Xenia Lainscsek, BSc.

Delay differential analysis for automatic sleep scoring and sleep apnea classification from a single electrode EEG

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieurin

Master's degree programme: Technical Physics

submitted to

Graz University of Technology

Supervisor

Univ. -Prof. Dipl. -Phys. Dr.rer.nat Wolfgang von der Linden

Institut für Theoretische Physik - Computational Physics

in cooperation with ZOMA Health Inc.

Graz, October 2019

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

19.09.2019

Date

Signature

Abstract

Sleep is of utmost importance in order to maintain a healthy brain and body and a continuous monitoring of sleep could help not only people suffering from sleep disorders, but could also contribute the overall well-being of the general population. In order to understand and classify sleep architecture, computer algorithms are implemented to try and capture dynamical information present in biological time series data. Naturally occurring systems, such as the human brain, are nonlinear in nature and exhibit chaotic, deterministic behavior. For this reason, nonlinear approaches are advantageous when dealing with such data. Here, a novel nonlinear analysis technique, delay differential analysis (DDA), will be introduced and applied to single electrode sleep electroencephalography (EEG) data with the goal of creating an automatic sleep stage detection algorithm as well as finding a biomarker for obstructive sleep apnea (OSA) through mere EEG signal. Ultimately, the hope is to be able to implement this algorithm into a wearable home EEG sleep monitoring device, as traditional polysomnography (PSG) is extremely expensive and time consuming.

Acknowledgement

I would like to express my sincerest gratitude to ...

... my thesis advisor, Prof. Wolfgang von der Linden for his unconditional support and for giving me the opportunity to delve into a rather unconventional topic for a masters thesis in technical physics.

... ZOMA Health Inc. for providing me with a summer internship position.

... Prof. Christophe Letellier, who took a lot of his time to help me grasp the basic theoretical concepts discussed in this thesis.

... my family for their love and support, both emotionally and financially, throughout my entire studies.

... my friends, for their support and for providing relief in times I felt under great pressure. You made my student days both stimulating and unforgettable.

... and lastly my mother, for fascinating me to explore science and for inspiring me in more ways than I could think possible.

Contents

Outline **1**

1 Nonlinear Dynamics and Chaotic Systems **2**
1.1 Nonlinear Dynamical Systems 2
1.1.1 Spring Mass System 5
1.1.2 Power Spectrum 15
1.2 Bifurcation theory 16
1.2.1 Saddle-Node Bifurcation 17
1.2.2 Pitchfork Bifurcation 18
1.2.3 Transcritical Bifurcation 19
1.2.4 Hopf Bifurcation 21

2 Delay Differential Analysis **24**
2.1 Delay Differential Embedding 25
2.2 Repeated Random Subsampling k-fold Cross Validation 30
2.2.1 Receiver Operating Characteristic Curve 34

3 Automatic Sleep Scoring **35**
3.1 The Sleep Cycle 35
3.1.1 N1 NREM Sleep (R & K S1) 36
3.1.2 N2 NREM Sleep (R & K S2) 36
3.1.3 N3 NREM Sleep (R & K S3 and S4) 37
3.1.4 REM Sleep 37
3.2 ZOMA Health Inc. 38
3.3 Polysomnography 39
3.4 DDA for automatic sleep scoring 40
3.4.1 DDA outputs as marker for sleep stage classification . . 40
3.5 Datasets 41

