions from which the gaps are derived from are given by

$$G_S(s; m_s, \omega_s) = \frac{1}{\omega_s\sqrt{2\pi}} e^{-(s-m_s)^2/2\omega_s^2} \tag{3.1}$$

$$G_P(p; m_p, \omega_p) = \frac{1}{\omega_p\sqrt{2\pi}} e^{-(p-m_p)^2/2\omega_p^2} \tag{3.2}$$

Here G_S is the Gaussian distribution of gap sizes with mean m_s and standard deviation ω_s. Instead of frequency of gaps we consider a related quantifier, termed gap position, modelled by the Gaussian G_P. The mean and standard deviation of the Gaussian distribution are given by m_p and ω_p, respectively. The gap position essentially gives a measure of the gapless regions in between two gaps.

Once the gaps are introduced, we merge the data across the gaps. Hence we are left with a time series that has a finite number of discontinuities. The time series of the Rössler system after introducing gaps and after merging is shown in Figure 3.1. The algorithm we develop has the following steps.

1. Generate a large evenly sampled time series.
2. Calculate the quantifier of interest, say Q for the time series. Lets call this Q^{es}
3. Introduce gaps into the time series where the gap size and frequency are drawn from two Gaussian distributions, $G_S(s; m_s, \omega_s)$ and $G_P(p; m_p, \omega_p)$
4. Ignore the gaps and merge the time series
5. Calculate Q for the time series with gaps, say Q^{gaps}.
6. Quantify the deviation of Q^{gaps} from Q^{es}

[2]This independence can be thought to be reasonable. For instance how frequently a telescope fails is likely to be independent of how frequently it fails.

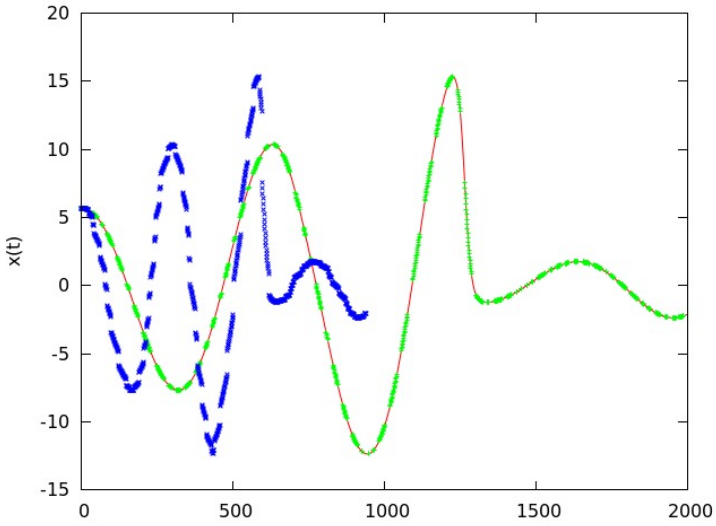


Figure 3.1: The time series of the Rössler system before introducing gaps (full curve), after introducing gaps (green) and after merging the data (blue).

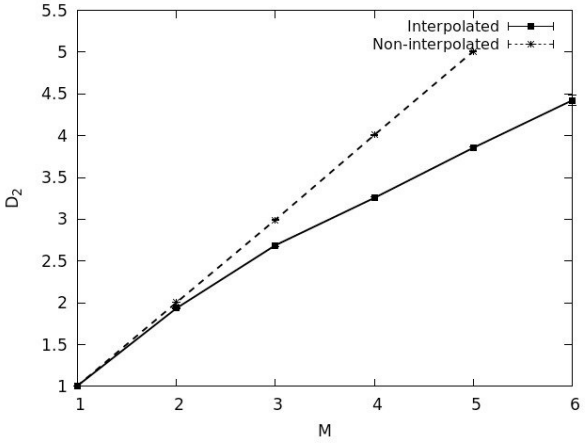


Figure 3.2: D_2 vs M plot for white noise with and without interpolation. The white noise without interpolation shows no saturation, while in the presence of interpolation, the D_2 vs M plot tends to saturate, mimicking nonlinear and chaotic behavior.

We first illustrate the dangers of interpolation in the context of a white noise time series that has been interpolated through. We compare the correlation dimensions of the interpolated and non interpolated time series, and show that the interpolated time series seems to show an artificial saturation in the D_2 vs M plot. The initial time series is a Gaussian white noise process with 20,000 points. We introduce gaps, with an average size (m_s) varying between 3 and 350 data points, every 300 data points($m_p = 300$ time steps) on an average. These gaps are interpolated through using a cubic spline. We then calculate $D_2(M)$ for dataset. The D_2 values do not saturate when interpolation is absent. When the gap size is large, the D_2 values for interpolated datasets show tendencies for saturation, hence suggesting underlying nonlinear dynamics. This is illustrated in Figure 3.2.

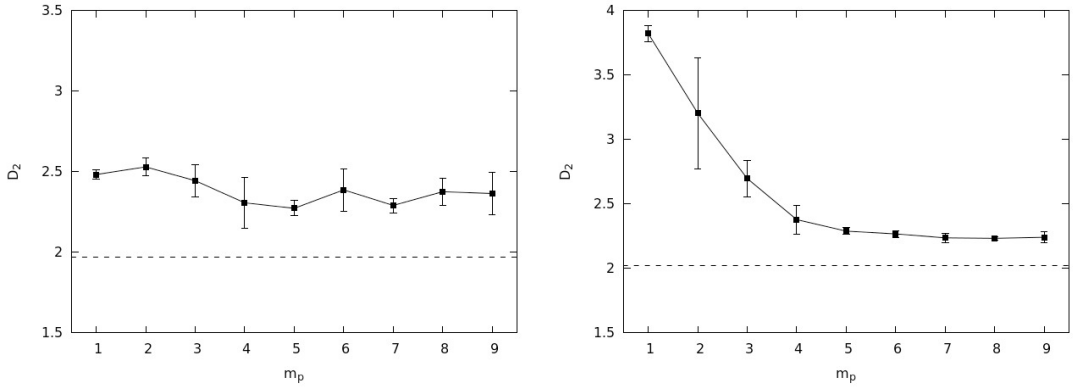


Figure 3.3: Variation of saturated D_2 with increasing mean position, m_p for the Lorenz (left) and Rössler (right) systems. m_s if fixed at 2τ. The error bars are standard deviations over five realizations.

3.3 Effect on time series quantifiers

In this section we will quantify the effect of missing data on a variety of nonlinear quantifiers. Specifically we will consider the effect it has on D_2 and the $f(\alpha)$ spectrum.

3.3.1 Correlation dimension

We introduced the correlation dimension as one of the measures of the fractal nature of the state space of a dynamical system. We will systematically examine the effect datagaps have on the calculation of this quantifier. For this analysis, we ensure that the datasets have $\approx$ 20,000 points after introduction of gaps. The natural time scale in the problem is the delay time used for reconstructing the state space of the system, τ. We vary m_s and m_p in terms of τ, since it remains invariant with the sampling time of the system. It is worth noting that the autocorrelation time τ, will fall faster in the presence of gaps. However for the range of gap sizes and frequencies we've considered here, the variation in τ remains small. The standard deviations, ωs, are fixed at 0.1 times the mean. We will conduct a systematic study of the effects of standard deviation later in this section.

We first look at the variation of D_2 with changing gap position, m_p. We consider the data from Lorenz and the Rössler system. The gap position m_p is increased from 1τ to 10τ. We observe that the D_2 falls as the m_p is increased. This is to be expected, as increasing m_p results in less frequent gaps in data. Figure 3.3 shows the change in D_2 with changing m_p.

The variation with m_s is slightly more complex. As we change m_s from 0.1τ to 10τ, the D_2 first increases, then reaches a peak at about 1τ and subsequently falls for higher m_s. We illustrate this for a particular value of m_p in Figure 3.4.

We now examine the m_s-m_p parameter plane for the Lorenz and Rössler systems, plotted in Figure 3.5. We observe that with increasing m_p the D_2 decreases for all values of m_s. Further, we find that the peak at 1 τ continues to persist across values of m_p. We will now examine this more closely.

The peak at 1τ is largely due to the way we have reconstructed the state space of the dynamical system. Gaps are chosen in units of τ, the point where the autocorrelation falls to $\frac{1}{e}$. Let the state space vectors reconstructed from the evenly sampled time series

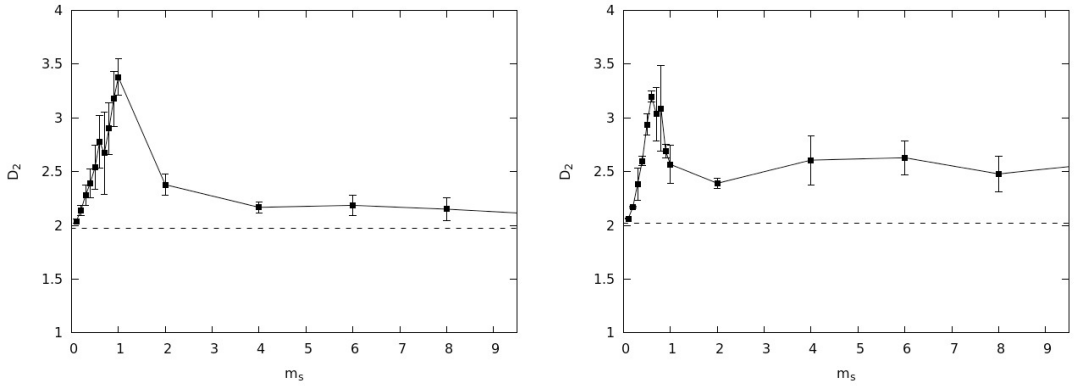


Figure 3.4: Variation of saturated D_2 with increasing mean position, m_s for the Lorenz (left) and Rössler (right) systems. m_p if fixed at 9τ. The error bars are standard deviations over five realizations.

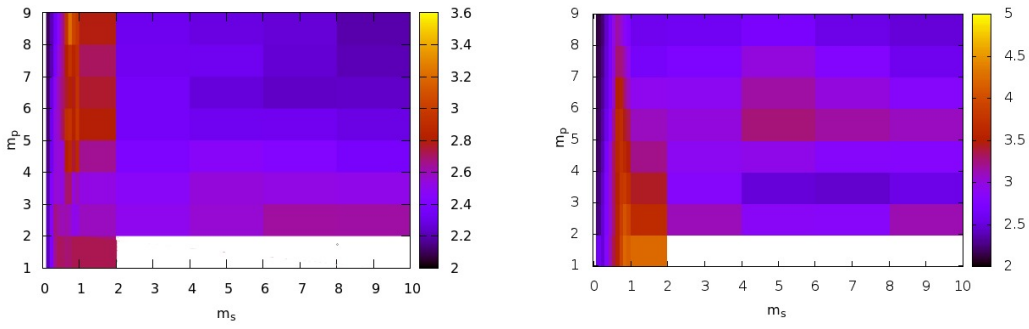


Figure 3.5: m_s, m_p parameter plane showing variation of saturated D_2 for the Lorenz (left) and Rössler (right) systems. The value of D_2^{sat} is color coded in the plane.

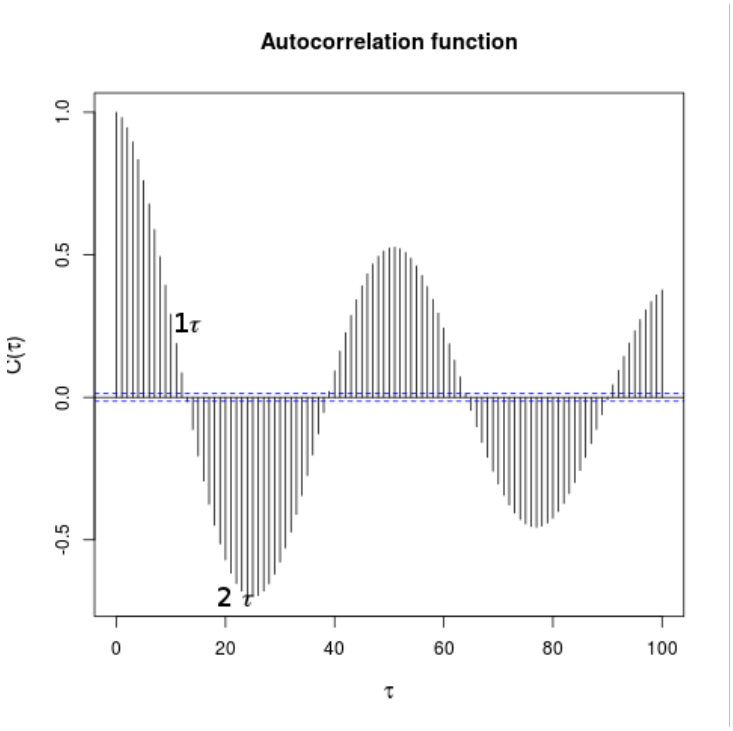


Figure 3.6: Autocorrelation function as a function of lag, τ for $m_s = 1\tau$, $m_p = 10\tau$.

be labeled as $\vec{v}^e$, say in $3-d$ space.

$$\vec{v}^e = (x_i, x_{i+\tau_e}, x_{i+2\tau_e}) \tag{3.3}$$

where τ_e is the evenly sampled delay time. Now, say the mean gap size m_s, is $1\tau_u$ (we label τ as τ_u, but recollect that for the gap ranges considered, τ_u and τ_e are similar). For a vector close to the gap, the reconstructed state space vector is

$$\vec{v}^u = (x_i, x_{i+2\tau_u}, x_{i+3\tau_u}) \tag{3.4}$$

Because of the periodicities in the attractor, $x_{i+2\tau_u}$ is anti correlated to x_i, as can be easily seen from the autocorrelation function for the data with gaps in shown in Figure 3.6. This results in a line of vectors lying along the diagonal $x = -y$, as can be seen in Figure 3.7. Hence the attractor is reconstructed with more space filled than it actually is, resulting in increased D_2. This does not affect the component at $x_{i+3\tau_u}$ as much, since the correlation dies down quickly.

An interesting question is whether the value of the embedding dimension at which D_2 saturates,M_d, changes as we introduce gaps. As we vary m_s and m_p , M_d varies between 3 and 6 for both systems. We notice that the deviation of M_d coincides with the deviation of D_2^{sat}. This is a point worth noting, as in a real dataset with gaps, when the actual embedding dimension is not known, relying on the M_d to find the dimensionality of the system may lead to a larger value than required.

We now come back to the question of how the standard deviation of the Gaussians affect the D_2 calculations. We fix the mean and vary standard deviations from 0.2 to 2.0 times the mean. Our results indicate that varying the standard deviations, ω_p and ω_s does not affect the D_2 value too much. This is illustrated for a specific value of m_p and m_s in Figure 3.8

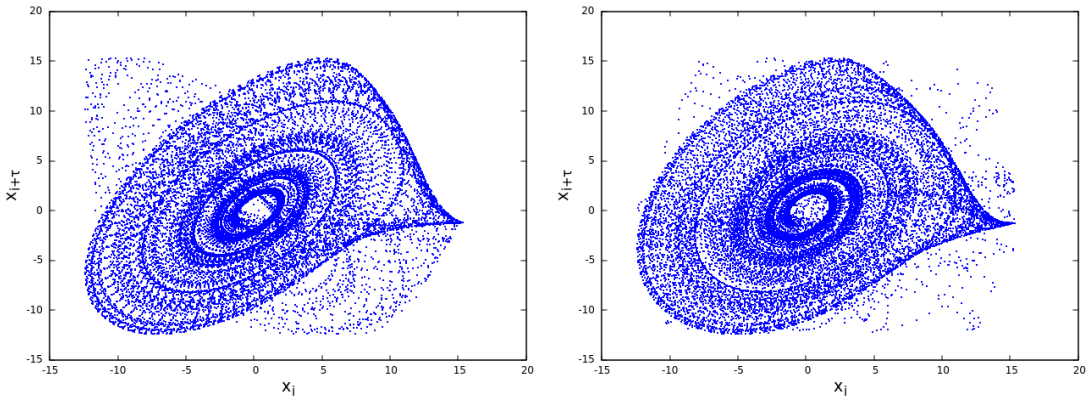


Figure 3.7: Reconstructed Rössler system for $m_s = 1\tau$ (left) and $m_s = 2\tau$ (right). The m_p is fixed for both cases at 10τ

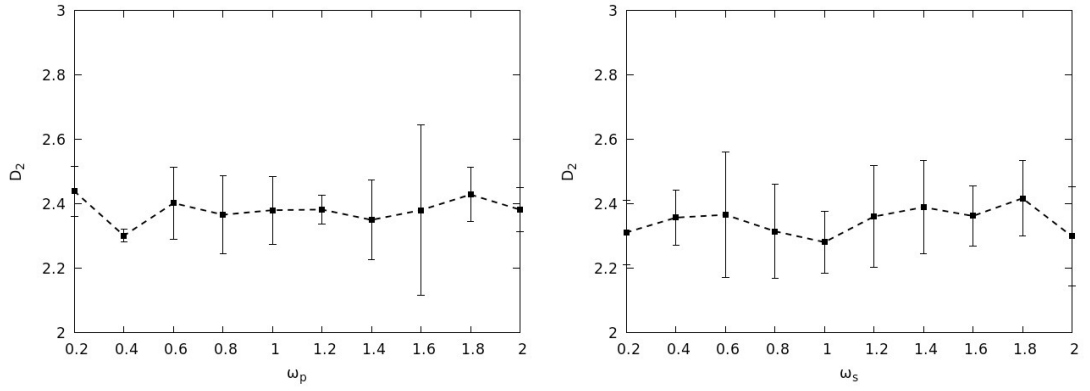


Figure 3.8: Variation of D_2 with varying standard deviations ω_p(left) and $omega_s$(right) for the Lorenz system. m_p and m_s are both fixed at 9τ in both figures.

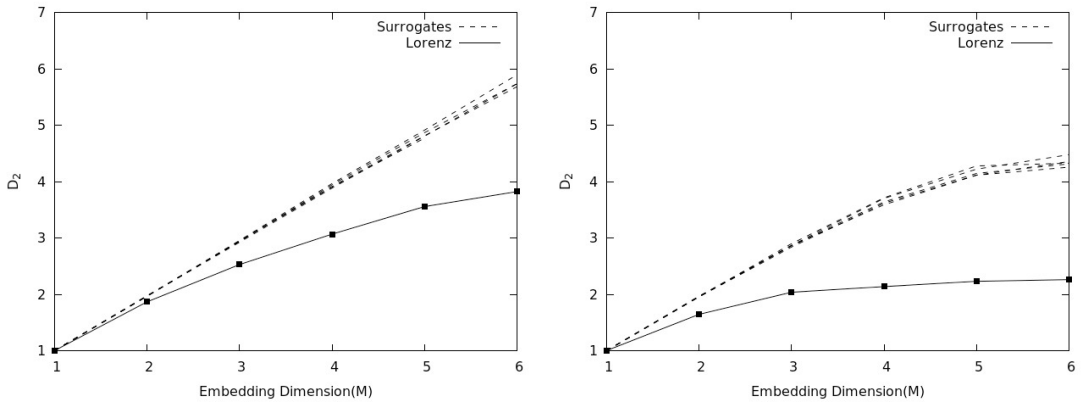


Figure 3.9: D_2 vs M plots for the Lorenz system and surrogates, with gaps. The mean gap size and positions are (a)$m_s = 0.8\tau, m_p = 9\tau$ (left) and (b)$m_s = 10\tau, m_p = 9\tau$ (right)

Finally, we consider the question of how the presence of datagaps is different from noise. To do this, we resort to the method of surrogate analysis. The surrogate datasets are generated using the method of Iterative Amplitude Adjusted Fourier Transform (IAAFT) [95]. We use the TISEAN package to generate the surrogates [96]. Since the IAAFT algorithm is suited only for evenly sampled datasets, we first generate surrogates from the evenly sampled datasets. The same profile of gaps used on the original dataset is used to introduce gaps into the 5 surrogate datasets. We then calculate $D_2(M)$ for the original data and surrogate datasets. We plot the $D_2(M)$ vs M plots for 2 cases (a)$m_s = 0.8\tau, m_p = 9\tau$ close to the peak in Figure 3.4 and (b)$m_s = 10\tau, m_p = 9\tau$ away from the peak, in Figure 3.9. We see that in both cases the surrogates are away from the data, suggesting that even though D_2^{sat} deviates highly from the evenly sampled value in most cases, it continues to behave distinctly different from noise [97].

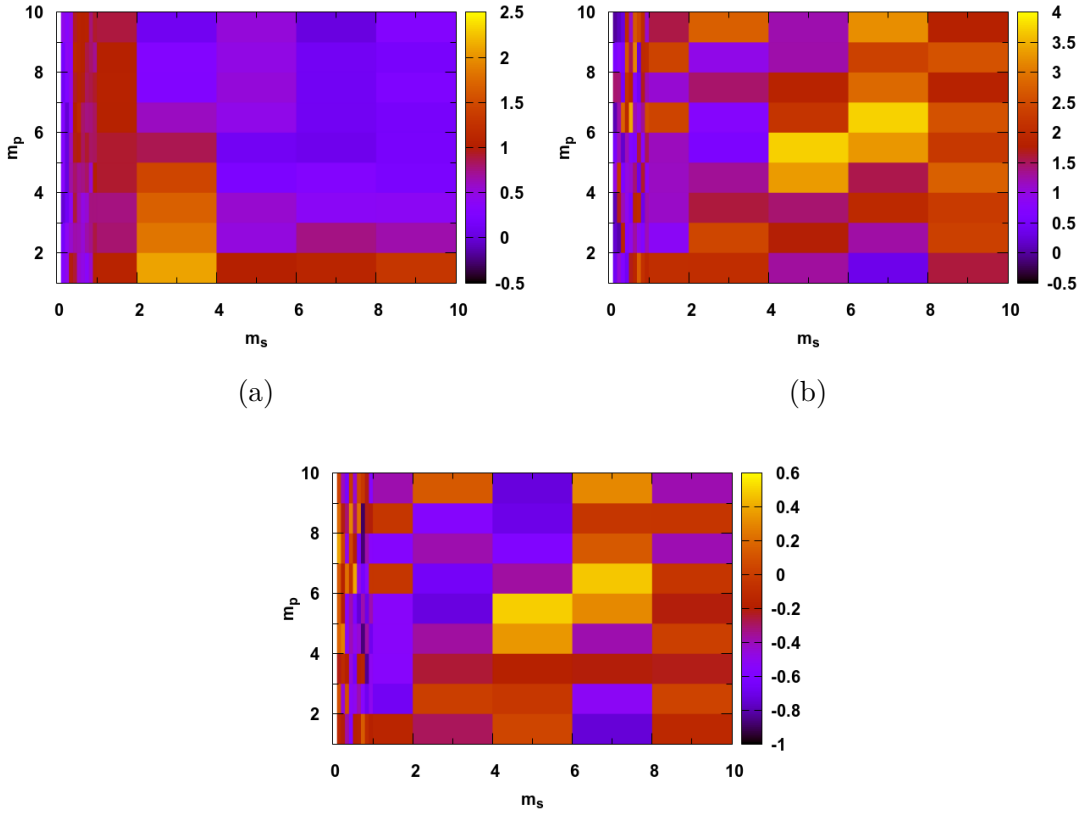
3.3.2 Multifractal spectrum

We will now proceed to consider how datagaps affect calculation of the $f(\alpha)$ or multifractal spectrum. We discussed the $f(\alpha)$ spectrum in detail in the first chapter. The $f(\alpha)$ curve can be characterized using the function given in equation 1.48, in chapter 1. Recollect that α_1, α_2, γ_1 and γ_2 uniquely characterized the $f(\alpha)$ curve [41]. An important measure is the width of the $f(\alpha)$ curve which gives a measure of the complexity or scales involved in the system. This is given by $\Delta\alpha = \alpha_2 - \alpha_1$. We initially define the deviation of the quantifiers from those of the evenly sampled data

$$\begin{aligned} \delta\Delta\alpha^{rel} &= \frac{\Delta\alpha - \Delta\alpha_E}{\Delta\alpha_E} \\ \delta\gamma_1^{rel} &= \frac{\gamma_1 - \gamma_{1E}}{\gamma_{1E}} \\ \delta\gamma_2^{rel} &= \frac{\gamma_2 - \gamma_{2E}}{\gamma_{2E}} \end{aligned} \tag{3.5}$$

Here $\Delta\alpha_E$, γ_{1E} and γ_{2E} are the evenly sampled values of these quantifiers.

We conduct the analysis the same way we did for D_2, by starting from large evenly sampled datasets. Gaps are introduced in terms of the delay time τ and varied as before, such that the resulting datasets all have ≈ 20000 points each. The $m_s - m_p$ parameter plane for the Rössler system is shown in Figure 3.10.



(a) (b)

(c)

Figure 3.10: $m_s - m_p$ parameter planes for the variation of (a) $\delta\Delta\alpha^{rel}$ (b)$\delta\gamma_1^{rel}$ and (c) $\delta\gamma_2^{rel}$ for the Rössler system.

We see that $\Delta\alpha$ deviates highly from the evenly sampled value at low m_p and approaches the evenly sampled value at higher m_p. For m_s, the $\Delta\alpha$ increases initially, reaches a peak at 1τ and then falls again at higher values. The behavior for the Lorenz system is similar, but with a less prominent peak around $m_s = 1\tau$.

The γ_1 and γ_2 planes indicate that in the presence of datagaps, the γ values deviate extensively from the evenly sampled value. This seems to suggest that in the presence of datagaps, the estimated γ values cannot be relied upon.

A wide multifractal spectrum alone does not guarantee the system is deterministic. Colored noise processes too show a saturating correlation dimension and a wide $f(\alpha)$ spectrum [41, 98]. We again use the method of surrogate data testing using IAAFT to determine whether the data behaves similar to noise or chaos. We first show how the $f(\alpha)$ spectrum for data and surrogates for standard systems and noise processes vary. This is shown in Figure 3.11, and listed in Table 3.1. It is clear that for chaotic data the $f(\alpha)$ is different from that of surrogate datasets, while for noise data and surrogates are very close in $f(\alpha)$.

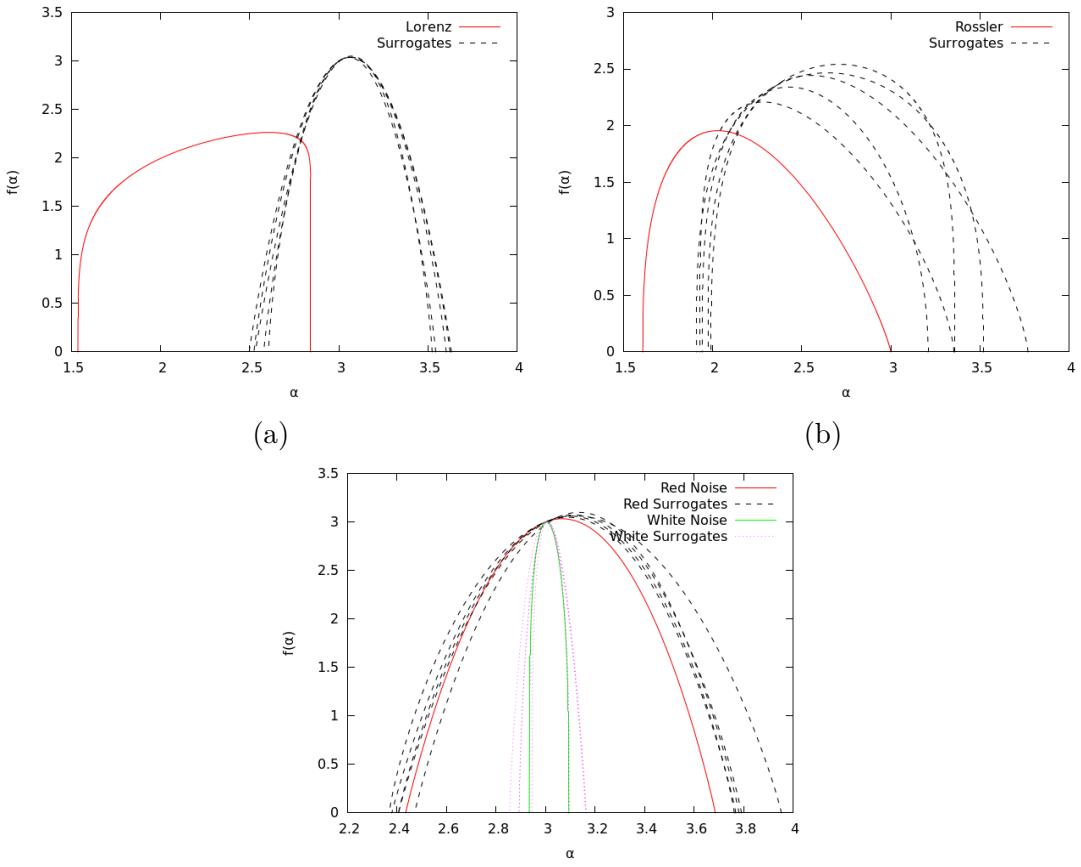


Figure 3.11: $f(\alpha)$ curves of (a) Lorenz (b) Rössler and (c)white and red noise along with 5 surrogate datasets for each.

We then examine the variation of data from surrogates when the time series is affected by datagaps. We take examples of Rössler data with mean gap sizes, $m_s = 1\tau$ and $m_s = 8\tau$, and mean gap position, $m_p = 7\tau$. This is shown in Figure 3.12. Table 3.2 shows the values for α_1 and α_2 for the data and its surrogates. We see that the data and surrogates merge for α_2 in both cases. For the case where $m_s = 8\tau$, $m_p = 7\tau$, α_1

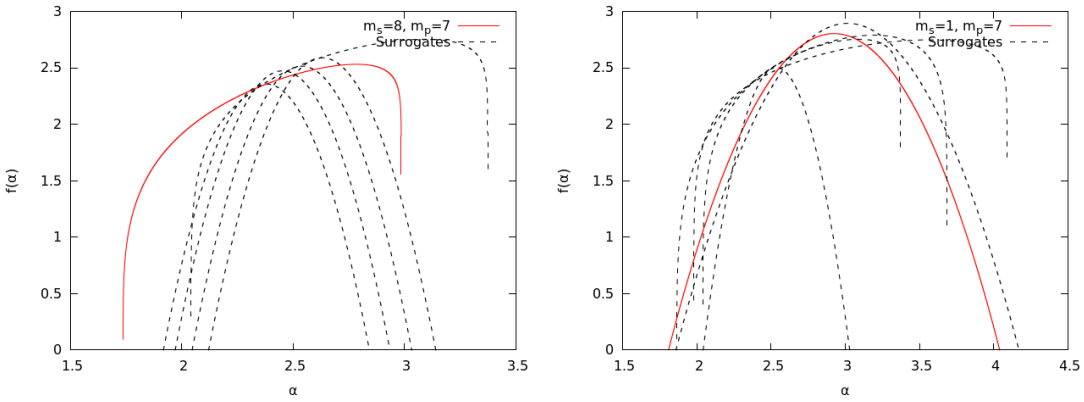
Table 3.1: α_1 and α_2 values for Lorenz, Rössler, red noise, white noise and their surrogates.

Data	α_1^d	α_1^s	α_2^d	α_2^s
Lorenz	1.54	2.55±0.04	2.84	3.58±0.04
Rössler	1.60	1.95±0.03	2.99	3.44±0.19
Red noise	2.43	2.41±0.04	3.68	3.81±0.07
White noise	2.93	2.90±0.03	3.09	3.13±0.03

Table 3.2: α_1, α_2 values for the Rössler system and its surrogates for two different values of m_s.

m_s	m_p	α_1^d	α_1^s	α_2^d	α_2^s
8τ	7τ	1.73	2.02±0.03	2.98	3.06±0.18
1τ	7τ	1.81	1.95±0.07	4.04	3.82±0.43

is distinctly different for data and surrogates. However for $m_s = 1\tau$, $m_p = 7\tau$ data and surrogates merge.



(a) (b)

Figure 3.12: The $f(\alpha)$ spectra of Rössler system with gaps for (a)$m_s = 8\tau, m_p = 7\ \tau$ and (b) $m_s = 1\ \tau, m_p = 7\ \tau$.

In the regions of the $m_s - m_p$ parameter plane where $\Delta\alpha$ deviates highly from the evenly sampled value, the $f(\alpha)$ curves are indistinguishable from noise. This is unlike in the case of D_2, where even when the D_2 value deviated away from the evenly sampled value, it saturated distinctly away from surrogates.

3.4 Application to Real World Datasets

We now apply the analysis conducted in the previous section to analyse two real world datasets. The first is composed of light curves of irregular variable stars observed by the American Association of Variable Star Observers (AAVSO) [99]. The second consists of a set of five datasets observed at the Station for Measuring Forest Ecosystem-Atmosphere Relation (SMEAR), Finland [100, 101].

3.4.1 AAVSO Variable Star Light Curves

Astrophysical datasets are highly prone to datagaps because of the nature in which they are collected. Since astronomy relies almost entirely on observational data, the control one has over external factors is limited. Cloud covers, eclipses and unfavourable weather conditions limit continuous observation in optical astronomy. Radio astronomy observations are affected highly by radio emissions from man made sources, ionospheric scintillation, lightning, solar emissions etc. While many of these sources of gaps can be eliminated in space based observatories, they too are plagued by datagaps, albeit to a lesser extent [102].

Data can also be completely unevenly sampled, and AAVSO data is a typical example of such datasets. There is no underlying sampling time, as the observations are made by different individuals at random times. The study on pure uneven sampling focuses on the distributions of sampling times, Δt [103, 104]. It has been shown that Δts follow a gamma distribution [105]. For very small mean gap positions, m_p, we expect the data to behave similar to datasets with no inherent sampling time. In fact we can show that for small m_s and m_p, the sampling time distribution does follow a gamma distribution.

For the AAVSO datasets, we attempt to convert purely unevenly sampled datasets to cases of time series with datagaps. We achieve this through binning. Binning introduces an artificial sampling time into the data, equal to the bin size b. It has two major effects. First it reduces the noise, and second it converts unevenly sampled data into a data that we've handled in the previous sections. For binning, we chose a threshold n_b^t. If the number of points within a bin, n_b, is less than n_b^t, we treat that as a gap. Otherwise the average over all the points in that bin is chosen as the value at that time bin. For our calculations, we choose $n_b^t = 3$. We choose two values for b, 2 days and 5 days. We consider light curves of four irregular variable stars namely R Scuti, U Monocerotis, SU Tauri and SS Cygni, all of which have been suspected of nonlinear dynamics in the past. Sections of light curves of all four stars are shown in Figure 3.13.

R Scuti

R Scuti (R Sct.) is an RV Tauri variable whose nonlinear properties have been extensively studied in the past [84, 106, 107]. We first calculate the delay time,τ, of the light curve which is about 20 days. The binning is first done for bin size, $b = 2$ days. This gives $m_p = 33$ days and $m_s = 8$ days. In terms of τ this gives $m_p \sim 1.65\tau$ and $m_s \sim 0.4\tau$. We see that the D_2 vs M curve does not yield a saturated value. This could be because of the low value of m_p. We then bin the light curve with $b = 5$ days. This gives $m_p \sim 10.6\tau$ and $m_s \sim 1.1\tau$. m_s still falls into the critical region of D_2 identified before, and we do not get a saturated value, D_2^{sat}. Finally, we try a bin size $b = 10$ days. This gives $m_p \sim 16.4\tau$ and $m_s \sim 1.4\tau$, and a saturated value of $D_2^{sat} = 5.67 \pm 0.10$. This value is higher than previously reported values of D_2 [84, 106]. We suggest that this could either be because of m_s being close to the critical range considered, or because of the smoothing methods employed in previous calculations. However conclusions regarding the nature of the dynamics remains the same. Graphs of D_2 vs M for the three bin sizes is shown in Figure 3.14.

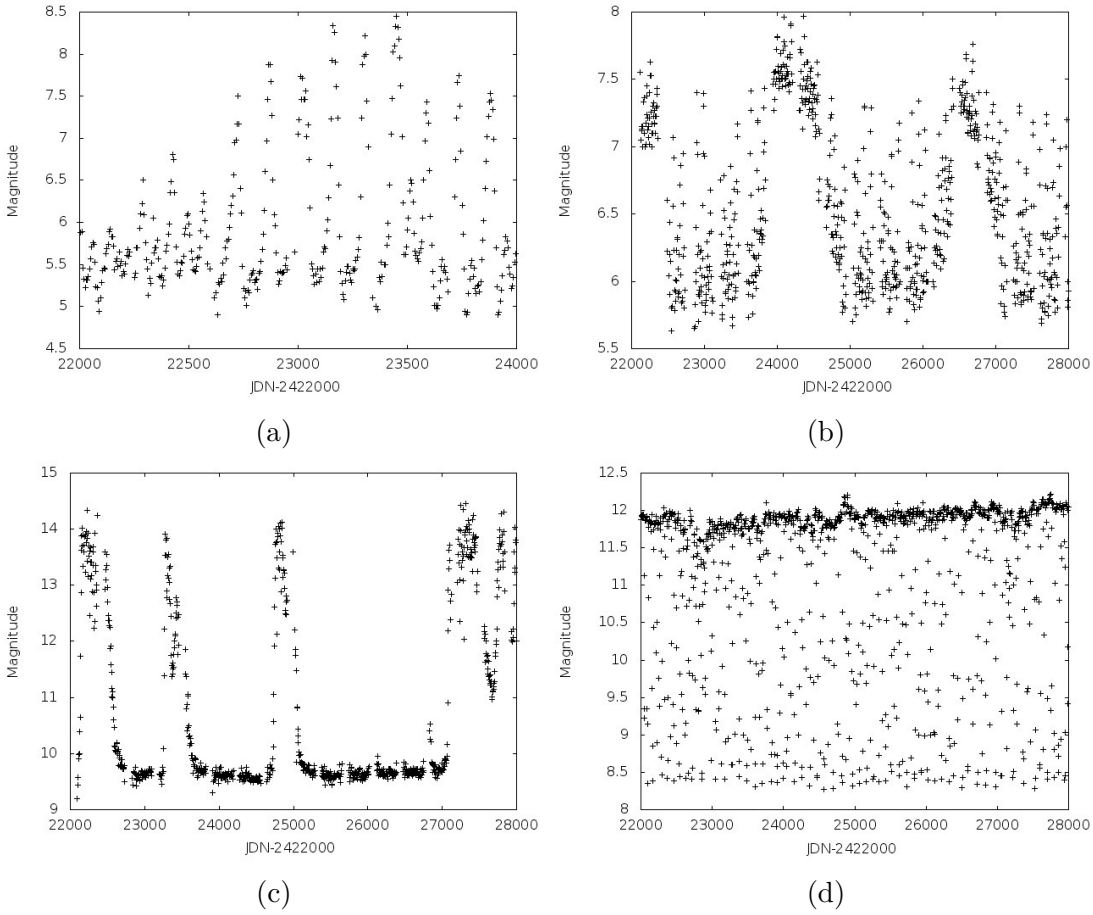


Figure 3.13: AAVSO light curves for variable stars (a)R Sct. (b)U Mon. (c)SU Tau. and (d)SS Cyg. binned over 5 day bins.

U Monocerotis

U Monocerotis(U Mon.) is a pulsating variable star that falls into the RVb Tauri subcategory of stars [108]. Variability in RV Tauri stars has been hypothesized to be due to their nonlinear behavior [109]. The delay time τ for U Mon. is found to be 15 days. With $b = 2$ days, we get $m_p = 11.4$ days or $\sim 0.76\tau$, and $m_s = 15.87$ days or $\sim 1\tau$. The m_s and m_p fall exactly into the critical region identified, and the D_2 vs M curve does not saturate. When we use $b = 5$ days, we get $m_p \sim 3.5\tau$ and $m_s \sim 2.9\tau$. This yields a $D_2^{sat} = 3.44 \pm 0.08$. The D_2 vs M curves are shown in Figure 3.15. To the best of our knowledge there are no existing calculations of D_2^{sat} for this star for comparison.

SU Tauri

SU Tauri(SU Tau) is a variable star of the R Coronae Borealis type [108].Its variations were suspected to be of chaotic origin previously [110]. The delay time of this star was found to be $\tau \sim 224 days$. With $b = 2$ days, the mean gap position and size were found to be $m_p \sim 0.06\tau$ and $m_s \sim 0.06\tau$. The calculated D_2^{sat} was 3.26 ± 0.08. With a bin size of $b = 5$ days the means changed to $m_p \sim 0.28\tau$ and $m_s \sim 0.15\tau$. This yielded $D_2^{sat} = 3.28 \pm 0.08$ Despite the high frequency of gaps, the consistent values of D_2^{sat}, indicate that these may be reliable. In this case also, no known calculations of D_2 exist for comparison. The D_2 vs M plots are shown in Figure 3.16.

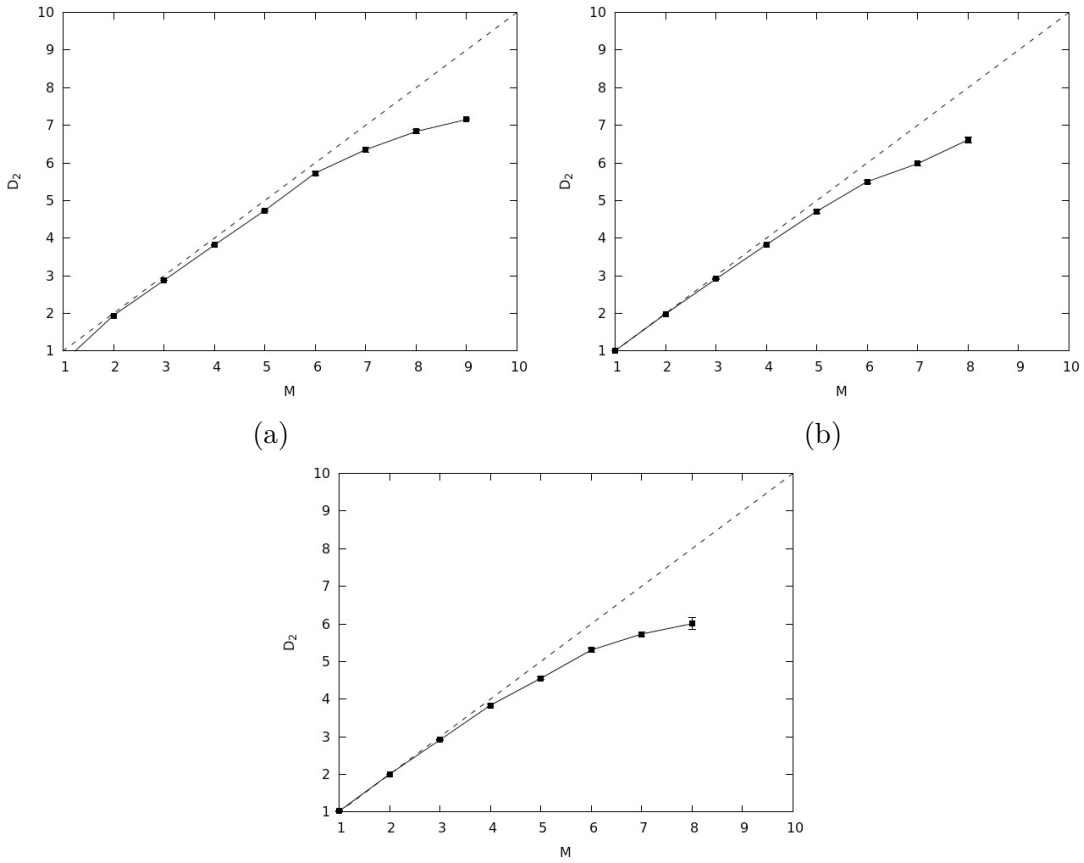


Figure 3.14: D_2 vs M plots for variable star R Sct. binned over (a)2 days (b)5 days and (c)10 days.

SS Cygni

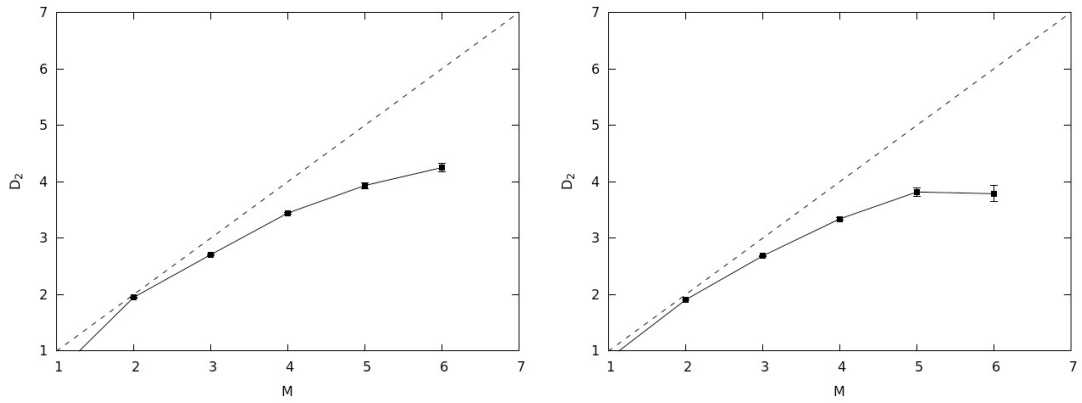
SS Cygni is a cataclysmic variable of the dwarf nova class [111]. Its dynamics was a subject of dispute in the early 1990s [112,113]. Subsequently no evidence for strangeness was found in its dynamics [114]. We follow the previous routine of binning the light curves using bin sizes $b = 2$ days and 5 days. The former yields $m_p \sim 10\tau$ and $m_s \sim 0.5\tau$, while the latter yields $m_p \sim 50\tau$ and $m_s \sim 1\tau$. Both datasets show no saturation, despite both m_s and m_p not falling into critical regions for $b = 2$ days.

We then point out the dangers of interpolation for this star, by interpolating through both datasets and recalculating D_2^{sat}. While the dataset with $b = 2$ does not show any saturation even after interpolation, the dataset with $b = 5$ yields a saturated value for $D_2 = 4.14 \pm 0.05$. Figure 3.17 shows the D_2 vs M for both the interpolated and non interpolated cases for two bin sizes.

The results for the correlation dimension calculations for all the stars are tabulated in Table 3.3.

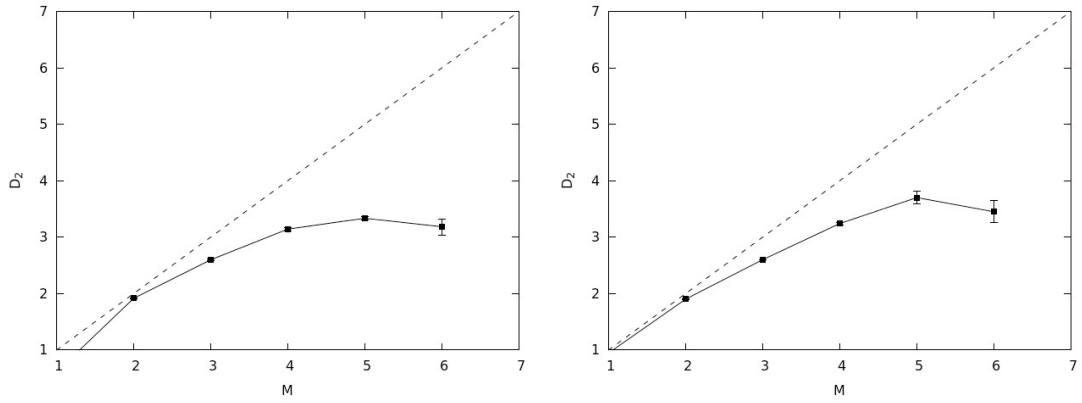
3.4.2 SMEAR Datasets

In this section, we will analyse the multifractal spectra of datasets with gaps, gathered at the Station for Measuring Forest Ecosystem-Atmosphere Relation (SMEAR) in Finland. We consider five datasets: three datasets from SMEAR I, namely CO_2 exchange, photo-



(a) (b)

Figure 3.15: D_2 vs M plots for variable star U Mon. binned over (a)2 days and (b)5 days.

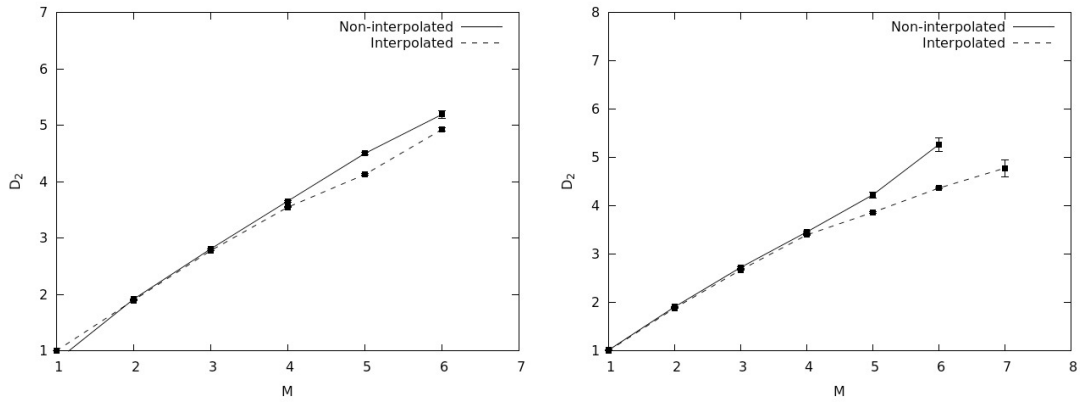


(a) (b)

Figure 3.16: D_2 vs M plots for variable star SU Tau. binned over (a)2 days and (b)5 days.

synthetically active radiation (PAR) and soil moisture, and two from SMEAR II, namely air temperature and dew point. SMEAR datasets are affected by datagaps primarily due to instrument failure due to breaks in electricity, thunderstorms and so on [115, 116]. In this subsection, we will attempt to estimate the $f(\alpha)$ spectra of these datasets from SMEAR. The SMEAR I station is located 200 km north of the Arctic circle, in the Varriö forest. We consider three main observational datasets from this station, i.e. the time series of CO_2 exchange, photosynthetically active radiation(PAR) and soil moisture. The first of these is known to be a good proxy for photosynthesis rate [117], which in turn is known to depend on the latter two. The photosynthesis rate of plants have been suspected to have chaotic dynamics by multiple authors, previously [118, 119]. The variation of photosynthesis rate is known to be affected by variables such as ambient temperature, CO_2 concentration, PAR, soil moisture etc [120, 121]. The dynamics of these variables in turn also affect each other and hence are in constant feedback.