3.5.1 MUSE EEG Headband 43
3.6 Results 44
3.6.1 SHHS 44
3.6.2 Physionet 47

4 Sleep Apnea Classification **49**
4.1 Apnea Hypopnea Index 49
4.2 Obstructive Sleep Apnea 50
4.3 Results 52

5 Conclusion **58**
5.1 Outlook 59

Outline

Delay differential analysis (DDA) is a tool based on nonlinear dynamics designed to capture large-scale dynamics present in time series data. In the first chapter of this thesis, an introduction to basic nonlinear dynamics and chaos is given, by taking a look at the state portraits of a few simple deterministic dynamical systems. These ideas are then further extended to embedding theory in chapter 2, which lies at the heart of DDA. This concept will be exemplified with delay and differential embeddings of the duffing equation. Since the governing equations of a complex system, such as the human brain, are unknown, the goal of DDA is to detect and classify the dynamical information contained in the system rather than try to model the system itself. The basic computational procedure of DDA and the statistics used to interpret the outputs will also be discussed chapter 2. Chapter 3 will introduce the physiology of the human sleep cycle and will present the results of using DDA on sleep EEG data. Two data sets were used, one of which contained patients suffering from obstructive sleep apnea, which is a condition that causes a partial or complete obstruction of breathing during sleep. Chapter 4 will demonstrate the classification performance of DDA on patients with varying sleep apnea severities. Chapter 5 will give a brief summary of all the obtained results presented in this thesis.

Chapter 1

Nonlinear Dynamics and Chaotic Systems

1.1 Nonlinear Dynamical Systems

In physics, the dynamics of a system is governed by a set of nonlinear differential equations. The time evolution of the variables included in this set, which span the n-dimensional *state space* of the system, depend on their initial conditions as well as on a set of constant parameters. The collection of trajectories produced by this system is called the *state portrait* and they exhibit, depending on parameter choice and initial conditions, steady-state (singular point), periodic, quasi periodic or chaotic behavior. Nonlinear systems are sensitive to initial conditions and neighboring trajectories who stay close to each other over a long finite period of time suddenly exponentially diverge.[18,70]

When dealing with highly complex nonlinear systems, such as the human brain, we do not have access to all the system variables and rely on global modeling techniques to get insight into the underlying dynamics of the system. Since such systems (e.g. the brain) are highly unlinear, embeddings can be used[33,66,83,77]. The Whitney embedding theorem (1936) states that a generic map from an n-dimensional manifold to a $2n+1$ dimensional euclidean space is an embedding[33]. This implies that each state space can be uniquely identified by $2n+1$ measurements and therefore reconstructed.

The Takens embedding theorem (1981) further states that the same can be done with a single measurement[83]. Rather than $2n+1$ generic signals,

$2n + 1$ delay embeddings (time delayed versions) of one generic signal are used to construct the embedding therefore reconstructing the state space.

Around the same time, similar theorems were published by Aeyels (1981)[2] and Packard (1980)[66]. These theorems were generalized in 1991 by Sauer et al.[77]. Similar to delay embeddings, differential embeddings can be defined by using $2n + 1$ successive derivatives of the signal.

DDA then combines the differential embedding with nonuniform functional embeddings, which are a generalization of traditional embeddings, to get insight into the overall nonlinear dynamical structure of the system. DDA does not strive to reconstruct the original dynamical system, but to detect and/or classify dynamical structure. To obtain this information, the term on the left hand side of a DDE is the differential embedding. The right hand side is the delay embedding, which is a small term polynomial approximation model (in this thesis 3-term model) containing low dimensional embeddings (typically two delayed versions of the signal). These ideal will be further discussed in chapter 2, let us first focus on more simple low dimensional oscillating systems.

The singular points of a dynamical system **S** are the points where there is no change in the system, i.e. the points at which the derivatives of the system variables are zero. The state portrait of a system for a certain choice of parameters can be obtained by determining the solution of the system at various initial conditions. This can however, more quickly be achieved, by investigating the stability of these singular points. The characteristic polynomial of the Jacobian matrix $\hat{J}$ of the system results in characteristic eigenvalues that, depending on which number set they belong to, correspond to the type of stability of a particular singular point from which the surrounding trajectory field can be inferred.