The SMEAR II station is located in the Hyytiälä forest. It primarily measures atmospheric aerosols, eco-physiology, soil and water measurements, solar and terrestrial radiation and meteorological measurements. We choose two meteorological dataset from SMEAR II, namely the dew-point and air temperature datasets, due to the high instance



(a) (b)

Figure 3.17: D_2 vs M plots for variable star SS Cyg. binned over (a)2 days and (b)5 days, with and without interpolation.

Table 3.3: Results for variable stars

Star	Bin($days$)	τ(days)	$m_s(\tau)$	$m_p(\tau)$	D_2
R Sct.	2	20	0.4	1.65	N.S
R Sct.	5	20	1.1	10..6	N.S
U Mon.	2	15	1	.38	N.S
U Mon.	5	15	2.9	0.7	3.436
SU Tau.	2	224	.06	.06	3.265
SU Tau	5	224	.15	.28	3.285
SS Cyg.	2	10	0.5	10	N.S
SS Cyg.	5	10	1.0	50	N.S

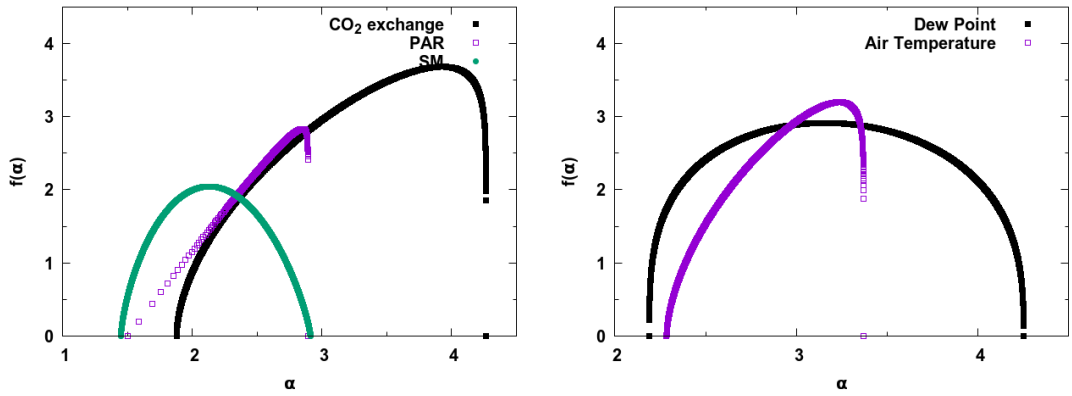
of datagaps in them [115]. Multifractal analysis of climate datasets has been conducted previously [122, 123].

SMEAR I and II datasets are sampled every half an hour. We use data between the period 2005-2017 for SMEAR I and from 2008-2017 for SMEAR II datasets. The D_2,τ, m_s and m_p values are shown for the five datasets in Table 3.4. We notice that the mean gap parameter ranges for soil moisture and air temperature fall into the critical region identified. To estimate the $f(\alpha)$ spectrum, we first embed the time series with an embedding dimension, M_d equal to the smallest integer greater than D_2. The calculated $f(\alpha)$ spectra are separately shown for the SMEAR I and II datasets in Figure 3.18

Table 3.4: τ, m_s, m_p and D_2 for time series from the SMEAR datasets.

Data	$\tau(hours)$	$\mathrm{m}_s(hours)$	$\mathrm{m}_p(hours)$	D_2
CO_2 Ex.	5	119.6	157.3	3.18
PAR	7	119.6	157.3	4.08
SM	970	499.5	971.1	1.15
Dew Pt.	1000	219.2	252.7	2.63
Air T.	667	483.2	535.1	3.28

We then look for signs of determinism, by constructing surrogates for these datasets,



(a) (b)

Figure 3.18: $f(\alpha)$ spectra for datasets from (a)SMEAR I (b)SMEAR II.

Table 3.5: α_1, α_2 values for SMEAR datasets and their surrogates. α_1^d and α_2^d are distinctly different from α_1^s and α_2^s for CO_2 exchange and PAR, while they are within errors of each other for the other three.

Data	α_1^d	α_1^s	α_2^d	α_2^s
CO_2 Ex.	1.87	2.55±0.01	2.26	4.75±0.01
PAR	2.10	3.39±0.07	3.08	3.88±0.09
SM	1.45	1.35±0.02	2.91	2.88±0.03
Dew Pt.	2.19	2.31±0.06	4.25	4.25±0.05
Air T.	2.29	1.95±0.01	3.79	4.16±0.04

and calculating $f(\alpha)$ spectra for them. Since these datasets are affected by datagaps, we construct surrogates in the following way. Initially, we take regions of the data that are continuous i.e there are no gaps present in the data. Surrogates are generated for these continuous segments and joined together to form a surrogate for the whole dataset. The presence of large continuously sampled segments in the SMEAR datasets makes this possible, but we assert that this may not always be the case. The surrogate datasets are then subjected to $f(\alpha)$ calculations as before. The α_1 and α_2 values for the five datasets and their surrogates are shown in Table 3.5. We see that the data and surrogates are distinctly different for CO_2 exchange and PAR, but are reasonably close to each other for the other three datasets. Of these air temperature and soil moisture have m_s falling into the critical region identified, while the dew point too has a high frequency of gaps. Hence it is difficult to identify if the data and surrogates co-vary due to lack of deterministic nonlinearity or due to the gap profile in these datasets [124].

3.5 Summary and Discussion

The presence of datagaps in time series has been a major impediment in their analysis. As we discussed previously in this chapter, smoothing techniques can affect estimates of nonlinear quantifiers, often leading to false positives. We considered the extend to which nonlinear time series quantification is affected by the presence of datagaps in time series data, when no such smoothing techniques are employed.

We quantified the resilience of two popular state space quantifiers, namely D_2 and $f(\alpha)$ spectrum, to datagaps. We observe that a tolerable region exists for D_2 and $\Delta\alpha$, within which reliable conclusions can be drawn about the nature of the dynamics of the system. We then proceeded to use the analysis we conducted, to deal with real world AAVSO and SMEAR datasets, and estimate the D_2 and $f(\alpha)$ spectra, respectively. We illustrate how suitable binning can help to shift datagaps to tolerable limits and then the estimated value of nonlinear quantifiers can give reliable conclusions.

Chapter 4

Distinguishing Dynamical States From Time Series Using Bicoherence

A major impediment in understanding the dynamics of systems that give rise to real world time series observations, is the presence of noise. In this thesis we define noise processes as those that have no correlations in their Fourier phases. Systems may have stochastic dynamics or be contaminated by a stochastic process. It may often be challenging to determine whether a system is primarily stochastic or deterministic, especially if a deterministic system is exhibiting irregular behavior like chaotic or strange non-chaotic dynamics. Further, it is often difficult to distinguish between different nonlinear dynamical states when the data is contaminated by noise or when the system is evolving with noise.

In this chapter, we will present the results of our study to distinguish between dynamical states in the presence of noise by using higher order spectra, specifically the bicoherence. We introduced the bicoherence in some detail in chapter 1. We will review it again now. The bicoherence function is defined in equation 1.33 in chapter 1. It measures the extent of quadratic coupling between frequencies at f_1 and f_2, by looking for a response at $f_1 + f_2$[1]. The bicoherence function is useful when dealing with noisy data because it is not phase blind unlike the power spectrum. Further, it does not distinguish between symmetrically distributed mean zero noises, irrespective of its color.

We discuss the use of the bicoherence function to distinguish noisy periodic phenomena from deterministic irregular phenomena, when noise contamination is involved. Specifically we use it to distinguish noisy limit cycles from chaotic behavior and noisy quasiperiodicity from strange non chaotic behavior.

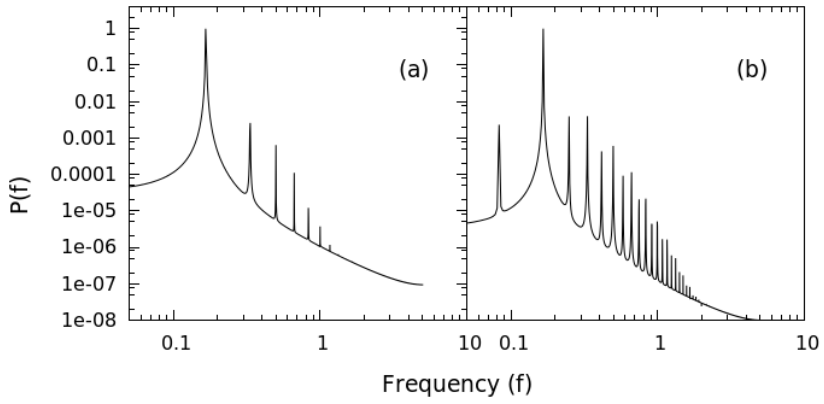
[1]One can motivate this by imagining a quadratic system with an x^2 response. Suppose the input signal is $x = sin(\omega_1 t) + sin(\omega_2 t)$. The system would cause components to appear at $sin(\omega_1 t + \omega_2 t)$ and $sin(\omega_1 t - \omega_2 t)$. The bicoherence checks if these components in the Fourier transform are related to each other.

4.1 Analysis of Noisy Limit Cycles

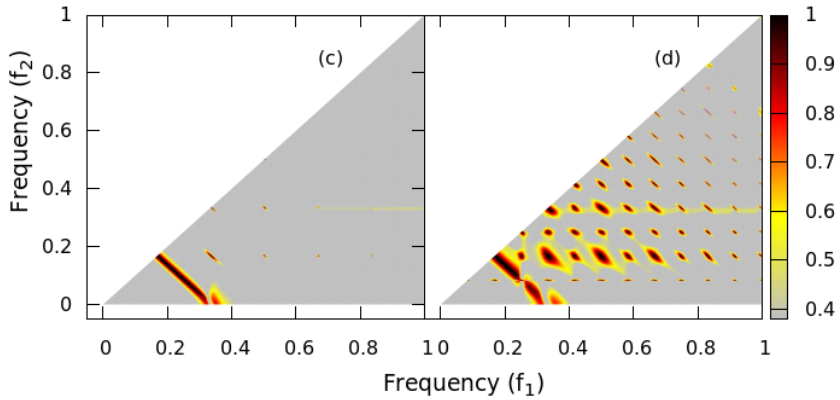
To study the behavior of noisy limit cycles, we use the Rössler system evolved in the presence of noise as our test system. It is given by

$$\begin{aligned} \dot{x} &= -y - z + \eta(t) \\ \dot{y} &= x + ay \\ \dot{z} &= b + z(x - c) \end{aligned} \tag{4.1}$$

where a=0.1, b=0.1, c is varied for the regime under consideration and η(t) is δ-correlated mean zero Gaussian white noise. We initially set $\eta(t)$ to 0, and check the bicoherence for the noiseless Rössler in the periodic regime [125]. The power spectra and bicoherence plane for the period 1 and 2 regimes is shown in Figure 4.1. The bicoherence is plotted only for the frequency pairs over the 99% significance threshold, given by $\sqrt{\frac{9.2}{dof}}$ [30]. Here $d.o.f$ is $2N$, where N is the number of segments into which the time series is divided. In our study,we divide the time series into 32 segments with 4096 points each.



(a)



(b)

Figure 4.1: Power spectrum for the Rössler system in (a)period 1 and (b)period-2 regimes. The corresponding bicoherence graphs zoomed to $[0, 1]$ region are shown in (c) and (d). The islands of high bicoherence corresponds to the finite width of the peaks in the power spectra.

We then proceed to analyze the difficulty in dealing with the case where $\eta(t)$ is non zero. We set the standard deviation of $\eta(t)$ to 2.0 for the periodic regimes($c = 4$ and 6.2)

and at 5.0 for the chaotic regime($c = 18$). The power spectra for the periodic Rössler evolved in the presence of noise, is shown in Figure 4.2 a and b. We see a continuous inter-peak peak power in them, that is absent in the noiseless case. This is similar to the noiseless chaotic power spectrum shown in Figure 4.2 c. Further, as can be seen from the chaotic Rössler evolved with noise (Figure 4.2 d), the fine features of the chaotic power spectrum is lost when contaminated by noise. Hence it becomes difficult to understand the underlying dynamical state of the system looking at the power spectrum alone.

The bicoherence corresponding to the four cases is shown in Figure 4.3. Unlike the power spectra, the bicoherence graphs show distinct features for the periodic regimes evolved with noise, different from the chaotic regime. For the periodic time series, the points corresponding to the primary frequency and its harmonics can be clearly identified[2].

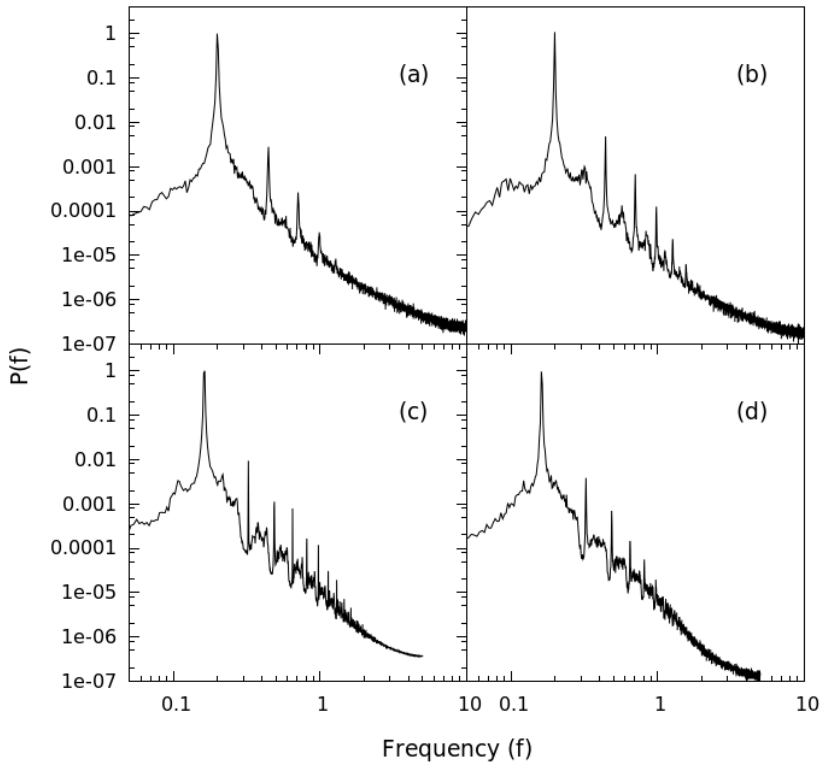


Figure 4.2: Power spectra for the Rössler system evolved in the presence of noise in (a)period-1 (b)period-2 and (d)chaotic regimes. (c) shows the noiseless Rössler in the chaotic regime.

4.2 Analysis of Noisy Quasiperiodicity

In this section, we will try and distinguish between noise contaminated quasiperiodicity and strange non chaotic behavior. Quasiperiodicity is a dynamical state exhibited by

[2]The small but significant bicoherence one sees along the horizontal lines is due to the finite width of the peaks due to the finite length of the data, which causes the stochastic peaks on its wings to be misidentified as being of chaotic origin.

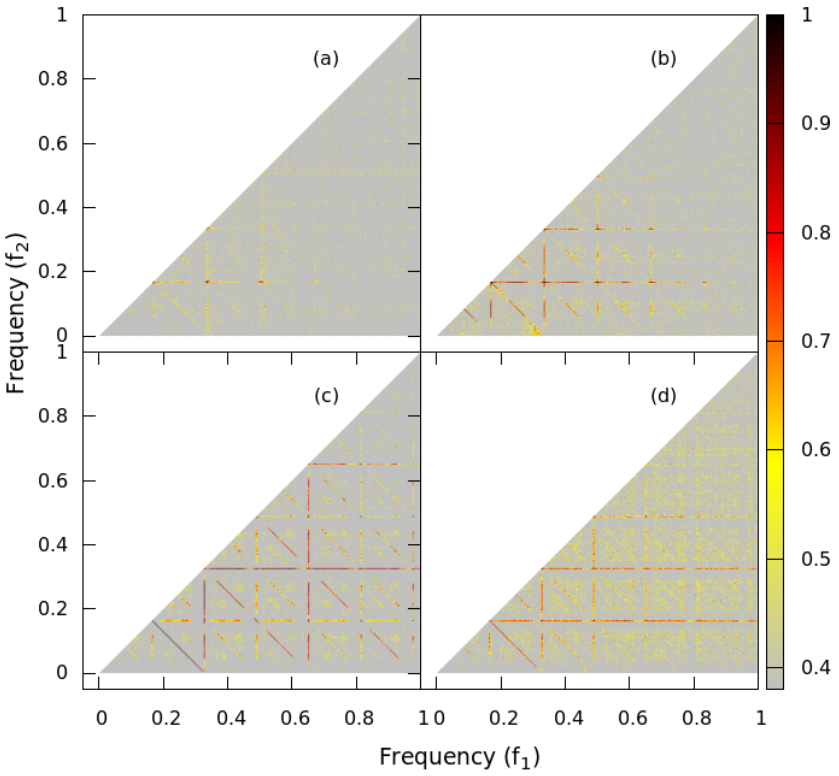


Figure 4.3: Bicoherence graphs for the Rössler system evolved in the presence of noise in (a)period-1 (b)period-2 and (d)chaotic regimes. (c) shows the noiseless Rössler in the chaotic regime.

dynamical systems possessing two or more independent frequencies of oscillations that are incommensurate. The power spectrum of such a state would contain peaks at the original frequencies f_1 and f_2 and all their linear combinations like f_1+f_2, f_1-f_2, f_1+2f_2 and so on [7]. Our definition of the bicoherence in equation 1.33 seems to suggest that in such a system, the bicoherence of pairs such as (f_1, f_2) and $(f_1 - f_2, f_2)$ would be significant. Hence the bicoherence function seems to be a natural quantifier to deal with systems exhibiting quasiperiodicity.

Since all possible frequencies in the power spectrum must arise from linear combinations of these fundamental frequencies, it would suffice for us to look at bicoherence pairs where one of these frequencies is a member. We hence define a quantifier, the main peak bicoherence,

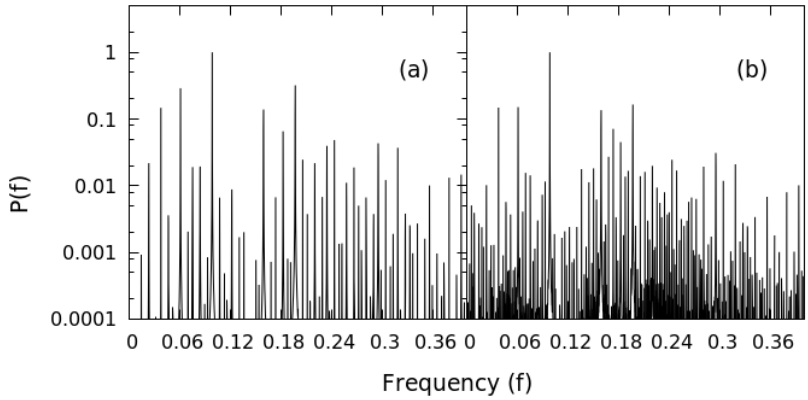
$$b_F(f) = \frac{|\sum_{i=1}^{k} A_i(F)A_i(f)A_i^*(f+F)|}{\sum_{i=1}^{k} |A_i(F)A_i(f)A_i^*(f+F)|} \tag{4.2}$$

where F is the maximal peak in the power spectrum.

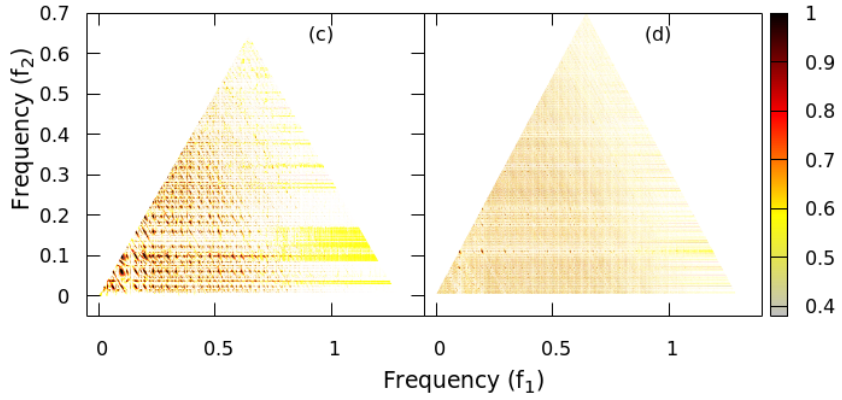
One of the possible dynamical states that can occur in systems that have two independent frequencies of oscillations is the strange non chaotic state. We discussed this briefly in chapter 1. A system is said to be exhibiting strange non chaotic behavior when it possesses a strange attractor with fractal geometry, but does not show divergence of nearby trajectories. The former is characterized by a non integer fractal dimension while the latter is characterized by the absence of any positive Lyapunov exponents.

We study quasiperiodicity and strange non chaotic behavior taking the dynamics of

a doubly driven pendulum as a standard system, already considered in chapter 1 and described by 1.12. This system exhibits quasiperiodicity for K=1.34, V=0.55, p=3.0, $\omega_1=\frac{\sqrt{5}-1}{2}$ and ω_2=1.0 , and strange non chaotic dynamics by changing K to 1.33 [8]. The power spectrum and bicoherence graphs for these regimes are shown in Figure 4.4. In this case we divide the time series into 32 segments of 16,384 points each. Differences in these dynamical states can be easily seen in both the power spectrum and bicoherence. To illustrate the usefulness of $b_F(f)$, we also plot the main peak bicoherence for these states in Figure 4.5. We clearly see that all the peaks in the power spectrum have significant bicoherence. The error on the bicoherence for N segments is given by $\frac{1}{N}$ [126].



(a)



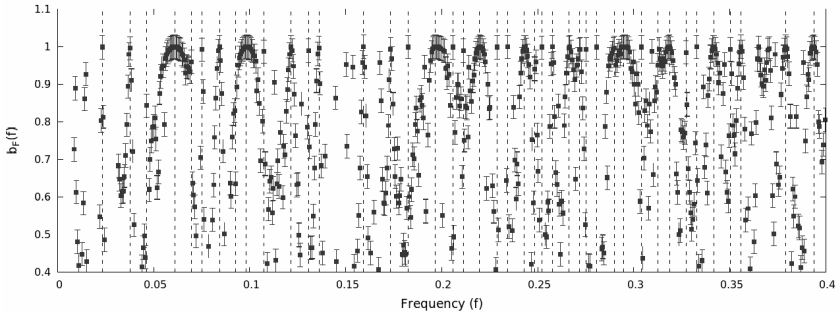
(b)

Figure 4.4: Power spectrum for the doubly driven pendulum in (a)quasiperiodic and (b)strange non chaotic regimes, normalised to the range $[0, 1]$. The corresponding bicoherence graphs are shown in (c) and (d).

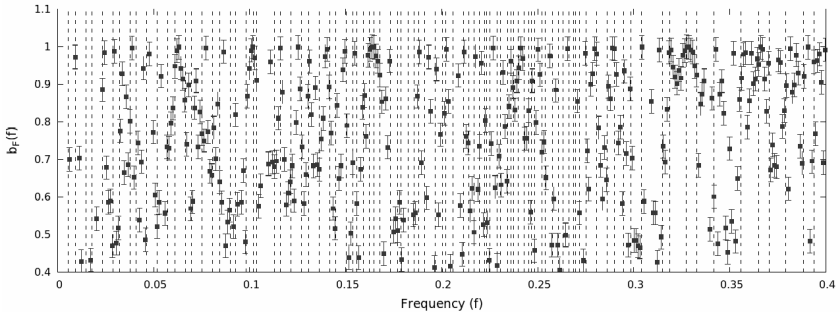
One of the most popular methods to detect strange non chaotic behavior in time series data has been through the spectral scaling of peaks in the strobed power spectrum [76, 127]. In this, the time series is first strobed[3] along its primary frequency. The power spectrum of this strobed time series is then calculated. The number of peaks[4], N_p is counted for a fixed threshold power. This threshold power is increased and the number of

[3]Strobing is the process of resampling the time series along a desired frequency. In this case the data is strobed along the primary frequency. This is equivalent to taking a Poincare section of the time series.

[4]A peak is defined as a local maxima in the power spectrum. This helps avoid artifacts due to finite peak widths in the power spectrum



(a)



(b)

Figure 4.5: Main peak bicoherence for the doubly driven pendulum in (a)quasiperiodic and (b)strange non chaotic regimes. The vertical lines represent the peaks in the power spectrum with power above 0.001. We see that in both cases, all the peaks in the power spectrum have significant main peak bicoherence as well.

peaks is counted each time. The variation of the number of peaks with threshold power can be shown to have a power law scaling, for strange non chaotic behavior [128]. We show that a system exhibiting quasiperiodicity, when contaminated with colored noise can show very similar scaling behavior.

We show that this is the case for the quasiperiodic pendulum when contaminated with red noise with a signal to noise ratio[5] of 1.15. The time step for integration of the system is chosen as one tenth of the primary frequency, $F_1(\frac{2\pi}{10\Omega_1})$. The time series of the y variable, given by $\frac{d\phi}{dt}$, is strobed by picking every tenth point. The power spectrum of this strobed time series for 32 segments is shown in Figure 4.6. The scaling of number of peaks with changing threshold is then calculated. We use logarithmic binning[6] to avoid statistical errors in the tail of the distribution [129]. The scaling of peaks for the noiseless strange non chaotic case and the noise contaminated quasiperiodic case is shown in Figure 4.7. The similarity in the scaling indices between the two is immediately obvious, and may lead to false conclusions regarding the underlying dynamics of the system. Hence, alternate methods must be employed to differentiate between the two cases. We suggest the method of applying a main peak bicoherence filter to consider only the peaks of dynamical origin in the counting. Recollect from Figure 4.5 that all the peaks in the strange non chaotic power spectrum also have significant bicoherence.

One of the main hurdles in the use of the main peak bicoherence filter, when dealing

[5]Defined as the ratio of standard deviation of the signal to the standard deviation of the noise

[6]In this the bin sizes are not chosen to be equal. Rather, they are chosen as powers of a number, n. Hence, the bins are equally spaced on a logarithmic axis(See Figure 4.7, for instance).

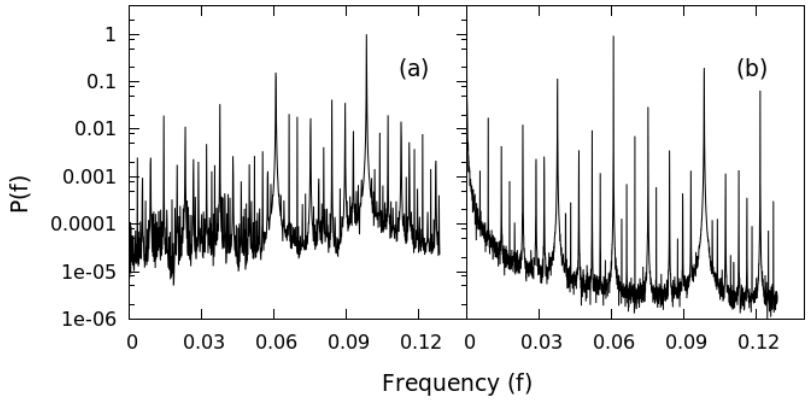


Figure 4.6: Power spectra for the y time series of the driven pendulum in (a)noiseless strange non chaotic state and (b)quasiperiodic state contaminated with red noise.

with strobed time series is the small value for the Nyquist frequency, f_m[7]. We get around this problem, by considering the significance of $b_F(f_m - F)$, if $f + F > f_m$. Hence we check if

$$b_F(f) > \sqrt{\frac{9.8}{2N}} \qquad , F + f < f_m$$
$$b_F(f_m - F) > \sqrt{\frac{9.8}{2N}} \qquad , F + f > f_m$$

before counting the peaks above a threshold in the power spectrum. The scaling behavior of the strange non chaotic state is shown in Figure 4.8(a). We see the scaling is identical before and after applying the filter. On the other hand the scaling of quasiperiodic power spectrum when contaminated with red noise is shown in Figure 4.8(b). One sees a distinct difference in the scaling before and after application of the filter in this case. We find that the filter reduces the number of peaks by a factor greater than seven at the smallest threshold [130]. Our method can hence be used to identify quasiperiodicity, even in cases when the system is contaminated by noise. We point out that the nature of the bicoherence function would make this filter applicable, irrespective of the color of the noise. We also point out that this filter may be useful in filtering out dynamical peaks even when the system evolves in the presence of noise [131].

4.3 Dynamics of RR Lyrae Stars

In this section we will examine the dynamics of RR Lyrae stars in more detail, especially in the light of the analysis considered above. One of the primary drawbacks of nonlinear analysis like the bicoherence, has been the need for long continuous high quality light curve data[8]. This has been made available with the advent of space telescopes for observations. The Kepler space telescope is one such telescope [132, 133]. With its advent, the understanding of the underlying nonlinearity in the dynamics of many stars has improved considerably. Some important results to this end has been the discovery of period

[7]For a strobing frequency F_s, the Nyquist frequency is $\frac{F_s}{2}$.

[8]The time series of the intensity variation of an astrophysical object is called its light curve

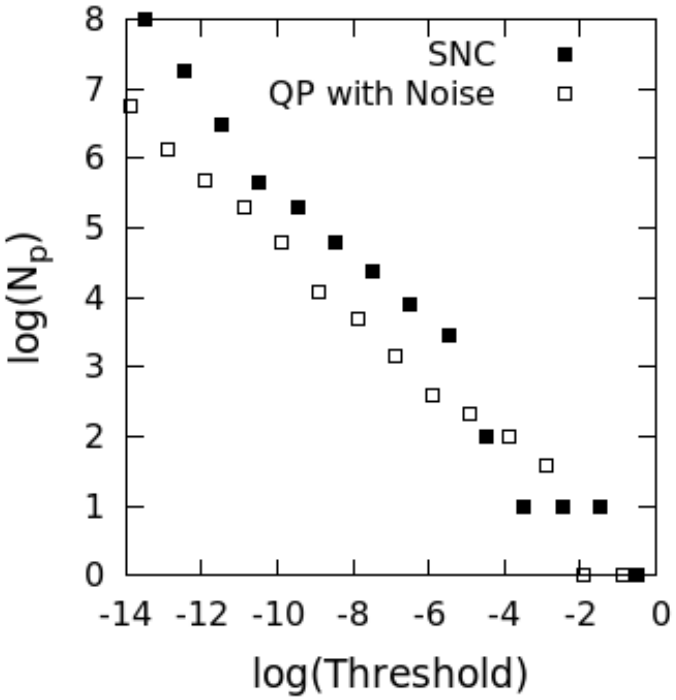


Figure 4.7: Scaling of peaks in the strobed power spectrum for the noiseless strange non chaotic (SNC) state and the quasiperiodic(QP) state contaminated with red noise. Both show similar power law like scaling.

doubling in RR Lyrae stars [73, 74],strange non chaotic behavior in RRc Lyrae stars [76], chaotic behavior in RV Tauri stars [134] etc.

4.3.1 RRab Lyrae Stars

RRab Lyrae stars are subclasses of RR Lyrae stars that pulsate in the radial fundamental mode, with long to moderately long periods [61]. These stars have a strong periodicity and were thought to be perfectly periodic until recently. With the Kepler space telescope, many RRab Lyrae stars were shown to have undergone period doubling [74]. Hydrodynamic simulations of RR Lyrae stars show possibilities of rich dynamics like period doubling, intermittency and chaos [70, 135, 136].

We explore signatures of chaos in the Kepler light curves of RRab Lyrae stars using the bicoherence function. We first compute the power spectrum from the star light curves. Specifically we consider two stars KIC 4484128 and KIC 7505345. Since the datasets come with gaps in observations, the calculations are done by averaging over n evenly sampled segments. The power spectra for the two stars are shown in Figure 4.9. Other than the peaks at the fundamental, harmonics and half harmonics, we also see minor peaks at other frequencies. These may be stochastic peaks or may be of chaotic origin. We explore this question further by considering the bicoherence for these stars. The full bicoherence planes for these stars are shown in Figure 4.10. We observe that the bicoherence plane has significant bicoherence at many points, besides the main peak, half harmonics and harmonics, suggesting that these stars may be exhibiting richer dynamics than simple period doubled behavior. We illustrate this specifically in the case of the star KIC 4484128, by plotting the main peak bicoherence along with the prominent power

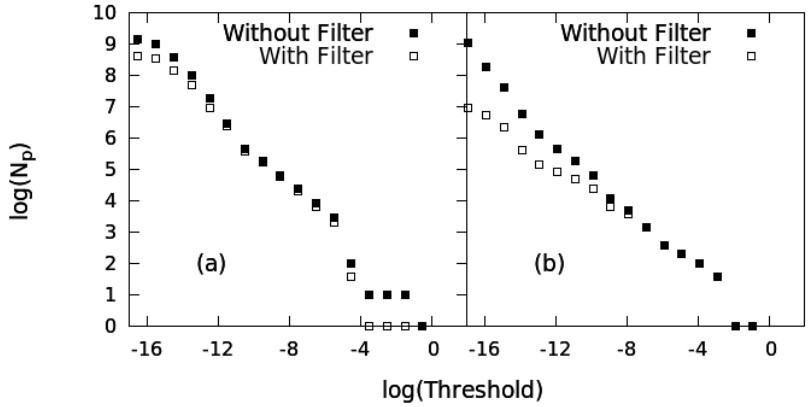


Figure 4.8: Scaling of peaks in the power spectrum with and without bicoherence filter for the (a)noiseless strange non chaotic state and (b)the quasiperiodic state contaminated with red noise.

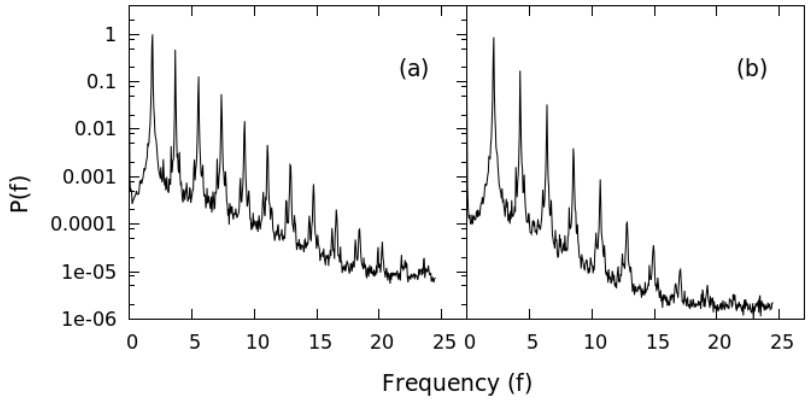


Figure 4.9: Power spectrum graphs for RR Lyrae stars (a) KIC 4484128 and (b)KIC 7505345.

spectral peaks in Figure 4.11.

In an attempt to understand the dynamics of these stars better, we also undertake correlation dimension analysis of these stars, along with its surrogates. The plots for the same are shown in Figure 4.12. The surrogates are generated separately for the individual segments and joined together to act as a surrogate for the whole time series, as described in Section 4.2 in chapter 3. We see clearly that in both cases the data deviates from surrogates, but the deviation is larger for KIC 4484128. We also show the $f(\alpha)$ plots for these stars in Figure 4.13. The D_2 values and the parameters of the $f(\alpha)$ spectrum for these stars are listed in Table 4.1. One sees that the width of the $f(\alpha)$ curve in the case in the case of KIC 7505345 is very narrow, similar to noise.

Table 4.1: D_2,α_1,α_2, γ_1 and γ_2 values for the stars considered in Figure 4.9.

KIC	D_2	α_1	α_2	γ_1	γ_2
4484128	2.88	1.69	2.10	0.31	.0005
7505345	3.81	1.65	1.69	1.0	.25

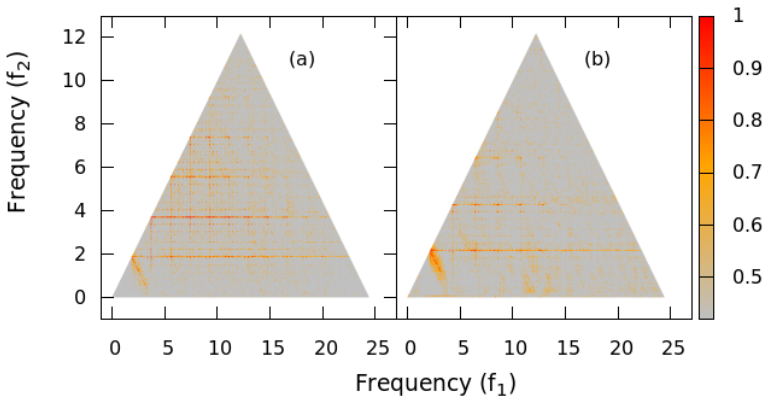


Figure 4.10: Full bicoherence plane for RR Lyrae stars (a) KIC 4484128 and (b)KIC 7505345.

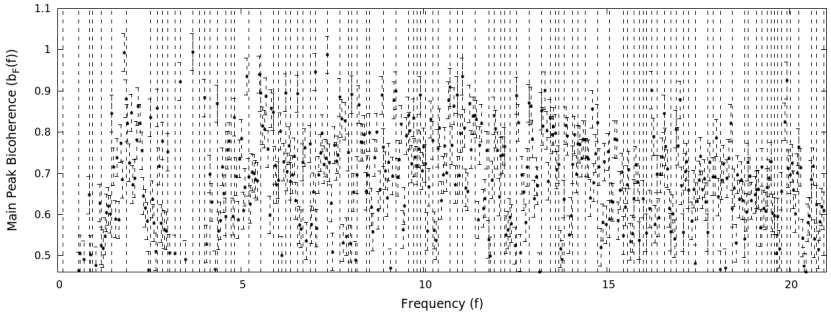


Figure 4.11: Main peak bicoherence function for RR Lyrae star KIC 4484128. The vertical lines indicate the power spectral peaks.

RRc Lyrae Stars

In this section, we analyze the dynamics of RRc Lyrae variable stars in the Kepler field of view. Four RRc Lyrae stars fall into this category, namely KIC 4064484, KIC 5520878, KIC 8832417 and KIC 9453114. Spectral scaling analysis of these stars had lead to the conclusion that all four may be exhibiting strange non chaotic variation [76]. Subsequently, spectral scaling of the power spectrum without strobing, was also shown to have similar power law scaling for these stars [137]. In fact this difference in spectral scaling between RRc and RRab Lyrae stars was suggested as a method of differentiating between the two subclasses [138].

We first examine the power spectra of these four stars. As before the power spectrum is calculated for n evenly sampled segments using the FFT algorithm. The plots for the same is shown in Figure 4.14. One sees immediately that all four stars show a rich spectrum, with peaks an interpeak continuum. We explore the question of whether these peaks are stochastic as in the noise contaminated quasiperiodic case or dynamic as in the strange non chaotic case. The values of the primary and secondary peaks and their ratio is shown in Table 4.2 [73]. We notice that all the stars seem to exhibit an irrational ratio for primary to secondary frequency, very close to the golden ratio. In order to check whether the peaks are of dynamical origin, we first consider the full bicoherence graphs corresponding to these stars. This is shown in Figure 4.15. We see that the bicoherence planes for the stars KIC 4064484 and KIC 9453114 are distinctly different from the stars KIC 5520878 and KIC 8832417. We quantify this difference by defining a

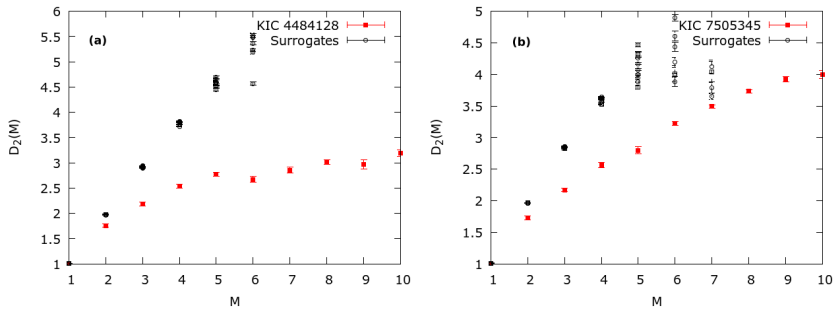


Figure 4.12: Correlation dimension (D_2) vs embedding dimension (M) plots for RR Lyrae stars (a) KIC 4484128 and (b)KIC 7505345, along with 10 surrogates.

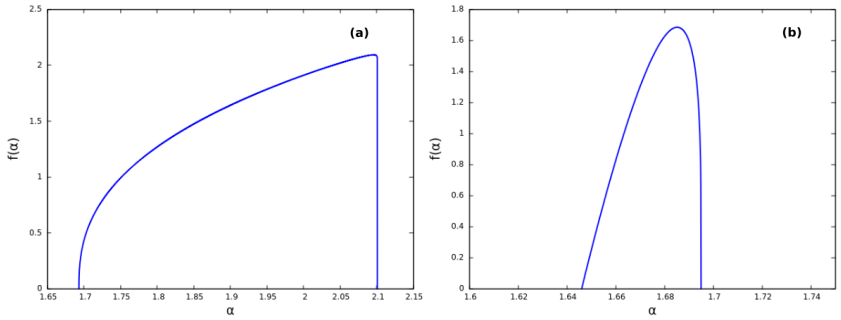


Figure 4.13: $f(\alpha)$ vs α plots for RR Lyrae stars (a) KIC 4484128 and (b)KIC 7505345.

quantifier called the significant pair fraction. This counts the relative number of peaks with significant bicoherence. We list this in Table 4.2, along with the average value of bicoherence through the plane. This confirms that the bicoherence is higher through the plane for KIC 4064484 and KIC 9453114, as compared to KIC 5520878 and KIC 8832417.

As representative cases, we study the main peak bicoherences of KIC 4064484 and KIC 5520878. We see that the main peak bicoherences too reflect the lack of quadratic coupling between the peaks in the power spectrum for KIC 5520878. KIC 4064484 on the other hand shows significant bicoherence for almost all the peaks present in the power spectrum. This is shown in Figure 4.16. Finally, we analyze the spectral scaling of the peaks in KIC 4064484 and KIC 5520878, in the presence of the bicoherence filter developed in Section 4.2. We see that while the scaling behavior with the filter is retained for KIC 4064484, it changes drastically for KIC 5520878, as seen in Figure 4.17. This indicates that KIC 4064484 is indeed undergoing strange non chaotic behavior as suggested by [76], while KIC 5520878 is not. In order to explore the dynamics of KIC 5520878 further, we also consider the tricoherence along the primary power spectral peak for this star. However no significant tricoherence was observed either. This suggests that the dynamics in KIC 5520878 may actually be noise contaminated quasiperiodicity. Similar results were obtained for the other two stars. KIC 8832417 showed behavior similar to KIC 5520878, with a drastic difference in the spectral scaling when a bicoherence filter was applied. However, KIC 9453114 did not show any such change, indicating strange non chaotic dynamics similar to KIC 4064484 [131].

The results of the bicoherence analysis done suggests that the stars KIC 4064484 and KIC 9453114 have distinctly different dynamics from KIC 5520878 and KIC 8832417. We explore this further, and see if this is reflected in other properties of these stars. We first check the astrophysical properties and observe that the two populations differ in three

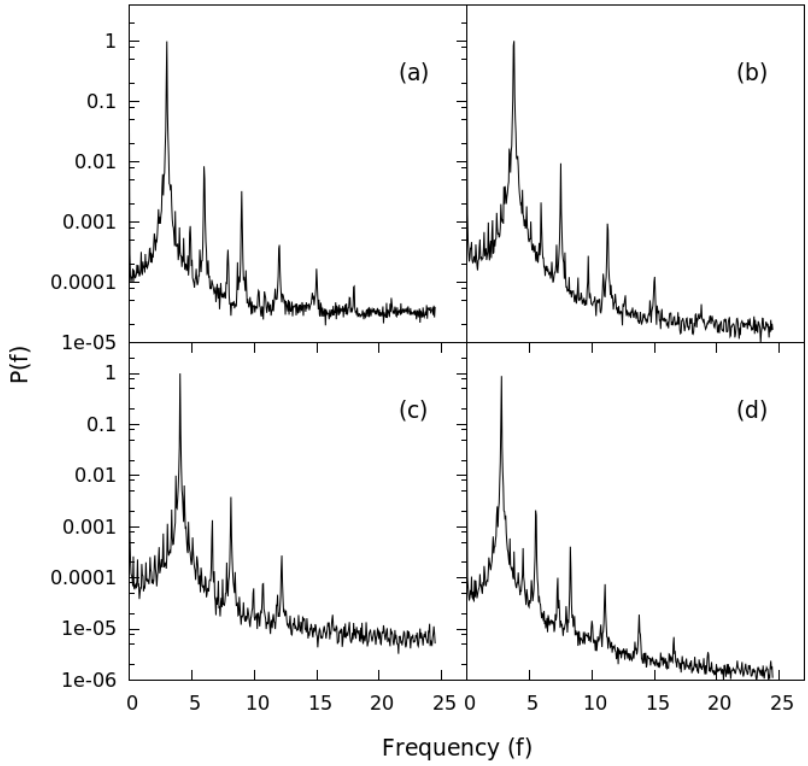


Figure 4.14: Power spectrum graphs for RRc Lyrae stars (a)KIC 4064484, (b)KIC 5520878, (c)KIC 8832417 and (d)KIC 9453114.

Table 4.2: Primary period(P_1), secondary period (P_2), the ratio of the primary to secondary, significant pair fraction(SPF), mean bicoherence(B_{avg}), effective temperatures(T_{eff}) and metallicities($[Fe/H]$) of the RRc Lyrae stars in the Kepler field of view.

KIC	$P_1(days)$	$P_2(days)$	$\frac{P_2}{P_1}$	SPF	B_{avg}	T_{eff}	$[Fe/H]$
4064484	0.337	0.207	0.615	0.77	0.64	6500	-1.58
5520878	0.269	0.170	0.631	0.31	0.40	7250	-0.18
8832417	0.248	0.152	0.612	0.24	0.33	7000	-0.27
9453114	0.336	0.224	0.614	0.70	0.55	6500	-2.13

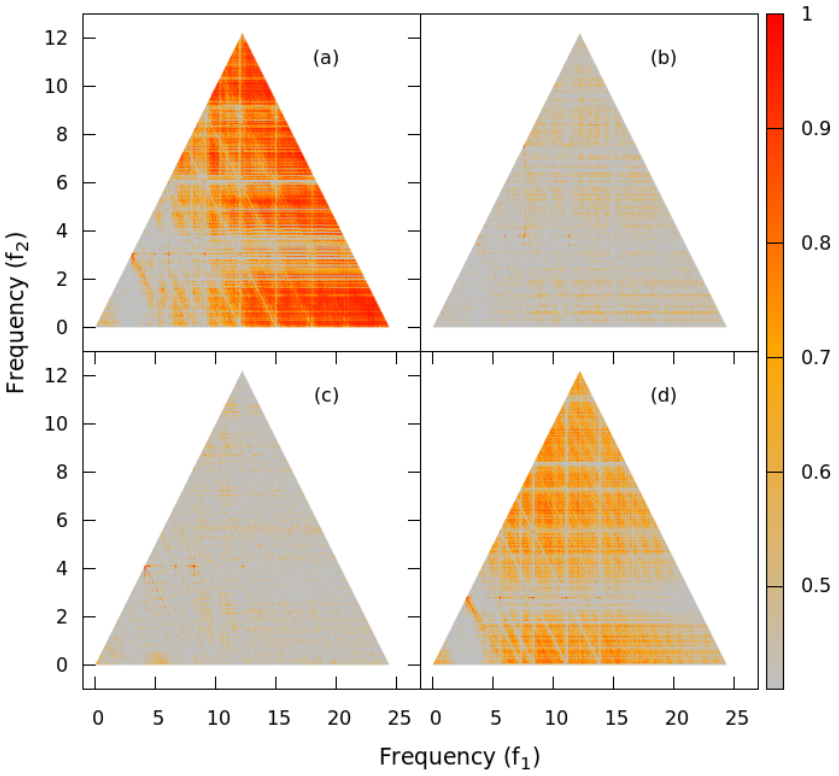


Figure 4.15: Full bicoherence graphs for RRc Lyrae stars (a)KIC 4064484, (b)KIC 5520878, (c)KIC 8832417 and (d)KIC 9453114.

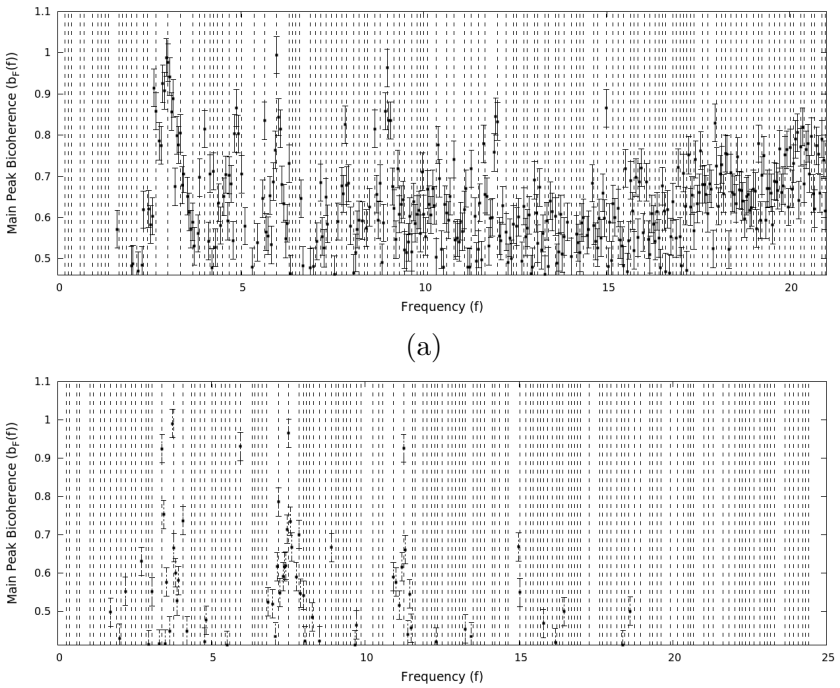


Figure 4.16: Main peak bicoherence for the Kepler RRc Lyrae stars (a)KIC 4064484 and (b)KIC 5520878. The vertical lines represent the peaks in the power spectrum. Almost all the power spectral peaks have significant bicoherence in (a), while very few of them show significant bicoherence in (b).