The term *attractor* was introduced in the second half of the 20th century by mathmeticians like Coddington and Levinson (1955) and Mendelson (1960)[64] although the limit cycle was already discovered by Poincaré in the 19th century[70]. An attractor is a point or a set of points in the state space, asymptotically reached by trajectories, which results in a geometrical structure depicting the long term behavior of systems that remain near the singular points. There exist also chaotic attractors who are produced by 3-dimensional dynamical systems in the absence of any singular points, they are described in further detail in the case of the Wei system[86] by Letellier and Malasoma (2016)[56]. The transient region in a state portrait is any part of the trajectories between the inital conditions and the attractor. Attractors

consiting of more than one point seem to have first been published in 1964 by Auslander, Bhatia and Seibert[64]. There are four types of attractors: a) point or fixed point attractor, b) limit cycle, c) torus attractor, d) strange attractor and the e) chaotic attractor which arises from physical systems (realistic and idealistic) who's trajectories a) evolve towards a fixed point like the center bottom position of a damped pendulum b) evolve towards a closed trajectory like those of oscillating systems c) evolve around a fixed point as a limit cycle with more than one frequency like planets orbiting a star while rotating about themselves as well and d & e) are structured near their fixed points in a bizarre fashion respectively. Most of the time, a chaotic attractors are also strange attractors, but there exist both non-chaotic strange attractors, and chaotic non-strange attractors. The term strange refers to the geometry or shape of the attracting set and such an attractor has a fractal dimension. The term chaotic refers to the dynamics of the trajectories, or orbits, on the attractor and such an attractor is sensitive to initial conditions[31].

These concepts are to be exemplified by means of damped and driven oscillators[18,37].

1.1.1 Spring Mass System

Fig. 1.1 represents a simple dynamical system without friction and only one degree of freedom. The motion of the mass can be described by a linear differential equation for a harmonic oscillator

$$\ddot{\theta} + \omega_0^2 \theta = 0 \tag{1.1}$$

where x is the deflection of the mass from its starting point 0 and $\omega_0 = \sqrt{k/m}$ is the resonance frequency of the oscillation. This can be rewritten as a set of first order differential equations with $x = \theta$

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\omega_0^2 x \end{aligned} \tag{1.2}$$

Setting Eq. 1.2 equal to zero we get the singular points **S**

$$\mathbf{S_0} = \left| \begin{aligned} x_0 &= 0 \\ y_0 &= 0 \end{aligned} \right. \tag{1.3}$$

The Jacobian matrix $\hat{J}$ is

$$\hat{J} = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{bmatrix} \tag{1.4}$$

The characteristic polynomial of $\hat{J}$ is

$$\lambda^2 + \omega_0^2 = 0 \tag{1.5}$$

solving for Eq. 1.5 results in the eigenvalues

$$\lambda_{\pm} = \pm i\omega_0 \tag{1.6}$$

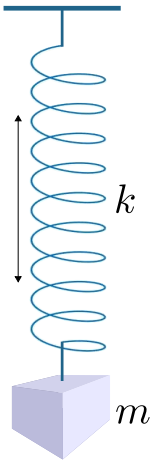

Figure 1.1: Mass on spring

Poincaré investigated the stability of 2-dimensional dynamical systems in relationship to the global aspect of trajectories[70]. For 2-dimensional systems, four types of singular points are classified namely: node, center, focus and saddle-point[74]. Fig. 1.2 depicts the state portraits of such systems according to the eigenvalues of the Jacobian of the system at hand.