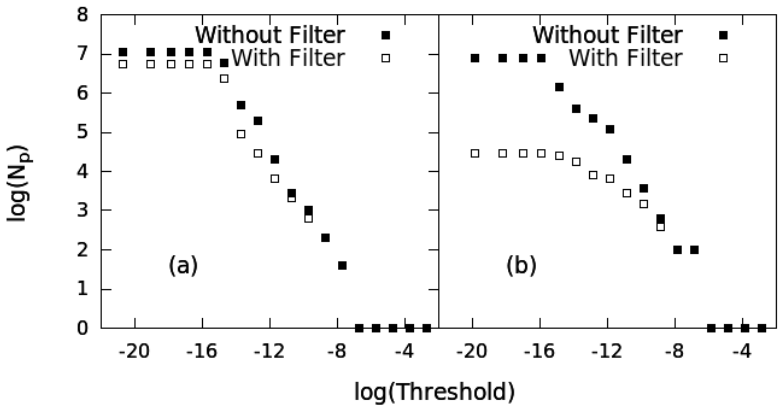


Figure 4.17: Spectral scaling with and without bicoherence filter for RRc Lyrae stars (a)KIC 4064484 and (b)KIC 5520878. (a) retains the scaling behavior after applying the filter, while (b) shows distinctly different behavior.

main parameters. These are the primary period P_1, effective temperature T_{eff} and the metallicity measured through $[Fe/H]$[9] [73, 75](See Table 4.2). We explore the nonlinear properties of these stars further, using recurrence quantification analysis in chapter 6.

4.4 Summary and Discussion

In this chapter, we present the results of our study on the bicoherence function in various forms to deal with noisy data. We primarily look at two cases when periodic data can be misinterpreted as being chaotic or strange non chaotic when contaminated with noise. The kind of noise may be additive or the system itself may be evolving in the presence of noise.

When the system evolves with noise, we show that limit cycles may seem similar to chaos. However, the bicoherence planes for the two cases help us differentiate between these two states. We subsequently show how the bicoherence may be used to identify the underlying dynamics in RRab Lyrae stars where the power spectrum shows an interpeak continuum which may be of chaotic or stochastic origin. Our conclusions suggest that RRab Lyrae stars may be undergoing chaotic dynamics. We supplement the bicoherence studies with D_2 and $f(\alpha)$ studies.

The second scenario where periodic behavior is misinterpreted is in quasiperiodically forced systems. We show that a system in quasiperiodic state, when contaminated with noise shows behavior similar to strange non chaotic behavior, when only conventional techniques are considered. We point out that the spectral scaling of peaks in the strobed power spectrum is similar for strange non chaotic and noisy quasiperiodic systems. These two distinct dynamical behaviors can be distinguished if one uses a bicoherence based filter. The filter, we see, recovers the true scaling behavior, even when the system is contaminated with noise.

We use this analysis to examine the dynamics of RRc Lyrae stars which are thought to exhibit strange non chaotic behavior. We show that two distinct groups can be observed in RRc Lyrae stars based on their bicoherence. One group consisting of KIC 4064484 and KIC 9453114 seem to be exhibiting strange non chaotic dynamics while KIC 5520878 and KIC 8832417 seem to be exhibiting quasiperiodic dynamics. Further, these groups also seem to correspond to distinct classes when astrophysical properties like the effective temperature, metallicity and primary period are considered. We suggest that these stars may be further subclassified using their nonlinear dynamical properties. We explore these stars further using recurrence based analysis in chapter 6.

[9]Measured as $[Fe/H] = log_{10}(\frac{N_{Fe}}{N_H})_{star} - log_{10}(\frac{N_{Fe}}{N_H})_{sun}$

Chapter 5

Nonlinear Dynamics of Contact Binary Stars

In this chapter we will consider the nonlinear properties of overcontact binary stars. Overcontact binary stars are binaries in which both component stars overfill their Roche lobes. This implies that both matter and energy can be transferred between the companions. Hence the light curves of overcontact binary stars are not purely periodic, but exhibit irregularities in the form of uneven maxima and eclipse time variations. The origins of these are still not well understood and are attributed to a number of different factors. The nonlinear properties of overcontact binary stars remain relatively unexplored. On the other hand, one of the striking examples where nonlinear time series analysis has been successfully implemented in astrophysics is in the study of compact binary stars [87, 139, 140]. In this chapter we will attempt to understand the light variations of overcontact binary stars by systematically studying their nonlinear dynamics. We will also attempt to correlate the astrophysical properties of these stars with their nonlinear dynamical properties.

5.1 Dynamics of overcontact binary stars

As a preliminary step, we compute the power spectrum of four sample over contact binary stars. These are shown in Figure 5.1. We immediately notice the peaks at half integer multiples of the primary peak. This is indicative of period doubling in the underlying dynamics. Similar period doubling behavior has been observed in a number of different stars including RR Lyrae stars, W Virginis stars, black holes and so on [74,141,142]. The existence of half integer peaks in contact binary stars seems to suggest a strong indication of nonlinear dynamics in their light curves. Hence we apply nonlinear time series methods to get measures of nonlinearity in them.

5.1.1 Correlation Dimension

In this section we will calculate the correlation dimension for the 463 overcontact binary stars in the Kepler field of view [143]. To eliminate the wide differences arising from the amplitude distributions of the different stars, we first convert the amplitude distributions to uniform distributions. This is done by taking the uniform deviate of the light curve. We then proceed to reconstruct the state space using Taken's theorem. If $I(t)$ is the intensity of the star at time t and $I_u(t)$ is its uniform deviate, the M dimensional delay

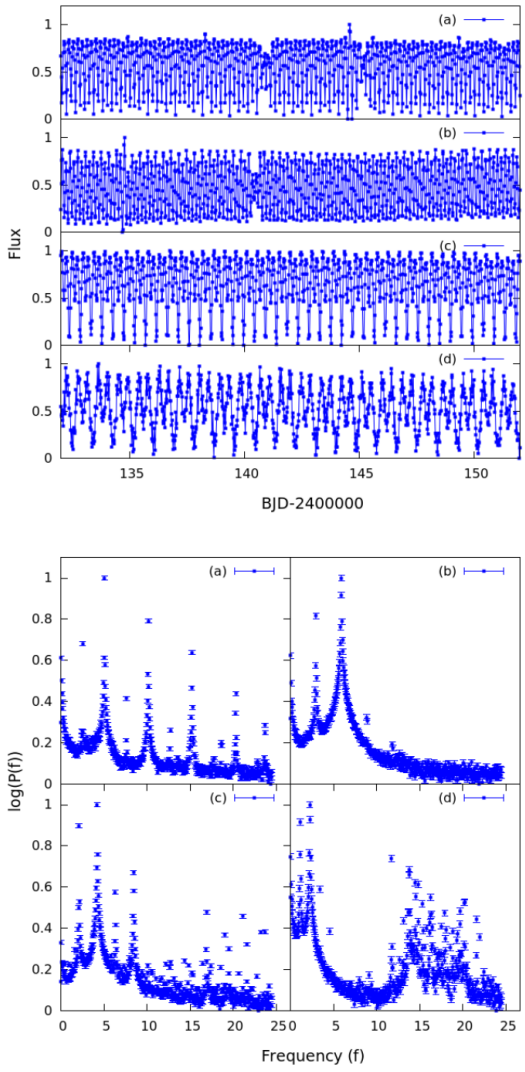


Figure 5.1: Light curves and power spectra of four typical over-contact binary stars ((a)KIC 4909422, (b)KIC 6368316, (c)KIC 7657914, and (d)KIC 8800998). The upper panel corresponds to the light curves and the lower panel shows the corresponding power spectra.

vectors are given by

$$\vec{v}_i = [I_u(t_i), I_u(t_i + \tau), ..., I_u(t_i + M\tau)] \tag{5.1}$$

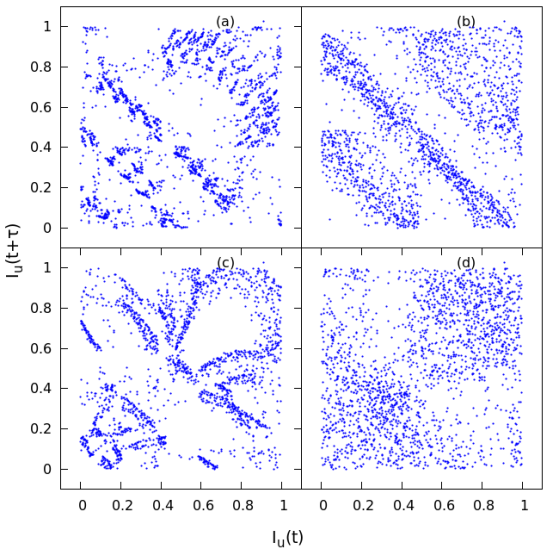


Figure 5.2: Reconstructed state space of four typical over-contact binary stars projected in 2D ((a)KIC 4909422, (b)KIC 6368316, (c)KIC 7657914, and (d)KIC 8800998)

The reconstructed state space projected to two dimensions for the four sample stars are shown in Figure 5.2. One immediately sees a structure in the state space of these stars, arising from the uneven maxima and eclipse time variations described above. We quantify this structure using the correlation dimension, D_2. A histogram of the values of D_2 for all 463 stars is shown in Figure 5.3. We see from the figure that the D_2 shows a range of values for the stars under consideration.

We also generate IAAFT surrogate datasets for all the light curves as described in previous chapters [144, 145]. Surrogate datasets are constructed by constructing individual

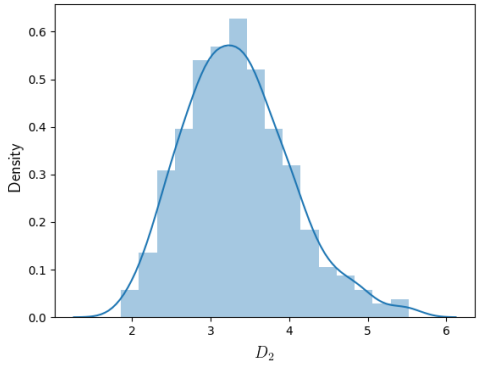


Figure 5.3: Histogram of D_2 values for all the overcontact binary stars under consideration.

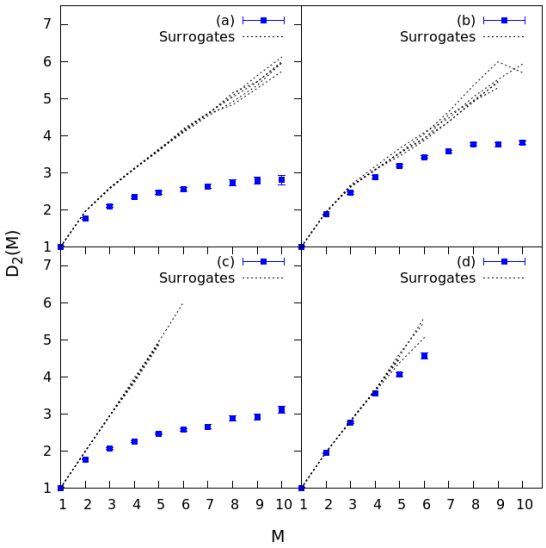


Figure 5.4: D_2 vs M curves for four typical over-contact binary stars ((a)KIC 4909422, (b)KIC 6368316, (c)KIC 7657914, and (d)KIC 8800998). We see the data is distinctly different for surrogates for (a),(b) and (d), while for (d) data and surrogates co-vary.

surrogates for the evenly sampled segments that are joined together to form a surrogate for the entire dataset. The $D_2(M)$ vs M curves for the four sample datasets and their surrogates is shown in Figure 5.4. We also calculate the *nmsd* measure for all the stars considered.

We make an important observation regarding the D_2 and *nmsd* for the stars under consideration, by plotting the D_2 with *nmsd*. This is shown in Figure 5.5. We see that the D_2 and *nmsd* are inversely related, such that higher D_2 values come with lower *nmsd*. One of the reasons behind lower *nmsd* values is the noise contamination in the system. Hence our analysis may suggest that systems with higher D_2 tend to have stochastic factors influencing their dynamics. We analyze this further as we consider the bicoherence studies of these stars.

Eclipse Time Variation

Periodic phenomena are very common in nature. However, very often the periodicities are not exact. Hence these phenomena are only nearly periodic. In such cases one of the properties that is widely studied is the variation or deviation from the period. This analysis is called by different names in different contexts. For instance in cardiology, the heart rate variability is studied and in ecology variations in flowering time and population dynamics are studied [146–148]. The origins of these variations are often explained using nonlinear dynamics [149–151]. In the context of eclipsing variable stars this deviation from the primary period is called eclipse time variation. The general method to study this phenomena is through the study of O-C curves (Observed-Calculated).

In order to analyze the eclipse time variations in these stars we generate O-C curves using the following equation.

$$\Delta = T_i - T_0 - i \times P^s \tag{5.2}$$

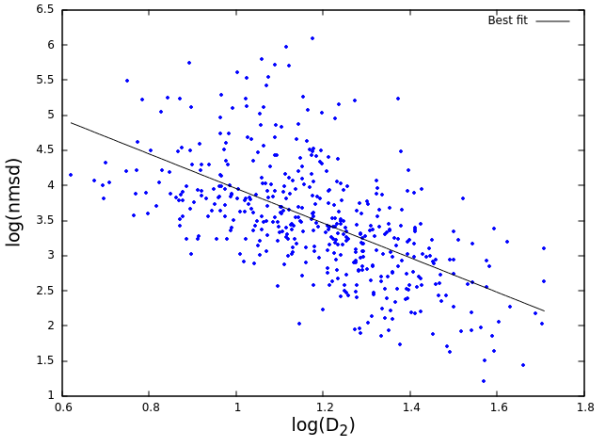


Figure 5.5: Variation of the log of the *nmsd* with log of D_2. We see a fall in *nmsd* as D_2 increases. Hence the increase D_2 may be attributed to an increase in stochasticity or a change in the underlying equations of the system.

Table 5.1: D_2 and *nmsd* values for the O-C curves of the four overcontact stars shown in Figure 5.1, and two chaotic dynamical systems.

Kepler ID	D_2	*nmsd*
4909422	1.74	2.77
6368316	1.50	3.56
7657914	1.92	7.42
8800998	1.74	2.05
Rössler	3.84	1.63
Pendulum	3.39	3.33

Here, T_i is the observed time of the i^{th} maxima, T_0 is the initial time of observation, and P^s is the mean period of occurrence of the maxima.

Eclipse time variation curves of overcontact binary stars were previously reported as showing random walk like behavior. Here we explore this further by considering the O-C curves generated by two standard chaotic systems with strong periodicities, namely the driven pendulum and the Rössler system. We see that the O-C curves generated from these systems too show random walk like characters. This is evident from the power spectrum, which shows $\frac{1}{f^2}$ characteristics(Figure 5.6). We see similar characteristics for the O-C curves generated from the overcontact Kepler stars.

Finally, we also consider the D_2 for the O-C curves generated from the standard systems and the overcontact binaries. We compare them to the five IAAFT surrogate datasets generated. The D_2 and *nmsd* for the O-C curves generated from the sample stars considered and the standard systems are shown in Table 5.1. The similar values of *nmsd* seems to suggest that the eclipse time variations may also be of deterministic origin.

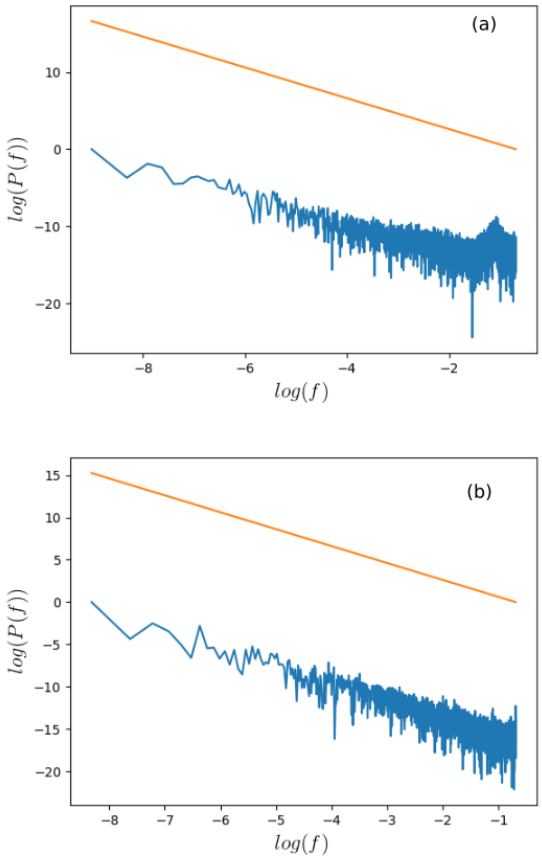


Figure 5.6: Power spectra for the eclipse time variation for (a)Rössler system and (b)KIC 4909422. The $\frac{1}{f^2}$ line is shown in both cases for comparison.

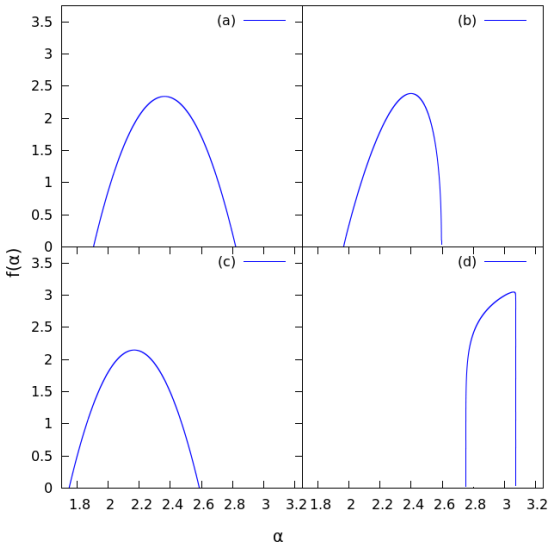


Figure 5.7: $f(\alpha)$ vs α plots for the 4 sample eclipsing binary stars considered. The figures correspond to (a)KIC 4909422, (b)KIC 6368316, (c)KIC 7657914, and (d)KIC 8800998. The narrow spectrum corresponding to (d) is indicative of noisy behavior, as suggested by Figure 5.4

5.1.2 Multifractal Properties

As mentioned in earlier chapters, multifractal spectra or $f(\alpha)$ curves are also calculated for the whole dataset of overcontact binary stars. We can characterize the $f(\alpha)$ curve using the function given in equation 1.48. The $f(\alpha)$ curves can be characterized completely using the four parameters α_{max}, α_{min}, γ_1 and γ_2.

The embedding dimension for calculation of the $f(\alpha)$ properties is chosen to be $M = 4$, as over 80% of D_2 values lie below this range in Figure 5.3. The calculated $f(\alpha)$ curves for the four sample stars considered is shown in Figure 5.7. α_{max} corresponds to the rarest parts of the attractor and as such needs a very large number of points for an accurate calculation. Hence the calculation of this quantifier may be susceptible to numerical errors.

5.1.3 Bicoherence Properties

Both the earlier subsections dealt with the fractal properties of the reconstructed state space of the overcontact binaries. In this subsection, we will look at the bicoherence properties, which measures the extent of coupling between the different frequency pairs[1] present in the system. The full bicoherence planes corresponding to the four sample stars we considered is shown in Figure 5.8.

As discussed in chapter 4, in many cases studying the main peak bicoherence function suffices [131].The main peak bicoherence function $b_F(f)$ is given by the equation 4.2. For overcontact binary stars, F is the eclipsing frequency. The $b_F(f)$ for the four sample stars under consideration is shown in Figure 5.9. In order to extract a single number,

[1]Upto quadratic order!

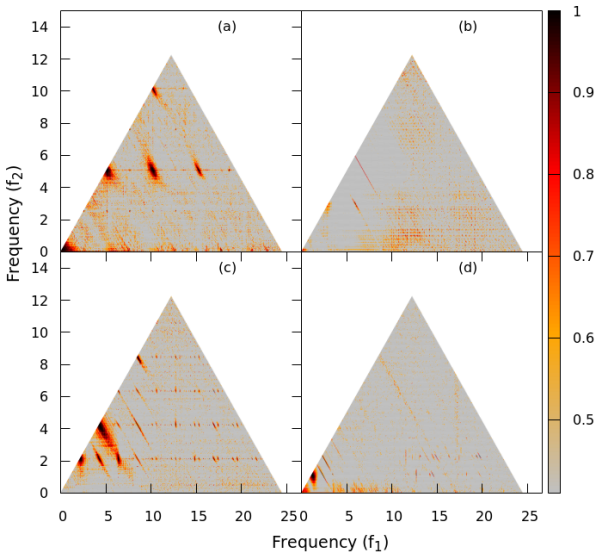


Figure 5.8: Full bicoherence plots for the four typical eclipsing binary stars considered before. Very few frequency pairs show significant bicoherence in (b) and (d), whereas (a) and (c) show significant bicoherence for many frequency pairs.

representing the extent of contact we define a quantity called the significant bicoherence fraction. It is the number of frequencies that have a significant (above 99%) bicoherence with the eclipsing frequency. It is important to note that the eclipsing frequency is equal to the maximal peak in the power spectrum in all cases. The plot of the histogram of the significant bicoherence fraction SBF is shown in Figure 5.10. The range of SBF observed seems to suggest that the coupling with the eclipsing frequency varies widely within the dataset. The stars where no coupling with the eclipsing frequency is seen, seem to be primarily dominated by stochastic factors or with a nonlinearity that is not of quadratic order.

5.1.4 Correlations Between Astrophysical and Nonlinear Properties

The results presented in the previous sections establish the nonlinearity in the light curves of overcontact binary stars. The quantifiers we get suggest that the stars may be exhibiting low dimensional chaotic behavior. We see that almost all the stars show this low dimensional chaos irrespective of the physical properties they possess. It then seems plausible that the nonlinear properties of the star may evolve as the physical properties of the star evolves. We address this question in this section by hunting for correlations between the nonlinear dynamics of a star and its astrophysical properties.

An important property of an overcontact binary star is the extent of contact that exists between the component stars. This is measured using the fill-out factor, ff. We discussed the fill-out factor in chapter 2, described in equation 2.4. ff is also associated with the evolution of a binary, with a higher ff indicative of a more evolved star [78]. Hence a correlation with ff is indicative of evolution and age of the star.

We first look at the correlation between D_2 and ff. For this we split the stars into two

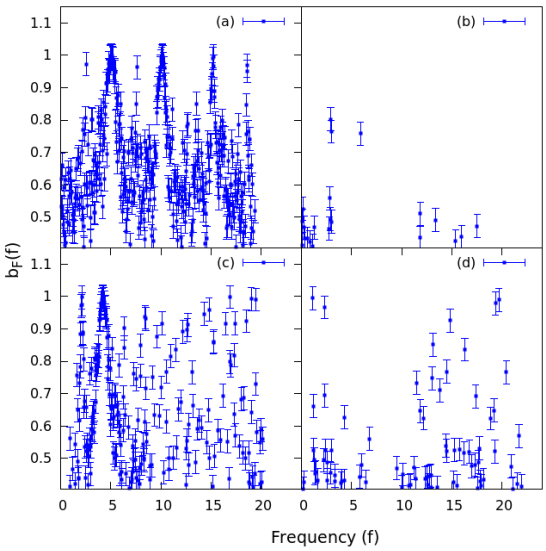


Figure 5.9: $b_F(f)$ vs f plots for the four typical eclipsing binary stars considered above. Only bicoherence values above 99% significance are plotted. As in the case of Figure 5.8, (a) and (c), show significant coupling with the eclipsing frequency whereas (b) and (d) have much less frequencies that show significant value for bicoherence. The error bar on the bicoherence for N segments is given by $\frac{1}{N}$ [126].

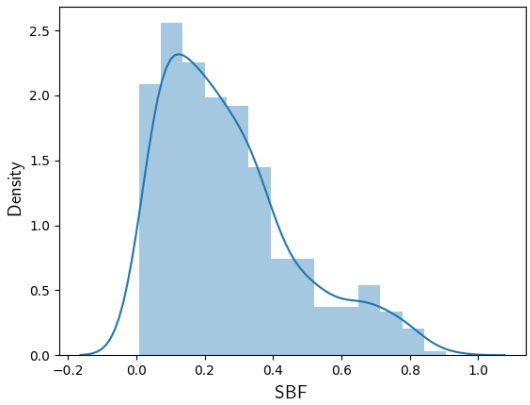


Figure 5.10: Normalized histogram and kernel density estimate of Significant Bicoherence Fraction (SBF) for all the light curves considered.