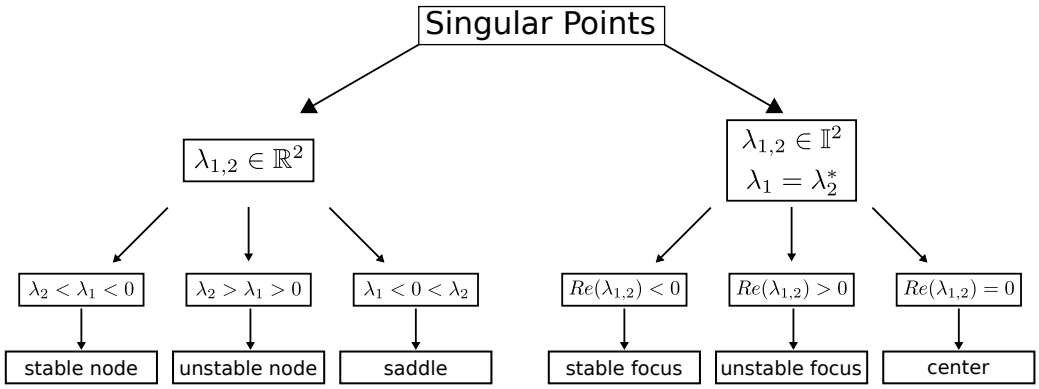

Figure 1.2: Overview concerning the 4 types of singular points for 2-dimensional dynamical systems

In the case of the singular points Eq. 1.6, there exist only imaginary parts and the singular points are marginally stable and center.

The resulting state portrait of this two dimensional system is thus a closed loop who's radius depends on the oscillation frequency as well as initial conditions x_0 and y_0. Fig. 1.3 is the solution of the system Eq. 1.2 for initial conditions $x_0 = 0$, $y_0 = 0.5$ and $\omega_0 = 0.5$ solved with ode45 in MATLAB.

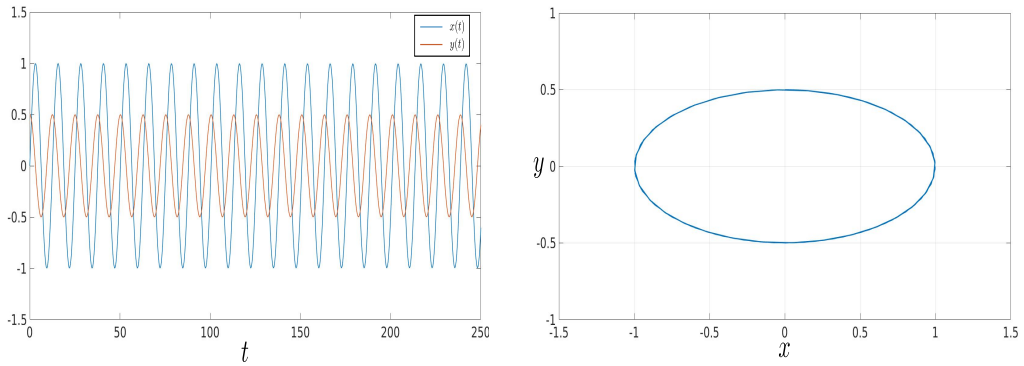

Figure 1.3: Time series of mass on spring system described by Eq. 1.2 (top) and corresponding state portrait (bottom) for initial conditions $x_0 = 0$, $y_0 = 0.5$ and $\omega_0 = 0.5$.

In the case of a damped oscillator with constant damping coefficient ζ Eq. 1.1 becomes

$$\ddot{\theta} + 2\zeta\omega_0\dot{\theta} + \omega_0^2\theta = 0 \tag{1.7}$$

and can be rewritten as a set of first order ordinary differential equations

$$\begin{array}{rcl} \dot{x} & = & y \\ \dot{y} & = & -2\zeta\omega_0 y - \omega_0^2 x \end{array} \tag{1.8}$$

The singular points are the same as in the undamped case

$$\mathbf{S} = \left| \begin{array}{rcl} x & = & 0 \\ y & = & 0 \end{array} \right. \tag{1.9}$$

The Jacobian matrix $\hat{J}$ is:

$$\hat{J} = \left[\begin{array}{cc} 0 & 1 \\ -\omega_0^2 & -2\zeta\omega_0 \end{array} \right] \tag{1.10}$$

The characteristic polynomial is

$$\lambda^2 + 2\zeta\omega_0\lambda + \omega_0^2 = 0 \tag{1.11}$$

with the resulting eigenvalues

$$\lambda_\pm = -\zeta\omega_0 \pm \omega_0\sqrt{\zeta^2 - 1} \tag{1.12}$$

For $0 < \zeta < 1$ (Fig. 1.4 A & B), the eigenvalues become complex conjugated and the singular points are a stable foci. For $\zeta < 0$ these foci would become unstable. For $\zeta \geq 1$ (Fig. 1.4 C & D), the eigenvalues are real and negative and the singular points are stable nodes.