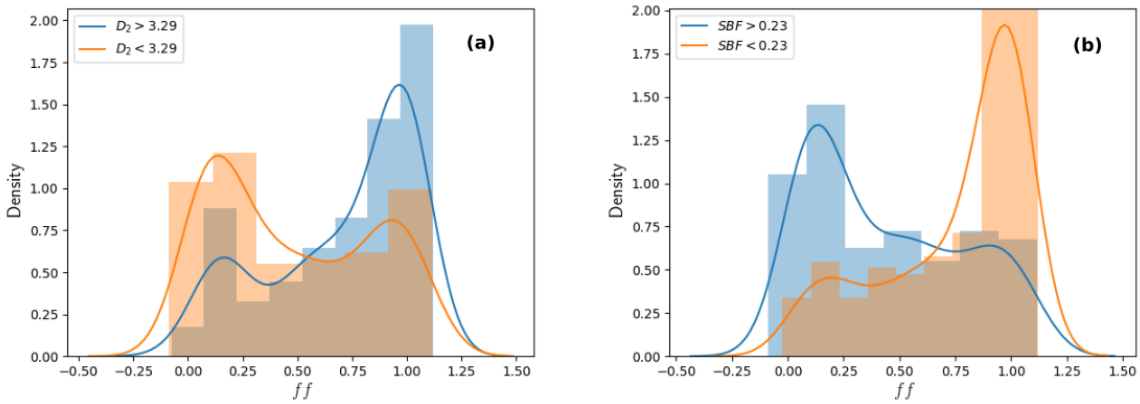


Figure 5.11: Plots of kernel density estimates of fill-out factors for (a)$D_2 < 3.29$ and $D_2 > 3.29$ (b)$SBF < 0.23$ and $SBF > 0.23$. One can see two different distributions when we subdivide the parameters into two. In (a) the skews of the distributions corresponding $D_2 < D_2^{med}$ and $D_2 > D_2^{med}$, towards higher ff and lower ff respectively, suggests that higher D_2 implies a more evolved system with higher ff values. Similarly in (b) we see that the stars start to loose coupling with the eclipsing frequency as it evolves.

categories based on their D_2. This is done by taking the median of the D_2 distribution plotted in Figure 5.3 given by $D_2^{med} = 3.29$. We compare the fill-out factors of the two categories in Figure 5.11a. We find that the stars with a lower D_2 tends to have a lower value for ff, while the stars with higher D_2 have a higher values for ff.

We then consider the Spearman correlation coefficient between D_2 and ff. For two distributions X_i and Y_i, the Spearman correlation coefficient is defined as follows. First the distributions are converted into ranks, giving g_{X_i} and g_{Y_i}. Then the Spearman rho, ρ_S is given as

$$\rho_S = \frac{cov(g_X, g_Y)}{\sigma_{g_X, g_Y}} \tag{5.3}$$

where $cov(g_X, g_Y)$ is the covariance of the rank distributions. rho_S between D_2 and ff is found to be $\rho_S = 0.33$. The significance of this correlation is given by the p-value. One determines this from the value of

$$t = \rho_S \sqrt{\frac{n-2}{1-\rho_S^2}} \tag{5.4}$$

which can be shown to be Student t-distributed with $n-2$ degrees of freedom [13]. The significance essentially gives the probability that the correlation arose from a random distribution. Generally, a correlation is assumed to be significant, if its p-value is less than .001 The p-value for the correlation between D_2 and ff is 6.9×10^{-13}, and is hence likely to be highly significant.

For the correlation between SBF and ff, as before we divide the dataset into two categories according to the median of the SBF distribution. The value of the median is $SBF^{med} = 0.23$. The kernel density plots for the two distributions is given in Figure 5.11b. Unlike for the D_2 case, we see that higher SBF corresponds to lower ff, while lower SBF corresponds to higher ff. This is also reflected in the ρ_S, which is given by $\rho_S = -0.44$ with a p-value of 2.2×10^{-23}. We show this variation of ff with D_2 and SBF in Figure 5.12.

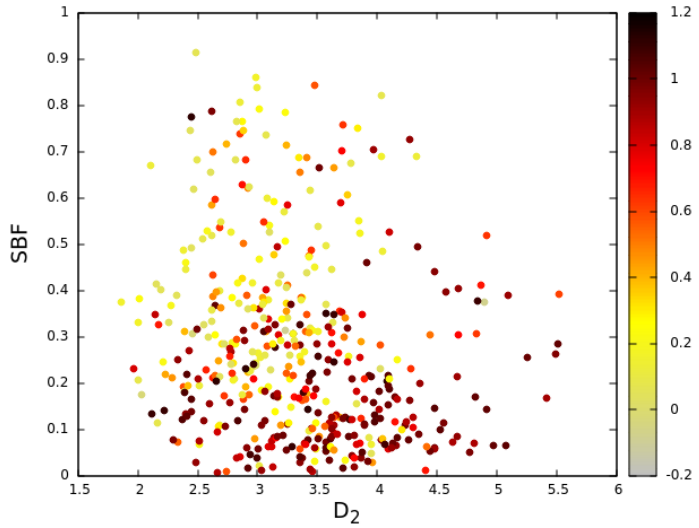


Figure 5.12: Scatter plot of all the overcontact stars as a function of D_2 and SBF. The color code shows the range of ff values. We see that the top left of the graph corresponding to higher SBF and lower D_2, shows a lower ff whereas the bottom right corresponding to higher D_2 and lower SBF shows a higher value for ff.

Miscellaneous correlations

In this section we explore the correlations that exist between different nonlinear time series measures with the physical parameters and between themselves. Such a study is important as it indicates the independent correlations in the study. In addition to the fill-out factor in this section we consider the following physical parameters. These are the effective temperature T_{eff}, the mass ratio, q and the period, P_0. The nonlinear parameters are the D_2, SBF and the parameters of the $f(\alpha)$ spectrum. The table of correlations is shown in table 5.2

We can see for instance that the correlations between ff and D_2 and ff and SBF do not follow from each other as the correlation between D_2 and SBF is much smaller at -0.24. Some of the correlations in the table, especially between P_0 and T_{eff} and between q and ff are already studied in the context of contact binary physics [152, 153]. Such a

Table 5.2: Spearman correlations between D_2, SBF, α_{min}, α_{max}, $\Delta\alpha$, ff, Period(P_0), T_{eff},and mass ratio, q. Correlations that are significant are displayed in bold. We consider a correlation as significant if the p-value is less than .001.

Property	D_2	SBF	α_{min}	α_{max}	$\Delta\alpha$	ff	P_0	T_{eff}	q
D_2	1								
SBF	**-0.24**	1							
α_{min}	**0.40**	**-0.33**	1						
α_{max}	−0.001	0.02	**0.28**	1					
$\Delta\alpha$	**-0.20**	**0.21**	**-0.26**	**0.75**	1				
ff	**0.33**	**-0.44**	**0.20**	−0.08	**-0.16**	1			
P_0	−0.12	**-0.24**	−0.13	−0.04	0.06	0.04	1		
T_{eff}	**-0.21**	**-0.17**	**-0.21**	−0.06	0.07	0.05	**0.61**	1	
q	0.04	−0.02	0.03	−0.04	−0.05	**0.59**	0.03	−0.01	1

table of correlations is important as it helps us identify spurious correlations that arise in data.

Subcategories

Till now we considered overcontact stars without considering differences they may have in physical properties. For instance, multiple stars of the same ff may be very different mass ratios. These could in turn contribute to differences in the underlying physics of the contact binary. For instance stars with high fill out factors and low mass ratios are thought to be progenitors of binary mergers [78]. Hence we will now concentrate on specific subcategories of contact binary stars, based on their mass ratios and effective temperatures.

We first narrow down the stars based on mass ratio. We avoid binaries with skewed mass ratios by taking the range $\frac{1}{2} < q < 2$. This gives rise to an increased correlation between D_2 and ff with $\rho_S = 0.47$ (p-value $= 6.9 \times 10^{-20}$, sample size = 333). ρ_S for SBF and ff increases to -0.50 (p-value $= 9.9 \times 10^{-23}$). Outside the range we consider no correlation remains between D_2 and ff. This implies that for $q < \frac{1}{2}$ and $q > 2$, D_2 and ff are no longer correlated($\rho_S = -0.01$, p-value = 0.911, sample size = 120 stars). Hence stars with highly skewed mass ratios seem to have little correlation between D_2 and ff.

Next we restrict the effective temperature, T_{eff} between 6000 and 7000, while keeping the restriction on q. This increases the correlation between D_2 and ff to $\rho_S = 0.69$ (p-value $= 5.2 \times 10^{-14}$, sample size = 91) [154]. With SBF, ρ_S becomes -0.60(p-value $= 2.6 \times 10^{-10}$). We show the kernel density plots of ff for (a) $D_2 > D_2^{med}$ and $D_2 < D_2^{med}(D_2^{med} = 3.12)$ and (b)$SBF > SBF^{med}$ and $SBF < SBF^{med}$ $(SBF^{med} = 0.21)$, for this case in Figure 5.13. Contact binary stars with the ranges of T_{eff} considered roughly fall into the F spectral class. The stars from these earlier spectral types are thought to constitute the A-type W UMa subclass [155].

We finally take the range, $\frac{2}{3} < q < \frac{3}{2}$ and $6000 < T_{eff} < 7000$. This gives $\rho_S = 0.75$ for D_2 and ff (p-value $= 2.87 \times 10^{-11}$, sample size= 56) and $\rho_S = -.65$ for SBF and ff (p-value $= 7.2 \times 10^{-8}$). These correlations are shown in Table 5.3.

Table 5.3: Spearman correlations of D_2 and SBF with ff and the corresponding p-values, for different sub-populations. ρ_S^1 corresponds to $.5 < q < 2$, ρ_S^2 to $6000 < T_{eff} < 7000$; $.5 < q < 2$ and ρ_S^3 to $6000 < T_{eff} < 7000$; $.67 < q < 1.5$. The sample sizes for the three sub-populations are 333, 91 and 56 respectively.

Measure	ρ_S^1	p	ρ_S^2	p	ρ_S^3	p
D_2	.47	6.9×10^{-20}	0.69	5.2×10^{-14}	0.75	2.87×10^{-11}
SBF	-0.50	9.9×10^{-23}	-0.60	$2.6 \times 10^{-}10$	-0.65	7.2×10^{-8}

One of the instances where the correlations we observe can be put to use is in prediction of approximate values of ff. For large datasets ff is calculated using machine learning algorithms, like neural networks. However using the nonlinear properties may be an alternative approach. Relations like the period-mass correlations derived in [156], can be calculated from the nonlinear dynamical quantifiers. We show this by considering a linear regression of SBF and D_2. We use the best fit line to predict the value of ff. We plot the cumulative distribution of the deviation of the predicted ff from the ff derived in [143] in Figure 5.14. We observe that the fill-out factors of more than 70% of stars can

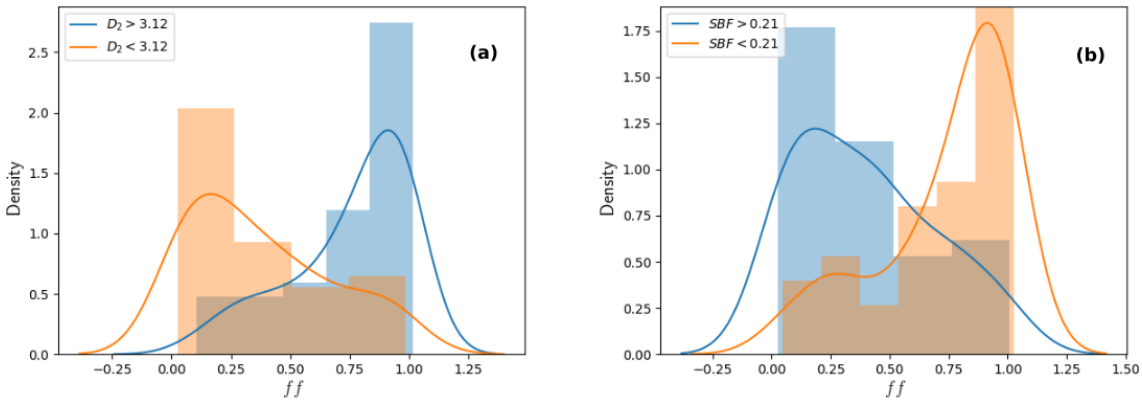


Figure 5.13: Plots of kernel density estimates of fill-out factors for (a)$D_2 > 3.12$ and $D_2 < 3.12$ and (b)$SBF > 0.21$ and $SBF < 0.21$ for $\frac{1}{2} < q < 2$; $6000 < T_{eff} < 7000$. The significantly different distributions in these plots suggests that as one goes into more restricted sub-populations, the dimension of the system and the coupling of frequencies in the system with the eclipsing frequency become more closely linked to the fill-out factor.

be predicted within an accuracy of 0.2 using linear regression within the subcategories considered [157].

5.2 Summary and Discussion

In this chapter we present the results on the quantifiers of the nonlinear dynamics of overcontact binaries, which are close binaries with both companions in contact with each other.

Overcontact binary stars have many partially explained features like eclipse time variations and unequal maxima in their light curves[2]. Hence the light variation is not purely periodic. We check whether the origin of the light variation is chaotic. The possibility of period doubling chaos in these stars is suggested by strong peaks at half integer peaks of the primary frequency evident in the power spectrum. We find a saturating D_2, distinctly different from surrogate datasets in almost all the stars, indicating deterministic chaotic behavior. We also explore the behavior of the time series of eclipse time variations captured by $O-C$ curves. We find similar behavior for the $O-C$ curves generated from period doubling systems like the Rössler and driven pendulum.

Our analysis also indicates that the nonlinear properties of overcontact binaries are related to their astrophysical characteristics. This is especially evident in the correlation observed for D_2 and the bicoherence with the fill-out factor ff. D_2 is positively correlated with the fill-out factor, ff, while SBF is negatively correlated with ff. These correlations become more prominent when we subcategorize these binaries based on mass ratio, q and effective temperature, T_{eff}. We show how this result can be used to derive approximate values for ff. Further these correlations must be reproducible in models or simulations of contact binary light curves and may serve as important checks for the validity or accuracy of binary star models.

[2]Called the O'Connell effect [79] after its discoverer.

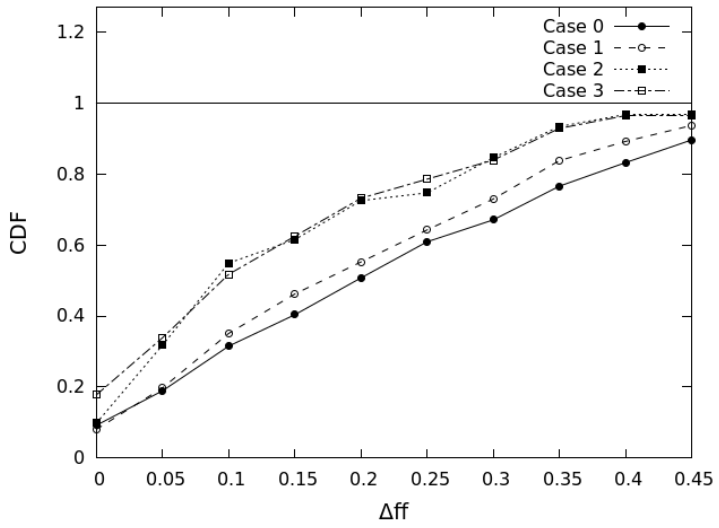


Figure 5.14: Cumulative density function of the deviations of ff from the best fit line for (a) Case 0 : No restrictions on parameters, (b) Case 1 : $\frac{1}{2} < q < 2$ (c) Case 2 : $\frac{1}{2} < q < 2$, $6000 < T_{eff} < 7000$ and (d)Case 3: $\frac{2}{3} < q < \frac{3}{2}$, $6000 < T_{eff} < 7000$. (a) and (d) suggests that, using simple linear regression, we can find a value of ff in restricted sub-populations of binary stars, with reasonable accuracy.

Chapter 6

Recurrence Networks and Analysis of Stars

Recurrence is a property of bounded dynamical systems. The pattern of recurrence of trajectories in state space is captured using a technique called recurrence plot, discussed in chapter 1. One can also construct a network from this plot. The properties of this network would reflect the properties of the underlying recurrence patterns. In this chapter we will use recurrence based analysis to classify different astrophysical objects. Recurrence based analysis has shown promise in differentiating between different kinds of dynamical states [158, 159]. For instance it has been used to capture differences between resting and seizure states in EEG signals, between stochastic and chaotic states in blackhole systems and so on [54, 160]. One of the primary advantages recurrence based analysis shows over conventional methods of time series analysis is that recurrence networks give reliable results with smaller datasets [39].

We first try to differentiate between the two classes of RRc Lyrae stars classified in chapter 4 using bicoherence measures, using recurrence based analysis. We also look for astrophysical parameters that are similar within these subgroups. This correlation between the astrophysical and nonlinear properties of astrophysical objects has been studied recently in the context of compact objects [161]. We also use recurrence networks to differentiate between the different kinds of close binary stars. We discussed close binaries in Chapter 2, as a group of binary stars where the companion stars can exchange matter or energy. We further classify binary stars based on whether the companion stars fill their Roche lobes as detached, semi-detached, over contact and ellipsoidal binary stars. We attempt to classify close binary stars based on their recurrence network properties and check whether this classification correlates with the classification based on Roche lobe filling.

6.1 Recurrence Analysis of RRc Lyrae stars

In chapter 4, we isolated two distinct populations of RRc Lyrae stars based on bicoherence based analysis. We saw that the stars are either quasiperiodic or strange non chaotic, from their spectral scaling behavior. To further understand the dynamics of these stars, we conduct a recurrence based analysis [44]. For this we first reconstruct the dynamics of the system from its light curves as described in previous chapters. The amplitude

Table 6.1: Delay time (τ), saturated correlation dimension (D_2), mean square difference ($sq.diff.$), Recurrence Rate (RR), Determinism (DET), Laminarity (LAM) and Laminarity Determinism ratio($\frac{LAM}{DET}$) for the four RRc Lyrae stars

KID	τ(days)	D_2	$sq.diff.$	RR	DET	LAM	$\frac{LAM}{DET}$
4064484	0.08	3.05	0.43	0.02	0.71	0.07	0.1
5520878	0.06	3.58	0.25	0.02	0.61	0.02	0.04
8832417	0.06	2.87	0.52	0.01	0.60	0.02	0.04
9453114	0.08	2.89	0.54	0.02	0.61	0.08	0.13

distribution of the light curves are first converted to a uniform distribution[1]. We use only 3000 points for recurrence analysis. These numbers have been shown to give robust estimates of recurrence parameters [39]. To determine the embedding dimension, we first calculate the value of the correlation dimension of these light curves. The correlation dimension and deviation of the correlation dimension from the surrogates, measured by the square difference [2] is listed in Table 6.1. We find the smallest embedding dimension to use to retain the dynamics of the system is 4. Hence we embed the stars in a $4-d$ state space. The recurrence threshold, ϵ used for constructing the recurrence plots is 0.14 in accordance with the criteria laid out in [39].

6.1.1 Recurrence Plots

We initially calculate the recurrence plot from the reconstructed space. A point on the recurrence plot is 1 if two points fall within ϵ distance of each other. The recurrence plots for the four RRc Lyrae variables considered is shown in Figure 6.1. An examination of the plots themselves seem to suggest that the plots corresponding to KIC 4064484 and KIC 9453114 seems to differ from the recurrence plots corresponding to KIC 5520878 and KIC 8832417. We quantify this difference using three main quantifiers of the recurrence plot, namely the laminarity, determinism and determinism to laminarity ratio. These quantities are listed in Table 6.1. The values in the last column are distinctly different for the two sets.

6.1.2 Recurrence Networks

One can use the recurrence plot as an adjacency matrix to construct a recurrence network [162]. Each state space vector is a node and a link exists between two nodes if they are geometrically separated by less than ϵ distance. We set all the diagonal elements to zero to avoid self loops. The recurrence networks constructed from the dynamics of the four RRc Lyrae stars is shown in Figure 6.2. We consider three main properties of the recurrence network so constructed. These are the average degree (d_{avg}), the clustering coefficient (CC_{avg})[3] and the characteristic path length (CPL). These properties are listed in Table 6.2. We see that KIC 4064484 and KIC 9453114 have a lower path length and higher average degree when compared to KIC 5520878 and KIC 8832417 [163]. The CPL-d_{avg} plane is shown in Figure 6.3.

[1]Converting the light curve to a uniform deviate eliminates differences arising from the amplitude distributions of the individual stars

[2]The square difference is calculated as follows $sqdiff = \frac{1}{M_{max}-1}\sum_{M=2}^{M_{max}} D_2(M) - < D_2^{surr}(M) >^2$

[3]The average clustering coefficient is used here.

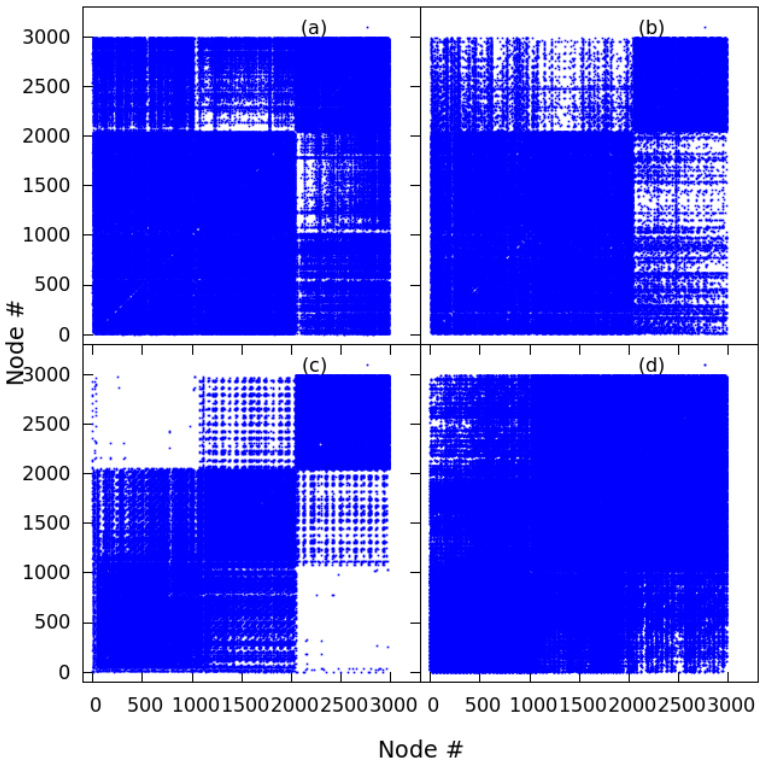


Figure 6.1: Recurrence plots for RRc Lyrae stars (a)KIC 4064484, (b)KIC 5520878, (c)KIC 8832417 and (d)KIC 9453114

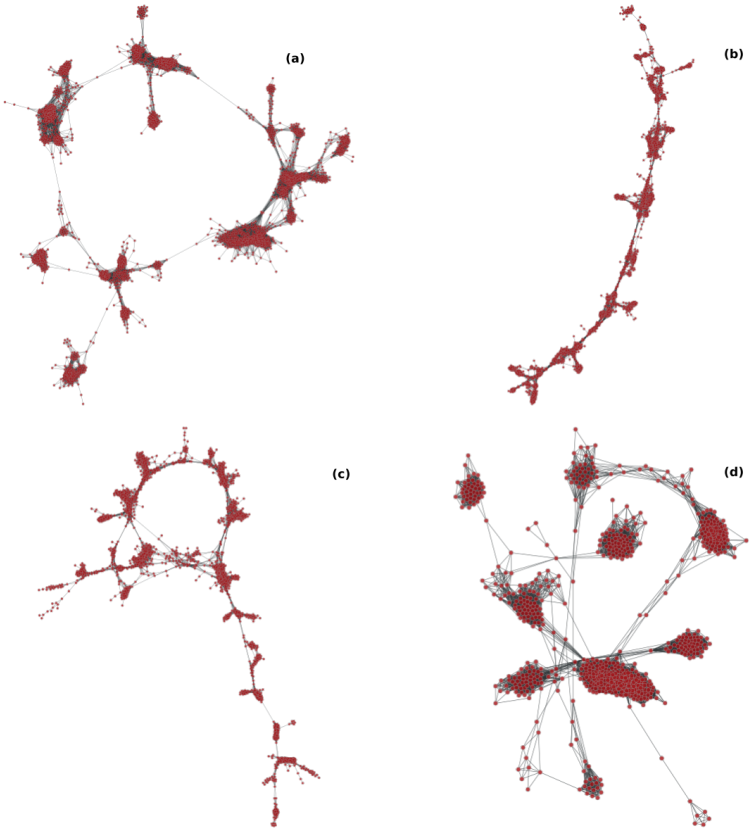


Figure 6.2: Recurrence networks constructed for the four RRc Lyrae stars considered.(a)KIC 4064484 (b)KIC 5520878 (c)KIC 8832417 and (d)KIC 9453114

We hence see that the classes identified above have distinct differences in recurrence based parameters as well. In chapter 4 we found that some of the astrophysical properties of these stars are also distinctly different in three main parameters, the primary period P_1, effective temperature T_{eff} and the metallicity. Our classification technique in chapter 4 was based on bicoherence based measures which need datapoints upwards of 20000(depending on the sampling rate) for a reasonable calculation. On the other hand, recurrence based measures are found to distinguish between the two classes using just 3000 points. We will explore the power of classification using recurrence networks further in the next section on close binary stars.

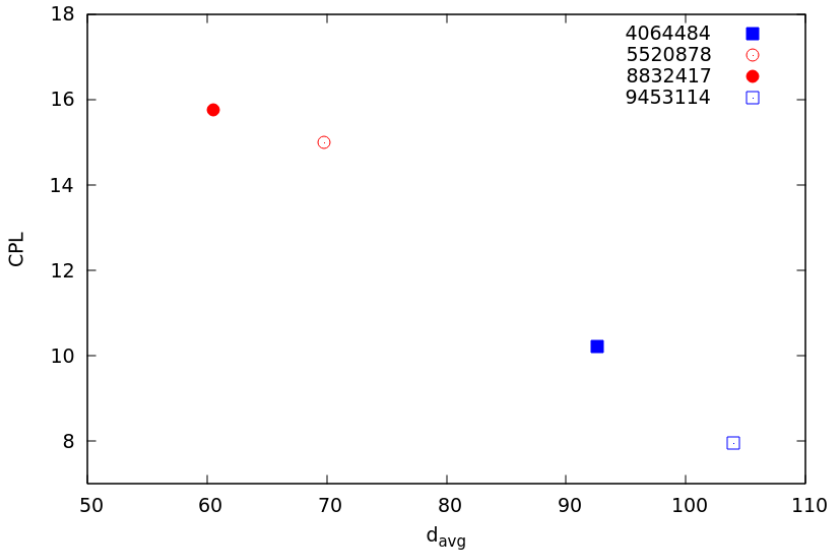


Figure 6.3: d_{avg}-CPL plot for the four RRc Lyrae stars considered. We see that there are two clusters formed with KIC 4064484 and 9453114 occupying a region away from KIC 5520878 and 8832417.

Table 6.2: Average degree (d_{avg}), Characteristic path length (CPL) and average clustering (CC_{avg}) for the RRc Lyrae stars considered.

Kepler ID	d_{avg}	CPL	CC_{avg}
4064484	92.6	10.22	0.73
5520878	69.8	15.0	0.75
8832417	60.5	15.77	0.75
9453114	104	7.95	0.80
white	7.1	10.6	0.49
red	104	7.95	0.80

6.2 Classification of Close Binary Stars

In the last section we explore the possibility of chaotic behavior in contact binary stars. From our discussions in chapter 2, we realize that there are three possible states for binary stars.

- **Detached binary stars:** When both stars do not fill their Roche lobes.
- **Semi detached stars:** When one of the components fills its Roche lobe.
- **Overcontact stars:** When both components fill their Roche lobes.

The last two configurations are called close binary stars. If the companion stars are such that their mutual gravitation distorts their shapes and such that no eclipse is visible from our line of sight, we have a category of stars called ellipsoidal binary stars. From our discussions in chapter 2, we know that the underlying processes governing light variations in different classes of close binaries are different. In this section we question whether this implies that the underlying nonlinear quantifiers are also different. If so, these properties could be used as an alternate classification scheme to identify the different kinds of close binary stars.

We first construct recurrence networks from the light curves of the close binary stars. We then study the properties of the recurrence networks, mainly the characteristic path length and clustering coefficient. We then explore the clusters in the plane of CPL and CC (and CPL alone) using clustering algorithms and machine learning techniques.

We find that the clusters found using these algorithms largely correspond to the presently accepted astrophysical classification. We can especially distinguish semi detached and ellipsoidal stars from over contact binary stars. Hence, we suggest that the nonlinear properties of close binaries may be used as an alternative method to separate them into semi detached, over contact and ellipsoidal binary stars. For our analysis we use the embedding dimension 4 for all the stars. The analysis aims to check whether the parameters of the recurrence network constructed from the embedded $4-d$ state space are distinct for different classes of close binary stars. The ϵ is chosen to be 0.14 using the criteria suggested in [39], for $4-d$ systems.

We quantify the recurrence networks using the characteristic path length (CPL) and average clustering(CC) of the network. We first examine the characteristic path length of the different kinds of close binaries. We show the kernel density plots for the three types of close binary stars in Figure 6.4. We see that the overcontact stars can be distinctly distinguished from the other two categories using CPL alone.

As mentioned earlier, ellipsoidal binaries are inclined such that no eclipses are possible with the line of sight. Hence these stars have distinctly different inclinations from semi detached stars. This is shown in Figure 6.5. We speculate that in many cases low inclination semi detached stars may be wrongly classified as being ellipsoidal.

6.2.1 Clusters in the CPL-CC Plane

We further classify the stars based on the clustering coefficients of their recurrence networks. The location on the plane of the CPL and CC has been shown to be a robust way to distinguish between dynamical states [54, 164]. We calculate and plot the values of CPL and CC on a parameter plane in Figure 6.6.

We will now examine how well the different kinds of close binaries can be identified from the recurrence network parameters. Initially we consider classification of the binaries using just the CPL. We define the accuracy of the algorithm as corrected predicted fraction of the total set. The intervals are chosen using the distributions of the CPL for the three kinds of close binaries, identified using the kernel density estimate. One finds that the distributions merge for over-contact and semi-detached at $CPL = 9.8$

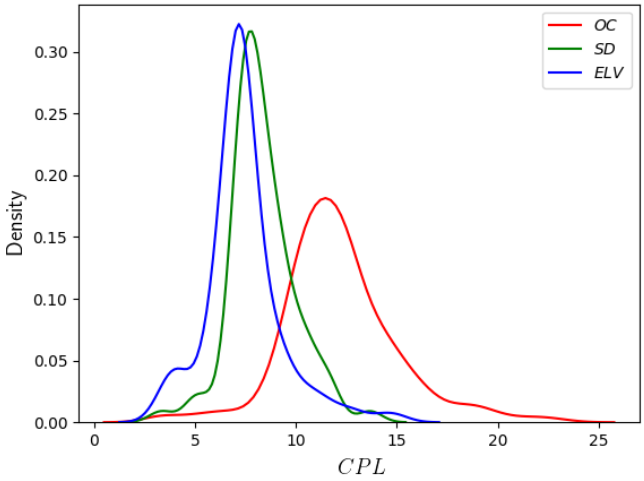


Figure 6.4: Distributions of CPL for the three kinds of close binary stars. The distributions for semi detached and ellipsoidal binary stars are distinctly different from the distribution for the overcontact stars.

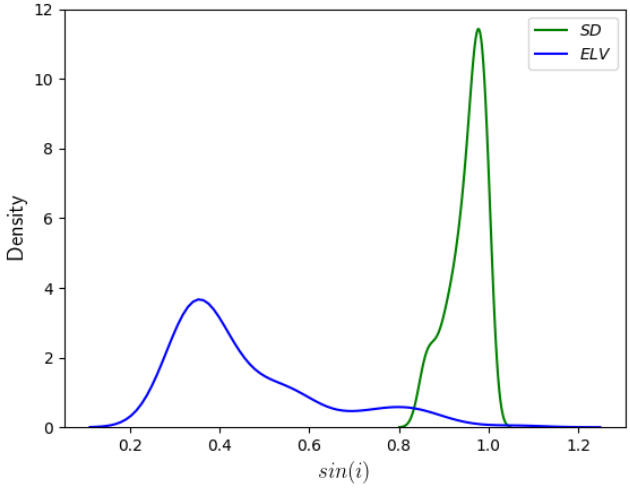


Figure 6.5: Distributions of $sin(i)$ for ellipsoidal and semi detached binary stars. Ellipsoidal stars have distinctly low inclinations whereas semi detached stars are almost exclusively detected at high inclinations.

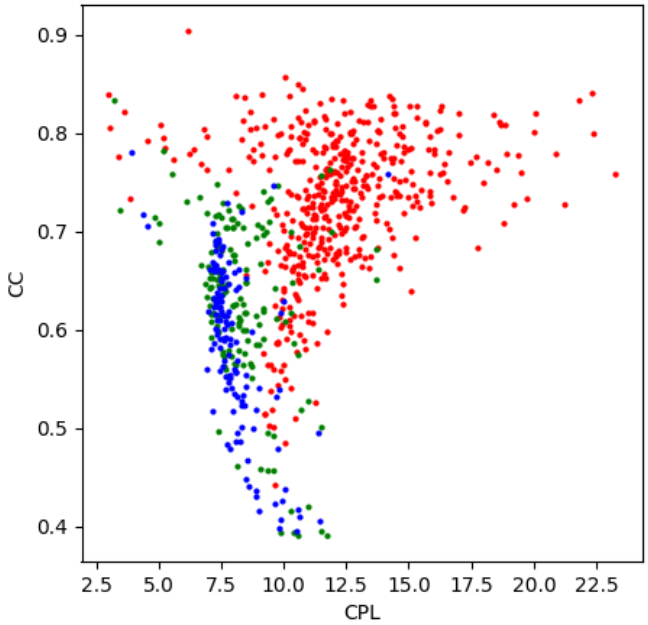


Figure 6.6: Locations of close binary stars plotted on the CPL-CC plane. Green corresponds to semi-detached, red corresponds to overcontact and blue to ellipsoidal binary stars. We immediately notice that overcontact stars lie in a distinctly different part of the plane from the semidetached and ellipsoidal stars.

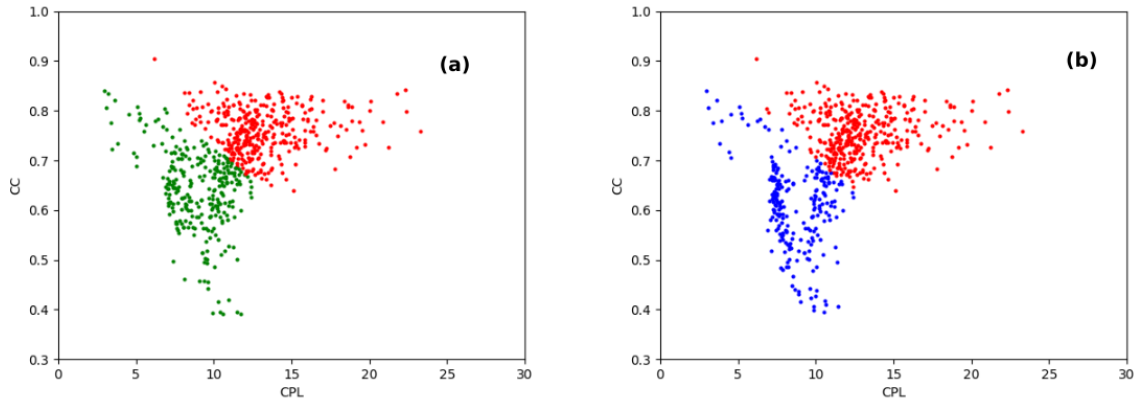


Figure 6.7: Clusters predicted by the $k-means$ clustering algorithm in the CPL-CC plane. Predicted clusters for (a)Overcontact-Semi detached case and (b)Overcontact-Ellipsoidal case.The color code is green: semi-detached, red: overcontact and blue: ellipsoidal binary stars.

and for overcontact and ellipsoidal at $CPL = 9.2$(Figure 6.4). For semi detached versus overcontact, this gives an accuracy of 0.86. We correctly classify 402 of 463 overcontact stars and 129 of 152 semi detached stars using these intervals. For the ellipsoidal case, we get an accuracy of 0.91 with 428 of 463 overcontact stars and 120 of 137 ellipsoidal stars falling into these intervals.

Classification using k-Means clustering

We next use a simple $k-means$ clustering algorithm to identify clusters in the $CPL-CC$ parameter space. The $k-means$ algorithm works in the following way [165].

- A random set of observations are chosen as initial $k-means$, $m_1, m_2, ..., m_k$[4].
- Each observation is then assigned to the nearest[5] mean.
- In the update step, the new means are calculated as the centroids of the clusters.

In our case we partition our CPL-CC plane into two. As before first we try to distinguish between overcontact and semi detached binary stars. Using the k-means algorithm, we get cluster the plane into two as shown in Figure 6.7a. This correctly classifies 330 of 463 overcontact binaries and 146 of 152 semi detached binaries. This leads to an accuracy of 0.77. Similar analysis for ellipsoidal binaries correctly classifies 357 overcontact binaries and 135 (of 137) ellipsoidal binaries with an accuracy of 0.82.

Classification using support vector machine

Support vector machine is a machine learning technique that attempts to partition a distribution into clusters based on a discriminating hyperplane. Suppose we have a

[4]This is called the Forgy method. Alternatively, clusters can be randomly assigned to each observation and the update can be conducted. This sets the initial mean to be the centroid of the cluster points. This method is called the Random Partition method.

[5]Nearest is calculated using least squared Euclidean distance

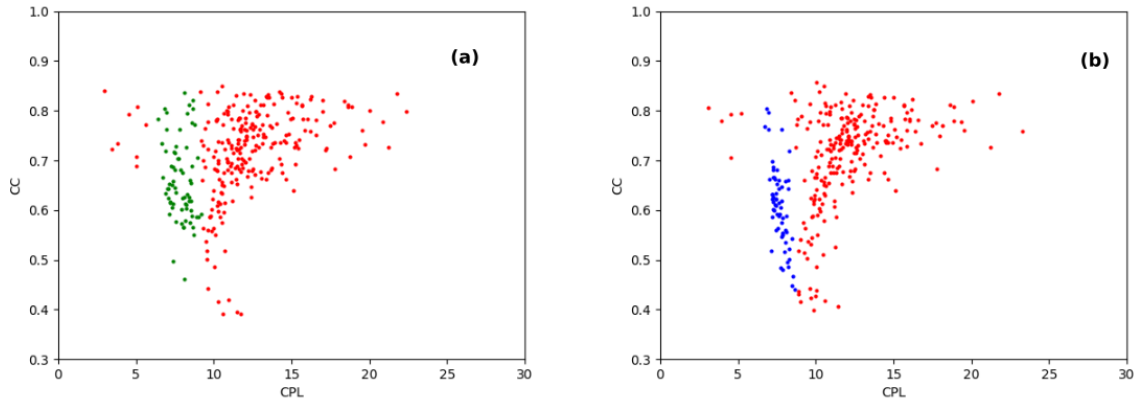


Figure 6.8: Clusters predicted by the support vector clustering algorithm in the CPL-CC plane. Predicted clusters for (a)Overcontact-Semi detached case and (b)Overcontact-Ellipsoidal case. The color code is green: semi-detached, red: overcontact and blue: ellipsoidal binary stars.

distribution of points in an $n-d$ space, separated into two clusters. The support vector machine tries to construct the optimal $n-1$ dimensional surface in this plane, such that the distance of the points closest to the surface from it, is maximized.

We use the support vector machine algorithm to classify the different classes of binaries in the $CPL-CC$ plane. 50% of the points are used for training the algorithm. The remaining 50% of data is used as the testing set. We see that the testing set gives an accuracy of 0.89 ± 0.01 for the semi detached versus over contact case, where as the ellipsoidal versus overcontact case gives an accuracy of 0.91 ± 0.01. The predicted clusters are shown in Figure 6.8.

Algorithm	*Accuracy*$(OC-ELV)$	*Accuracy*$(OC-SD)$
CPL	0.91	0.86
$K-means$	0.82	0.77
SVC	0.94 ± 0.01	0.89 ± 0.01

6.3 Summary and Discussion

In this chapter we present the results of recurrence based analysis of RRc Lyrae stars. In chapter 4 we saw that there are two distinct groups exhibiting strange non chaotic and quasiperiodic dynamics. These groups also show distinct astrophysical properties like effective temperature, metallicity and primary period. The recurrence based analysis conducted in this chapter shows that the two classes have highly differing recurrence plot and network measures. Hence this again supports our claim that there is a strong link between the astrophysical and nonlinear properties of RRc Lyrae stars. We explore this idea further in close binaries, where we find good correlation between the nonlinear dynamics and astrophysics of this category of binary stars.

We use measures of recurrence networks to distinguish between the different types of close binary stars. Different kinds of close binary stars have different mechanisms that are responsible for their variability. We check if these differences reflect in the

nonlinear properties of their light curves. If so, these properties can be used as an alternate classification scheme to identify the different kinds of close binary stars.

We first construct recurrence networks from the light curves of the close binary stars. We then study the properties of the recurrence networks, mainly the characteristic path length and clustering coefficient. We then explore the clusters in the plane of CPL and CC (and CPL alone) using clustering algorithms and machine learning techniques. We find that the clusters found using these algorithms largely correspond to the presently accepted astrophysical classification. We can especially distinguish semi detached and ellipsoidal stars from over contact binary stars. Hence, we suggest that the nonlinear properties of close binaries may be used as an alternative method to classify them into semi detached, over contact and ellipsoidal binary stars.

Chapter 7

Conclusions and Discussions

We summarize the results of the study reported in the thesis, its significance and the horizons to which it can be applied and extended to. We start by introducing the salient features of the methods used in nonlinear time series analysis and in subsequent chapters present their application to various types of datasets, specifically astrophysical data. We address how to manage datagaps and noise contamination in dealing with such data and their practical effectiveness in arriving at conclusions regarding the dynamics underlying the data.

7.1 Summary

Most of the methods of time series analysis are developed for evenly sampled datasets. Missing data, which is ubiquitous in observational data, is generally smoothed over by interpolation. Interpolation of data is known to create artifacts that resemble chaotic behavior, making interpolation unsuitable for nonlinear time series analysis. We tackle the problem of missing data by quantifying the extend to which they are affected by gaps in observations. We do this by artificially introducing gaps in standard systems and calculating deviations of quantifiers from the evenly sampled values. This helps us identify a region of gap sizes and frequencies, where the quantifier under consideration is resilient to gaps. We use this to identify regions where reliable conclusions can be drawn for correlation dimension D_2 and the multifractal spectrum $f(\alpha)$. We also use the results to compute these quantifiers for various datasets. We find signatures of chaos in the unevenly sampled light curves of pulsating stars as evidenced by saturating D_2. We also find multifractality resulting from deterministic nonlinearity in the time series of photosynthesis data and some meteorological datasets.

A second issue in all kinds of observational data is the presence of noise. Even perfectly deterministic systems would be affected by noise at the time of measurement. Apart from measurement noise, stochastic effects may be present at different points in the system's dynamics. The system itself may be evolving in the presence of noise, the measured variable may be contaminated by noise sources prior to measurement, the parameters of the system may be affected by noise and so on. We specifically consider the problem of distinguishing between different dynamical states when noise is present in the system. To do so, we use the bicoherence function. First we show that for a system evolving with noise, the bicoherence can be use to distinguish between noisy limit cycles and chaotic states. We use this to show that RRab Lyrae stars may be exhibiting richer dynamics than simple period doubled dynamics. We also point out that strange non chaotic states

behaves similar to noisy quasiperiodic states, when one checks the scaling behavior of peaks in the strobed power spectrum for them. We show that a bicoherence based filter can remove this ambiguity.

Finally we use the techniques developed, to study the dynamics of close binary stars. We initially establish chaotic behavior in contact binary stars. We show this by showing that most contact binaries in the Kepler field of view show saturating D_2, away from surrogate datatsets. We also explore the bicoherence properties of these stars. Our analysis also indicates that the D_2 and bicoherence properties of these stars are correlated to the extend of contact that exists between the component stars in a binary. We show that this correlation is higher when we restrict the spectral class and mass ratios from which the stars are sampled. This seems to be the first work that indicates how closely the astrophysical and nonlinear dynamical properties of a star may be related to each other.

We also show that we can use the nonlinear properties of stars to classify them. To achieve this we use the framework of recurrence plots and networks. We initially analyse the light curves of RRc Lyrae stars that are thought to show strange non chaotic behavior. In RRc Lyrae stars we discover two groups showing strange non chaotic and quasiperiodic behavior using bicoherence studies. Using recurrence plots and recurrence networks we establish that other nonlinear properties are also different for these groups. We use the properties of recurrence networks to distinguish between different types of close binary stars. We see that the properties of the recurrence network, especially the average path length (CPL) and clustering coefficient (CC), for contact binaries are different from other classes of binaries. We suggest using this technique to classify large datasets of close binaries into different standard types like semi detached, overcontact and ellipsoidal binaries.

7.2 Significance and Future Directions

The study reported in the thesis gives a recipe for examining the effect that datagaps have on nonlinear quantifiers. While we have used it only to examine three main quantifiers, the method is applicable to any nonlinear quantifier. It addresses the question of whether a quantifier can be used to analyze a time series with a particular profile of gaps. An important extension of this work would be to compare the effects of various kinds of interpolation with our method, and compare which works best. In this context it is imperative to mention the use of uniform deviates which converts the amplitude distribution to a uniform distribution. Our preliminary analysis with datagaps suggests that in the absence of uniform deviates, gaps tend to affect nonlinear quantifiers to a larger extend.

The bicoherence function computed from Fourier transforms has been utilized extensively in time series analysis. However barring a few early attempts in the 1990s, they have not been used to analyze chaotic systems in general. Bicoherence and the tricoherence functions seem to give important insights about systems and their dynamical states, especially in the case of period doubling chaos. One important extension would be to use cross bicoherence techniques to analyse nonlinear phenomena that involve relationships between multiple variables. Preliminary work in this direction seems to show good promise.

With the launch of the Transiting Exoplanet Survey Satellite (TESS), whose first light

was received on August 2018 larger datasets will soon become available for analysis. Then the relative abundance of quasiperiodic stars to strange non chaotic stars can be explored. An important step to take in the direction of understanding the dynamics of pulsating stars would be to have data driven models. An integration of time series techniques and numerical hydrodynamic simulations has not been considered in understanding pulsating stars, unlike in the case of compact objects [161]. Utilization of complimentary expertise to tackle these questions would result in an important leap forward in understanding the physics of these objects.

In chapter 5 we conducted an extensive analysis of contact binary stars. Apart from gaining important insights into the underlying physics of these stars, it also gives a practical method to classify large datasets and predict values for quantifiers like the fill out factor. While we have combed through the catalog of contact binary stars to find correlations between fill-out factor and nonlinear quantifiers, many more relations may emerge through a careful analysis of this dataset. Besides this dataset, the work opens up the possibility that physical parameters may be linked to nonlinear parameters. This may be the case in other natural systems exhibiting chaos. This method also gives important checks on the reliability of the numerical hydrodynamics models. While certain features of the light curves are reproduced currently, an important additional check would be to see if the nonlinear properties of the light curves produced from simulation and data are similar.

The recurrence based analysis in chapter 6 provides an alternate and more quantitative method of classification, based on the nonlinear dynamical properties instead of using broad features of the light curve. Further, the understanding for the reasons for the light variations in ellipsoidal binary stars is still lacking. The study of the nonlinear properties of ellipsoidal stars will form an important clue to develop an understanding of the reasons for light variations in these stars. Finally, merger events like the one reported in [62] are poorly understood. Recurrence networks have been shown to be good indicators for dynamical transitions in real world systems basically because they give reliable results with small datasets [160, 166]. So also continuous time monitoring of recurrence network parameters may be a good approach to study upcoming merger events in binary stars.

Recently machine learning algorithms are being increasingly used in the analysis of large number of datasets and classification of their results. We illustrate this in the context of the classification of binary stars based on the recurrence network measures. Very soon machine learning and artificial intelligence will become quite powerful tools in scientific research.

Time series analysis has also become an important tool for research in multiple diverse fields like medicine, finance, sociology, meteorology and ecology. There is an increasing recognition of the need to use the techniques of dynamical systems theory to fully understand various astrophysical phenomena. This thesis attempts to draw these fields together by using the techniques of time series analysis to astrophysical systems that show nonlinear and chaotic behavior. We strongly believe that this study makes an important addition to the research in nonlinear dynamics, astrophysics and time series analysis and will prove to be a starting point for future developments in these directions.

Bibliography

[1] Lorenz, E. N. Deterministic nonperiodic flow. *Journal of the atmospheric sciences* **20**, 130–141 (1963).

[2] Sparrow, C. *The Lorenz equations: bifurcations, chaos, and strange attractors*, vol. 41 (Springer Science & Business Media, 2012).

[3] Rössler, O. E. An equation for continuous chaos. *Physics Letters A* **57**, 397–398 (1976).

[4] Rossler, O. E. & WEGMANN, K. Chaos in the zhabotinskii reaction. *Nature* **271**, 89 (1978).

[5] Hilborn, R. C. *Chaos and nonlinear dynamics: an introduction for scientists and engineers* (Oxford University Press on Demand, 2000).

[6] Hirsch, M. W., Smale, S. & Devaney, R. L. *Differential equations, dynamical systems, and an introduction to chaos* (Academic press, 2012).

[7] Lakshmanan, M. & Rajaseekar, S. *Nonlinear dynamics: integrability, chaos and patterns* (Springer Science & Business Media, 2012).

[8] Romeiras, F. J. & Ott, E. Strange nonchaotic attractors of the damped pendulum with quasiperiodic forcing. *Physical Review A* **35**, 4404 (1987).

[9] Regev, O. *Chaos and complexity in astrophysics* (Cambridge University Press, 2006).

[10] Michael, S. *Applied nonlinear time series analysis: applications in physics, physiology and finance*, vol. 52 (World Scientific, 2005).

[11] Strogatz, S. H. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering* (CRC Press, 2018).

[12] Kantz, H. & Schreiber, T. *Nonlinear time series analysis*, vol. 7 (Cambridge university press, 2004).

[13] Press, W. H. *Numerical recipes 3rd edition: The art of scientific computing* (Cambridge university press, 2007).

[14] Rao, T. S. & Gabr, M. *An introduction to bispectral analysis and bilinear time series models*, vol. 24 (Springer Science & Business Media, 2012).

[15] Lohmann, A. W. & Wirnitzer, B. Triple correlations. *Proceedings of the IEEE* **72**, 889–901 (1984).

[16] Totsky, A. V., Zelensky, A. A. & Kravchenko, V. F. *Bispectral methods of signal processing: applications in radar, telecommunications and digital image restoration* (Walter de Gruyter GmbH & Co KG, 2015).

[17] Arfken, G. B. & Weber, H. J. Mathematical methods for physicists (1999).

[18] Bartlett, M. Smoothing periodograms from time-series with continuous spectra. *Nature* **161**, 686 (1948).

[19] Liu, Y., Gopikrishnan, P., Stanley, H. E. *et al.* Statistical properties of the volatility of price fluctuations. *Physical review e* **60**, 1390 (1999).

[20] Siwy, Z. & Fuliński, A. Origin of 1/f α noise in membrane channel currents. *Physical review letters* **89**, 158101 (2002).