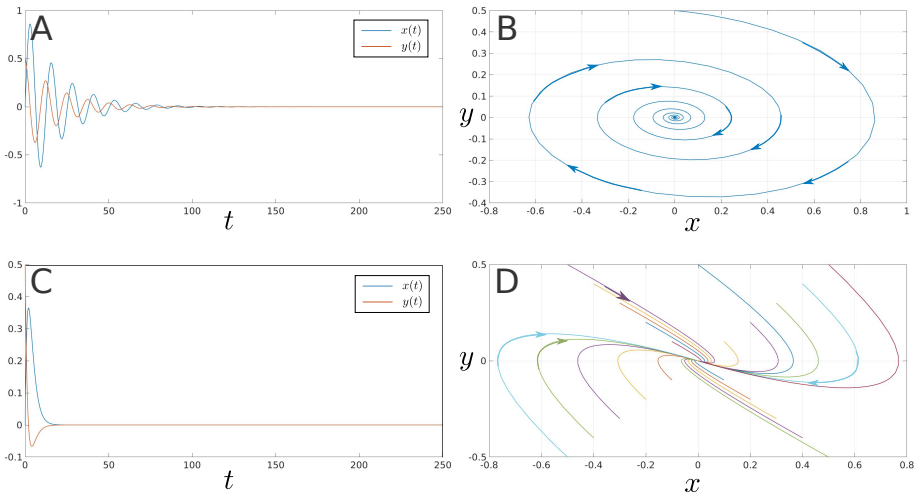

Figure 1.4: **A)** Time series of mass on spring system described by Eq. 1.8 for initial conditions $x_0 = 0$, $y_0 = 0.5$ and parameters $\omega_0 = 0.5$ and $\zeta = 0.1$. **B)** Resulting state portrait is a stable foci. **C)** Time series of mass on spring system described by Eq. 1.8 for initial conditions $(\boldsymbol{x_0}, \boldsymbol{y_0}) = [(0,0); (0.1,0.1); (0.2,0.2); (0.3,0.3); (0.5,0.5); (-0.1,-0.1); (-0.2,-0.2); \ldots (-0.3,-0.3); (-0.5,-0.5); (0.1,-0.1); (0.2,-0.2); (0.3,-0.3); (0.5,-0.5); \ldots (-0.1,0.1); (-0.2,0.2); (-0.3,0.3); (-0.5,0.5)]$ and parameters $\omega_0 = 0.5$ and $\zeta = 1.01$. **D)** Resulting state portrait are stable nodes.

If we now introduce a periodic external force into our system, we can compensate for the damping and the trajectories of the phase space evolve towards a stable closed trajectory or a limit cycle. We end up with an equation of the form

$$\ddot{\theta} + 2\zeta\omega_0\dot{\theta} + \omega_0^2\theta = u \tag{1.13}$$

We can consider $u = \omega_0^2\theta_0^2 \sin\omega_E t$ as a special solution to Eq. 1.1, where ω_E is the frequency of the external force. This is a non-autonomous system meaning the time t explicitly occurs in Eq. 1.13. The state space now is made up of two 2-dimensional systems, resulting in a 4-dimensional state space. The oscillator without driving force is spanned by variables x and y and the driving oscillator adds another two dimensions as was the case with the system described in Eq. 1.2. The driving signal u may be rewritten as a set of

two first order ordinary differential equations[61], with $u = a\sin\omega_E t + b\cos\omega_E t$ and the initial condition $u(t=0) = 0$ we have

$$\begin{array}{rcl} b &=& 0 \\ a &=& \omega_0^2\theta_0^2 \\ \dot{u} &=& \omega_0^2\theta_0^2\omega_E\cos\omega_E t \quad = \quad v \\ \dot{v} &=& -\omega_0^2\theta_0^2\omega_E^2\sin\omega_E t \quad = \quad -\omega_E^2 u \end{array} \tag{1.14}$$

The resulting total autonomous system is then governed by the 4-dimensional set of first order ordinary differential equations[61]

$$\begin{array}{rcl} \dot{x} &=& y \\ \dot{y} &=& -2\zeta\omega_0 y - \omega_0^2 x + u \\ \dot{u} &=& v \\ \dot{v} &=& -\omega_E^2 u \end{array} \tag{1.15}$$

with singular points

$$\mathbf{S} = \left| \begin{array}{rcl} x &=& 0 \\ y &=& 0 \\ u &=& 0 \\ v &=& 0 \end{array} \right. \tag{1.16}$$

and Jacobian matrix $\hat{J}_{\{x,y,u,v\}}$

$$\hat{J}_{\{x,y,u,v\}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\omega_0^2 & -2\zeta\omega_0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\omega_E^2 & 0 \end{bmatrix} \tag{1.17}$$

Since the driving oscillator is independent of the driven one, the stability analysis can be performed in two steps. In the subspace $\mathbb{R}^2(u,v)$ the characteristic polynomial of the Jacobian matrix $\hat{J}_{\{uv\}}$ is

$$\lambda^2 + \omega_E^2 = 0 \tag{1.18}$$

with eigenvalues

$$\lambda_{\pm\{u,v\}} = \pm i\omega \tag{1.19}$$

so the singular point in the uv-plane is again a center point (see Fig. 1.5). For the subspace $\mathbb{R}^2(x,y)$ the characteristic polynomial of the Jacobian matrix $\hat{J}_{\{xy\}}$ is

$$\lambda^2 + 2\zeta\omega_0\lambda + \omega_0^2 = 0 \tag{1.20}$$

with the same eigenvalues as before in Eq. 1.12

$$\lambda_{\pm\{x,y\}} = -\zeta\omega_0 \pm \omega_0\sqrt{\zeta^2 - 1} \tag{1.21}$$