[21] Pettersen, K. H., Lindén, H., Tetzlaff, T. & Einevoll, G. T. Power laws from linear neuronal cable theory: power spectral densities of the soma potential, soma membrane current and single-neuron contribution to the eeg. *PLoS computational biology* **10**, e1003928 (2014).

[22] Geisel, T., Zacherl, A. & Radons, G. Generic 1 f noise in chaotic hamiltonian dynamics. *Physical review letters* **59**, 2503 (1987).

[23] Dutta, P. & Horn, P. Low-frequency fluctuations in solids: 1 f noise. *Reviews of Modern physics* **53**, 497 (1981).

[24] Yadav, A. C., Ramaswamy, R. & Dhar, D. General mechanism for the 1/f noise. *Physical Review E* **96**, 022215 (2017).

[25] Venkataraman, B. *Elements of Nonequilibrium Statistical Mechanics* (Ane Books, 2008).

[26] Platen, E. A benchmark approach to finance. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* **16**, 131–151 (2006).

[27] Jona-Lasinio, G. & Mitter, P. K. On the stochastic quantization of field theory. *Communications in Mathematical Physics* **101**, 409–436 (1985).

[28] Damos, P., Rigas, A. & Savopoulou-Soultani, M. Application of markov chains and brownian motion models on insect ecology (2011).

[29] Zhang, M. A markov stochastic process theory of cosmic-ray modulation. *The Astrophysical Journal* **513**, 409 (1999).

[30] Elgar, S. & Guza, R. T. Statistics of bicoherence. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **36**, 1667–1668 (1988).

[31] Chandran, V., Elgar, S. & Pezeshki, C. Bispectral and trispectral characterization of transition to chaos in the duffing oscillator. *International Journal of Bifurcation and Chaos* **3**, 551–557 (1993).

[32] Takens, F. Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980*, 366–381 (Springer, 1981).

[33] Aguirre, L. A. & Letellier, C. Modeling nonlinear dynamics and chaos: a review. *Mathematical Problems in Engineering* **2009** (2009).

[34] Wolf, A., Swift, J. B., Swinney, H. L. & Vastano, J. A. Determining lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena* **16**, 285–317 (1985).

[35] Argyris, J., Andreadis, I., Pavlos, G. & Athanasiou, M. On the influence of noise on the largest lyapunov exponent and on the geometric structure of attractors. *Chaos, Solitons & Fractals* **9**, 947–958 (1998).

[36] Serletis, A., Shahmoradi, A. & Serletis, D. Effect of noise on estimation of lyapunov exponents from a time series. *Chaos, Solitons & Fractals* **32**, 883–887 (2007).

[37] Harikrishnan, K., Misra, R., Ambika, G. & Kembhavi, A. A non-subjective approach to the gp algorithm for analysing noisy time series. *Physica D: Nonlinear Phenomena* **215**, 137–145 (2006).

[38] Harikrishnan, K., Misra, R. & Ambika, G. Efficient use of correlation entropy for analysing time series data. *Pramana* **72**, 325–333 (2009).

[39] Jacob, R., Harikrishnan, K., Misra, R. & Ambika, G. Uniform framework for the recurrence-network analysis of chaotic time series. *Physical Review E* **93**, 012202 (2016).

[40] Jensen, M. H., Kadanoff, L. P., Libchaber, A., Procaccia, I. & Stavans, J. Global universality at the onset of chaos: results of a forced rayleigh-bénard experiment. *Physical review letters* **55**, 2798 (1985).

[41] Harikrishnan, K. P., Misra, R., Ambika, G. & Amritkar, R. Computing the multifractal spectrum from time series: an algorithmic approach. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **19**, 043129 (2009).

[42] Harikrishnan, K. P., Misra, R. & Ambika, G. Nonlinear time series analysis of the light curves from the black hole system grs1915+ 105. *Research in Astronomy and Astrophysics* **11**, 71 (2011).

[43] Shekatkar, S. M., Kotriwar, Y., Harikrishnan, K. P. & Ambika, G. Detecting abnormality in heart dynamics from multifractal analysis of ecg signals. *Scientific reports* **7**, 15127 (2017).

[44] Marwan, N., Romano, M. C., Thiel, M. & Kurths, J. Recurrence plots for the analysis of complex systems. *Physics reports* **438**, 237–329 (2007).

[45] Shekatkar, S. M., Bhagwat, C. & Ambika, G. Divisibility patterns of natural numbers on a complex network. *Scientific reports* **5**, 14280 (2015).

[46] Muñoz-Laboy, M., Garcia, J., Wilson, P. A., Parker, R. G. & Severson, N. Heteronormativity and sexual partnering among bisexual latino men. *Archives of sexual behavior* **44**, 895–902 (2015).

[47] King, A. D., Pržulj, N. & Jurisica, I. Protein complex prediction via cost-based clustering. *Bioinformatics* **20**, 3013–3020 (2004).

[48] Lacasa, L., Luque, B., Ballesteros, F., Luque, J. & Nuno, J. C. From time series to complex networks: The visibility graph. *Proceedings of the National Academy of Sciences* **105**, 4972–4975 (2008).

[49] Xu, X., Zhang, J. & Small, M. Superfamily phenomena and motifs of networks induced from time series. *Proceedings of the National Academy of Sciences* pnas–0806082105 (2008).

[50] Li, P. & Wang, B. An approach to hang seng index in hong kong stock market based on network topological statistics. *Chinese Science Bulletin* **51**, 624–629 (2006).

[51] Donner, R. V., Zou, Y., Donges, J. F., Marwan, N. & Kurths, J. Recurrence networksâĂŤa novel paradigm for nonlinear time series analysis. *New Journal of Physics* **12**, 033025 (2010).

[52] Zou, Y., Donner, R. V., Marwan, N., Donges, J. F. & Kurths, J. Complex network approaches to nonlinear time series analysis. *Physics Reports* (2018).

[53] Newman, M. *Networks* (Oxford university press, 2018).

[54] Jacob, R., Harikrishnan, K., Misra, R. & Ambika, G. Characterization of chaotic attractors under noise: A recurrence network perspective. *Communications in Nonlinear Science and Numerical Simulation* **41**, 32–47 (2016).

[55] Roy, A. E. & Clarke, D. *Astronomy: Principles and Practice, (PBK)* (CRC Press, 2003).

[56] Ostlie, D. A. & Carroll, B. W. *An introduction to modern stellar astrophysics* (Addison-Wesley, 2007).

[57] Phillips, A. C. *The physics of stars* (John Wiley & Sons, 2013).

[58] Perryman, M. A. *et al.* The hipparcos catalogue. *Astronomy and Astrophysics* **323**, L49–L52 (1997).

[59] Shu, F. *The physical universe: an introduction to astronomy* (University science books, 1982).

[60] Perryman, M. *et al.* The hyades: distance, structure, dynamics, and age. *Astronomy and Astrophysics* **331**, 81 (1998).

[61] Percy, J. R. *Understanding variable stars* (Cambridge University Press, 2007).

[62] Tylenda, R. *et al.* V1309 scorpii: merger of a contact binary. *Astronomy & Astrophysics* **528**, A114 (2011).

[63] Skuljan, L. & Cottrell, P. L. R coronae borealis stars. In *International Astronomical Union Colloquium*, vol. 193, 511–519 (Cambridge University Press, 2004).

[64] Kallrath, J. & Milone, E. F. *Eclipsing binary stars: modeling and analysis* (Springer, 2009).

[65] Kurtz, D. Stellar pulsation: an overview. In *Astrophysics of Variable Stars*, vol. 349, 101 (2006).

[66] Buchler, J. R. & Kovacs, G. Period doubling bifurcations and chaos in w virginis models. *The Astrophysical Journal* **320**, L57–L62 (1987).

[67] Serre, T., Kolláth, Z. & Buchler, J. Search for low-dimensional nonlinear behavior in irregular variable stars. the analysis of irregular w virginis model pulsations. *Astronomy and astrophysics* **311**, 845–851 (1996).

[68] Stellingwerf, R. Convection in pulsating stars-part two-rr-lyrae convection and stability. *The Astrophysical Journal* **262**, 339 (1982).

[69] Stellingwerf, R. Convection in pulsating stars. iii-the rr lyrae instability strip. *The Astrophysical Journal* **277**, 322–326 (1984).

[70] Plachy, E., Kolláth, Z. & Molnár, L. Low-dimensional chaos in rr lyrae models. *Monthly Notices of the Royal Astronomical Society* **433**, 3590–3596 (2013).

[71] Buchler, J. R. & Kollath, Z. On the blazhko effect in rr lyrae stars. *The astrophysical journal* **731**, 24 (2011).

[72] Stothers, R. B. A new explanation of the blazhko effect in rr lyrae stars. *The Astrophysical Journal* **652**, 643 (2006).

[73] Moskalik, P. *et al.* Kepler photometry of rrc stars: peculiar double-mode pulsations and period doubling. *Monthly Notices of the Royal Astronomical Society* **447**, 2348–2366 (2015).

[74] Szabó, R. *et al.* Does kepler unveil the mystery of the blazhko effect? first detection of period doubling in kepler blazhko rr lyrae stars. *Monthly Notices of the Royal Astronomical Society* **409**, 1244–1252 (2010).

[75] Nemec, J. M. *et al.* Metal abundances, radial velocities, and other physical characteristics for the rr lyrae stars in the kepler field. *The Astrophysical Journal* **773**, 181 (2013).

[76] Lindner, J. F. *et al.* Strange nonchaotic stars. *Physical review letters* **114**, 054101 (2015).

[77] Guinan, E. F. & Bradstreet, D. H. Kinematic clues to the origin and evolution of low mass contact binaries. In *Formation and Evolution of Low Mass Stars*, 345–375 (Springer, 1988).

[78] Podsiadlowski, P. The evolution of close binaries. In *RS Ophiuchi (2006) and the Recurrent Nova Phenomenon*, vol. 401, 63 (2008).

[79] O'Connell, D. The so-called periastron effect in eclipsing binaries. *Monthly Notices of the Royal Astronomical Society* **111**, 642–642 (1951).

[80] Conroy, K. E. *et al.* Kepler eclipsing binary stars. iv. precise eclipse times for close binaries and identification of candidate three-body systems. *The Astronomical Journal* **147**, 45 (2014).

[81] Molnar, L. A. *et al.* Prediction of a red nova outburst in kic 9832227. *The Astrophysical Journal* **840**, 1 (2017).

[82] Socia, Q. J. *et al.* Kic 9832227: Using vulcan data to negate the 2022 red nova merger prediction. *The Astrophysical Journal Letters* **864**, L32 (2018).

[83] Wilson, R. E. Eccentric orbit generalization and simultaneous solution of binary star light and velocity curves. *The Astrophysical Journal* **234**, 1054–1066 (1979).

[84] Buchler, J. R., Serre, T., Kolláth, Z. & Mattei, J. A chaotic pulsating star: The case of r scuti. *Physical review letters* **74**, 842 (1995).

[85] Misra, R., Harikrishnan, K., Mukhopadhyay, B., Ambika, G. & Kembhavi, A. The chaotic behavior of the black hole system grs 1915+ 105. *The Astrophysical Journal* **609**, 313 (2004).

[86] Misra, R., Harikrishnan, K. P., Ambika, G. & Kembhavi, A. The nonlinear behavior of the black hole system grs 1915+ 105. *The Astrophysical Journal* **643**, 1114 (2006).

[87] Karak, B. B., Dutta, J. & Mukhopadhyay, B. Search for chaos in neutron star systems: Is cyg x-3 a black hole? *The Astrophysical Journal* **708**, 862 (2009).

[88] Bachev, R., Mukhopadhyay, B. & Strigachev, A. A search for chaos in the optical light curve of a blazar: W2r 1926+ 42. *Astronomy & Astrophysics* **576**, A17 (2015).

[89] Eng, F. *Non-uniform sampling in statistical signal processing.* Ph.D. thesis, Institutionen för systemteknik (2007).

[90] Lomb, N. R. Least-squares frequency analysis of unequally spaced data. *Astrophysics and space science* **39**, 447–462 (1976).

[91] Scargle, J. D. Studies in astronomical time series analysis. ii-statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal* **263**, 835–853 (1982).

[92] Nicolis, C. & Nicolis, G. Is there a climatic attractor? *Nature* **311**, 529 (1984).

[93] Grassberger, P. Do climatic attractors exist? *Nature* **323**, 609 (1986).

[94] Nicolis, C. & Nicolis, G. Evidence for climatic attractors. *Nature* **326**, 523–523 (1987).

[95] Schreiber, T. & Schmitz, A. Improved surrogate data for nonlinearity tests. *Physical Review Letters* **77**, 635 (1996).

[96] Hegger, R., Kantz, H. & Schreiber, T. Practical implementation of nonlinear time series methods: The tisean package. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **9**, 413–435 (1999).

[97] George, S. V., Ambika, G. & Misra, R. Effect of data gaps on correlation dimension computed from light curves of variable stars. *Astrophysics and Space Science* **360**, 5 (2015).

[98] Osborne, A. R. & Provenzale, A. Finite correlation dimension for stochastic systems with power-law spectra. *Physica D: Nonlinear Phenomena* **35**, 357–381 (1989).

[99] Kafka, S. Observations from the aavso international database (2016).

[100] Haataja, J. & Vesala, T. Station for measuring forest ecosystem–atmosphere relations: Smear ii. *University of Helsinki, Department of Forest Ecology Publications* **17** (1997).

[101] Hari, P. *et al.* Station for measuring ecosystem-atmosphere relations: Smear. In *Physical and Physiological Forest Ecology*, 471–487 (Springer, 2013).

[102] García, R. *et al.* Impact on asteroseismic analyses of regular gaps in kepler data. *Astronomy & Astrophysics* **568**, A10 (2014).

[103] Grootes, P. & Stuiver, M. Oxygen 18/16 variability in greenland snow and ice with 10- 3-to 105-year time resolution. *Journal of Geophysical Research: Oceans* **102**, 26455–26470 (1997).

[104] Hu, C. *et al.* Quantification of holocene asian monsoon rainfall from spatially separated cave records. *Earth and Planetary Science Letters* **266**, 221–232 (2008).

[105] Rehfeld, K., Marwan, N., Heitzig, J. & Kurths, J. Comparison of correlation analysis techniques for irregularly sampled time series. *Nonlinear Processes in Geophysics* **18**, 389–404 (2011).

[106] Kolláth, Z. Chaotic behaviour in the light variation of the rv-tauri star r-scuti. *Monthly Notices of the Royal Astronomical Society* **247**, 377 (1990).

[107] Buchler, J. R., Kolláth, Z. & Cadmus, R. Chaos in the music of the spheres. In *AIP Conference Proceedings*, vol. 622, 61–73 (AIP, 2002).

[108] Ochsenbein, F., Bauer, P. & Marcout, J. The vizier database of astronomical catalogues. *Astronomy and Astrophysics Supplement Series* **143**, 23–32 (2000).

[109] Fokin, A. Nonlinear pulsations of the rv tauri stars. *Astronomy and Astrophysics* **292**, 133–151 (1994).

[110] Dick, J. & Walker, H. Fadings of r coronae borealis stars-random or chaotic? *Astronomy and Astrophysics* **252**, 701–704 (1991).

[111] Schreiber, M. R. & Lasota, J.-P. The dwarf nova ss cygni: what is wrong? *Astronomy & Astrophysics* **473**, 897–901 (2007).

[112] Cannizzo, J. K. & Goodings, D. Chaos in ss cygni? *The Astrophysical Journal* **334**, L31–L34 (1988).

[113] Hempelmann, A. & Kurths, J. Dynamics of the outburst series of ss cygni. *Astronomy and Astrophysics* **232**, 356–366 (1990).

[114] Cannizzo, J. K. & Mattei, J. A. On the long-term behavior of ss cygni. *The Astrophysical Journal* **401**, 642–653 (1992).

[115] Žliobaitė, I., Hollmén, J. & Junninen, H. Regression models tolerant to massively missing data: a case study in solar-radiation nowcasting. *Atmospheric Measurement Techniques* **7**, 4387–4399 (2014).

[116] Hari, P., Heliövaara, K. & Kulmala, L. *Physical and physiological forest ecology* (Springer Science & Business Media, 2012).

[117] Schneider, G. W. & Childers, N. Influence of soil moisture on photosynthesis, respiration, and transpiration of apple leaves. *Plant physiology* **16**, 565 (1941).

[118] Lüttge, U. & Beck, F. Endogenous rhythms and chaos in crassulacean acid metabolism. *Planta* **188**, 28–38 (1992).

[119] Krempaskỳ, J., Smrčinová, M. & Ballo, P. Periodicity and chaos in a photosynthetic system. *Photosynthesis research* **37**, 159–164 (1993).

[120] Farquhar, G. v., von Caemmerer, S. v. & Berry, J. A biochemical model of photosynthetic co2 assimilation in leaves of c3 species. *Planta* **149**, 78–90 (1980).

[121] Shimshi, D. Effect of soil moisture and phenylmercuric acetate upon stomatal aperture, transpiration, and photosynthesis. *Plant physiology* **38**, 713 (1963).

[122] De Lima, M. & Grasman, J. Multifractal analysis of 15-min and daily rainfall from a semi-arid region in portugal. *Journal of hydrology* **220**, 1–11 (1999).

[123] Kavasseri, R. G. & Nagarajan, R. A multifractal description of wind speed records. *Chaos, Solitons & Fractals* **24**, 165–173 (2005).

[124] George, S. V. & Ambika, G. Nonlinearity in data with gaps: Application to ecological and meteorological datasets. *Indian Academy of Sciences Conference Series* **1**, 85–91 (2017).

[125] Pezeshki, C., Elgar, S. & Krishna, R. Bispectral analysis of possessing chaotic motion. *Journal of Sound and Vibration* **137**, 357–368 (1990).

[126] Maccarone, T. J. & Coppi, P. S. Higher order variability properties of accreting black holes. *Monthly Notices of the Royal Astronomical Society* **336**, 817–825 (2002).

[127] Ditto, W. *et al.* Experimental observation of a strange nonchaotic attractor. *Physical Review Letters* **65**, 533 (1990).

[128] Heagy, J. & Ditto, W. Dynamics of a two-frequency parametrically driven duffing oscillator. *Journal of Nonlinear Science* **1**, 423–455 (1991).

[129] Newman, M. E. Power laws, pareto distributions and zipf's law. *Contemporary physics* **46**, 323–351 (2005).

[130] George, S. V., Ambika, G. & Misra, R. Detecting dynamical states using bicoherence function. In *Conference on Nonlinear Systems & Dynamics IISER Kolkata*, vol. 16, 18 (2016).

[131] George, S. V., Ambika, G. & Misra, R. Detecting dynamical states from noisy time series using bicoherence. *Nonlinear Dynamics* **89**, 465–479 (2017).

[132] Koch, D. G. *et al.* Kepler: a space mission to detect earth-class exoplanets. In *Space telescopes and instruments V*, vol. 3356, 599–608 (International Society for Optics and Photonics, 1998).

[133] Basri, G., Borucki, W. J. & Koch, D. The kepler mission: A wide-field transit search for terrestrial planets. *New Astronomy Reviews* **49**, 478–485 (2005).

[134] Plachy, E., Bódi, A. & Kolláth, Z. Chaotic dynamics in the pulsation of df cygni, as observed by kepler. *Monthly Notices of the Royal Astronomical Society* **481**, 2986–2993 (2018).

[135] Kolláth, Z., Molnár, L. & Szabó, R. Period-doubling bifurcation and high-order resonances in rr lyrae hydrodynamical models. *Monthly Notices of the Royal Astronomical Society* **414**, 1111–1118 (2011).

[136] Smolec, R. & Moskalik, P. Chaos in hydrodynamic bl herculis models. *Monthly Notices of the Royal Astronomical Society* **441**, 101–115 (2014).

[137] Lindner, J. F. *et al.* Simple nonlinear models suggest variable star universality. *Physica D: Nonlinear Phenomena* **316**, 16–22 (2016).

[138] Kohar, V., Lindner, J. F., Kia, B. & Ditto, W. L. Spectral scaling analysis of rr lyrae stars in ogle-iv galactic bulge fields. In *International Conference on Applications in Nonlinear Dynamics*, 65–76 (Springer, 2016).

[139] Levin, J. Gravity waves, chaos, and spinning compact binaries. *Physical review letters* **84**, 3515 (2000).

[140] Suková, P., Grzedzielski, M. & Janiuk, A. Chaotic and stochastic processes in the accretion flows of the black hole x-ray binaries revealed by recurrence analysis. *Astronomy & Astrophysics* **586**, A143 (2016).

[141] Plachy, E. *et al.* First observations of w virginis stars with k2: detection of period doubling. *Monthly Notices of the Royal Astronomical Society* stw2703 (2016).

[142] Rebusco, P., Moskalik, P., Kluźniak, W. & Abramowicz, M. Period doubling and non-linear resonance in the black hole candidate igr j17091-3624? *Astronomy & Astrophysics* **540**, L4 (2012).

[143] Prša, A. *et al.* Kepler eclipsing binary stars. i. catalog and principal characterization of 1879 eclipsing binaries in the first data release. *The Astronomical Journal* **141**, 83 (2011).

[144] Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Farmer, J. D. Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena* **58**, 77–94 (1992).

[145] Schreiber, T. & Schmitz, A. Surrogate time series. *Physica D: Nonlinear Phenomena* **142**, 346–382 (2000).

[146] Tejera, E., Nieto-Villar, J. & Rebelo, I. Unexpected heart rate variability complexity in the aging process of arrhythmic subjects. *Communications in Nonlinear Science and Numerical Simulation* **15**, 1858–1863 (2010).

[147] Jochner, S., Sparks, T. H., Laube, J. & Menzel, A. Can we detect a nonlinear response to temperature in european plant phenology? *International Journal of Biometeorology* **60**, 1551–1561 (2016).

[148] Esper, J., Büntgen, U., Frank, D. C., Nievergelt, D. & Liebhold, A. 1200 years of regular outbreaks in alpine insects. *Proceedings of the Royal Society B: Biological Sciences* **274**, 671–679 (2006).

[149] Wu, G.-Q. *et al.* Chaotic signatures of heart rate variability and its power spectrum in health, aging and heart failure. *PloS one* **4**, e4323 (2009).

[150] Iler, A. M., Høye, T. T., Inouye, D. W. & Schmidt, N. M. Nonlinear flowering responses to climate: are species approaching their limits of phenological change? *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **368**, 20120489 (2013).

[151] Iyengar, S. V., Balakrishnan, J. & Kurths, J. Co-existence of periodic bursts and death of cycles in a population dynamics system. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **26**, 093111 (2016).

[152] Terrell, D., Gross, J. & Cooney, W. R. A bvrcic survey of w ursae majoris binaries. *The Astronomical Journal* **143**, 99 (2012).

[153] Liu, L., Qian, S. & Xiong, X. A new mechanism of long-term period variations for w uma-type contact binaries. *Monthly Notices of the Royal Astronomical Society* **474**, 5199–5205 (2017).

[154] Kirk, B. *et al.* Kepler eclipsing binary stars. vii. the catalog of eclipsing binaries found in the entire kepler data set. *The Astronomical Journal* **151**, 68 (2016).

[155] Skelton, P. L. & Smits, D. Modelling of w uma-type variable stars. *South African Journal of Science* **105**, 120–126 (2009).

[156] Gazeas, K. & Stepien, K. Angular momentum and mass evolution of contact binaries. *Monthly Notices of the Royal Astronomical Society* **390**, 1577–1586 (2008).

[157] George, S. V., Misra, R. & Ambika, G. Fractal measures and nonlinear dynamics of over-contact binaries. *arXiv preprint arXiv:1805.08351* (2018).

[158] Jacob, R., Harikrishnan, K., Misra, R. & Ambika, G. Recurrence network measures for hypothesis testing using surrogate data: Application to black hole light curves. *Communications in Nonlinear Science and Numerical Simulation* **54**, 84–99 (2018).

[159] Gao, Z.-K. *et al.* Reconstructing multi-mode networks from multivariate time series. *EPL (Europhysics Letters)* **119**, 50008 (2017).

[160] Ouyang, G., Li, X., Dang, C. & Richards, D. A. Using recurrence plot for determinism analysis of eeg recordings in genetic absence epilepsy rats. *Clinical neurophysiology* **119**, 1747–1755 (2008).

[161] Adegoke, O., Dhang, P., Mukhopadhyay, B., Ramadevi, M. & Bhattacharya, D. Correlating non-linear properties with spectral states of rxte data: possible observational evidences for four different accretion modes around compact objects. *Monthly Notices of the Royal Astronomical Society* **476**, 1581–1595 (2018).

[162] Donner, R. V. *et al.* Recurrence-based time series analysis by means of complex network methods. *International Journal of Bifurcation and Chaos* **21**, 1019–1046 (2011).

[163] George, S. V., Misra, R. & Ambika, G. Nonlinear dynamics of rrc lyrae stars. In *NODYCON 2019*, 809–810 (2019).

[164] Jacob, R., Harikrishnan, K., Misra, R. & Ambika, G. Cross over of recurrence networks to random graphs and random geometric graphs. *Pramana* **88**, 37 (2017).

[165] MacKay, D. J. & Mac Kay, D. J. *Information theory, inference and learning algorithms* (Cambridge university press, 2003).

[166] Godavarthi, V., Unni, V., Gopalakrishnan, E. & Sujith, R. Recurrence networks to study dynamical transitions in a turbulent combustor. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **27**, 063113 (2017).

Index