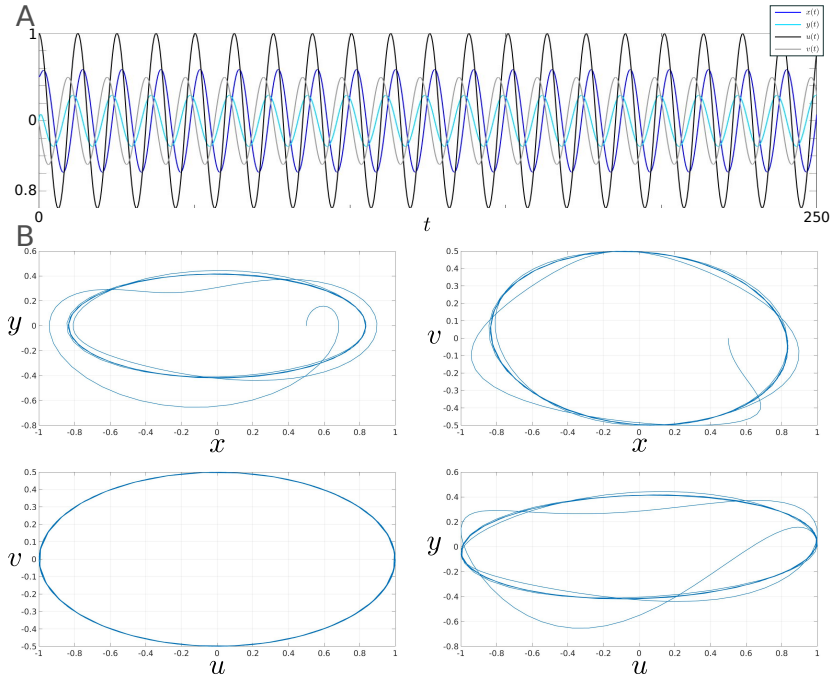

Figure 1.5: **A)** Time series of mass on spring system described by Eq. 1.15 for initial conditions $x_0 = 0.5$, $y_0 = 0$, $u_0 = 1$ and $v_0 = 0$ and for parameters $\omega_0 = 1.2$, $\zeta = 0.1$, $\omega_E = 0.5$. **B)** Corresponding state portrait of 4-dimensional state space is a limit cycle

Linear systems of equations of the sort Eq. 1.13 can be inegrated and their solution trajectories are dependent on the initial conditions. The system is semi-conservative, i.e. the driving system is conservative and feeds the dissipative driven system with energy. The amplitude of the driven system strongly depends on the driving one and the state portrait of the system is represented as a limit cycle. If we now regard the nonlinear case, we find ourselves dealing with far more complex systems with bizarre state space

representations. An example of a simple nonlinear system showing chaotic behavior is given by the Duffing-Equation[32,37]

$$\ddot{\theta} + 2\zeta\omega_0\dot{\theta} - \omega_0^2\theta + \delta\omega_0^2\theta^3 = \omega_0^2\theta_0^2 \sin\omega_E t \tag{1.22}$$

with the nonlinear damping constant δ. It is important to note here, that the nonlinear cubic term θ^3 is a necessary but not sufficient condition for a chaotic system[37]. In 1901, Bendixson, based on Poincarés work, proved that in 2-dimensions, the asymptotic behavior in a state portrait can only be one of four types of singular points (Fig. 1.2) or a limit cycle (Fig. 1.5). The Bendixson-Poincaré theorem states that as a corollary of this, at least 3-dimensions are required for a system to give rise to a chaotic attractor[9,70]. The following modified Duffing equation will be used for further investigation[37]

$$\ddot{\theta} + c\dot{\theta} - \beta\theta + \alpha\theta^3 = \gamma \sin\omega_E t \tag{1.23}$$

Writing this in a set of 4 first order ordinary differential equations

$$\begin{array}{rcl} \dot{x} &=& y \\ \dot{y} &=& -cy + \beta x - \alpha x^3 + u \\ \dot{u} &=& v \\ \dot{v} &=& -\omega_E^2 u \end{array} \tag{1.24}$$

with two singular points $\boldsymbol{S}_0$ and $\boldsymbol{S}_\pm$

$$\boldsymbol{S}_0 = \left| \begin{array}{rcl} x_0 &=& 0 \\ y_0 &=& 0 \\ u_0 &=& 0 \\ v_0 &=& 0 \end{array} \right. \qquad \boldsymbol{S}_\pm = \left| \begin{array}{rcl} x_\pm &=& \pm\sqrt{\beta/\alpha} \\ y_\pm &=& 0 \\ u_\pm &=& 0 \\ v_\pm &=& 0 \end{array} \right. \tag{1.25}$$

and Jacobian matrix $\hat{J}_{\{x,y,u,v\}}$

$$\hat{J}_{\{x,y,u,v\}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \beta - 2\alpha x^2 & -c & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\omega_E^2 & 0 \end{bmatrix} \tag{1.26}$$

The characteristic polynomial of $\hat{J}_{\{u,v\}}$ is

$$\lambda^2 + \omega_E^2 = 0 \tag{1.27}$$

with eigenvalues

$$\lambda_{\pm\{u,v\}} = \pm i\omega \tag{1.28}$$

The characteristic polynomial of $\hat{J}_{\{x,y\}}$ is dependent on the x-component of the singular points.

$$\lambda^2 + c\lambda - \beta + 2\alpha x^2 = 0 \tag{1.29}$$

In $\boldsymbol{S}_0$ the eigenvalues are

$$\lambda_{\pm\{x,y\}_0} = \frac{-c \pm \sqrt{c^2 + 4\beta}}{2} \tag{1.30}$$

since $c^2 + 4\beta > 0$, $\lambda_{\pm\{x,y\}_0} \in \mathbb{R}$, the singular point $\boldsymbol{S}_0$ is a saddle. The other two singular points $\boldsymbol{S}_\pm$ are symmetric under an inversion symmetry $\boldsymbol{S}_+ = \hat{\Psi}\boldsymbol{S}_-$ where $\hat{\Psi}$ is the negative of the identity matrix $\hat{\Psi} = -\hat{\mathbb{1}}$. Thus the type of these singular points is the same and we will consider only $\boldsymbol{S}_+$.

$$\lambda_{\pm\{x,y\}_+} = \frac{-c \pm \sqrt{c^2 - 4\beta}}{2} \tag{1.31}$$

So depending on the parameter c, we will get a different state portrait. There exist an infinite number of attractors in the four dimensional state space which exhibit period doubling bifurcations as well as chaotic behavior. Some examples of chaotic as well as non chaotic solutions are depicted in Fig. 1.6 by varying the initial condition $u_0 = \gamma$ in Eq. 1.23.

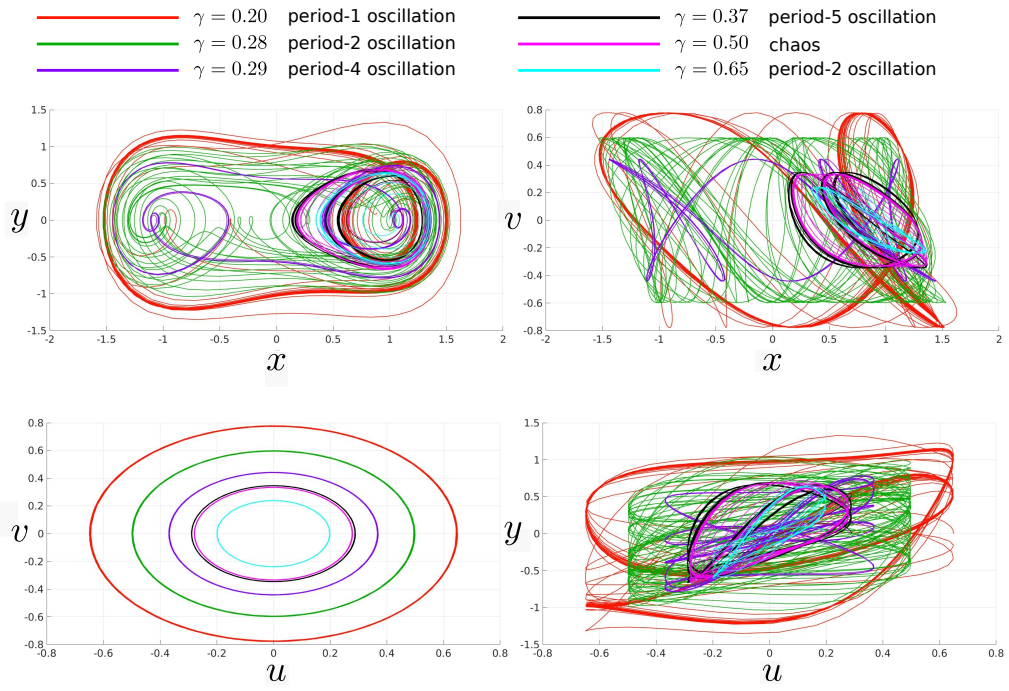

Figure 1.6: State portraits of the system given by Eq. 1.23 for parameters $c = 0.3$, $\beta = 1$, $\alpha = 1$, $\omega_E = 1.2$ and initial conditions $x_0 = 1$, $y_0 = 0$, $v_0 = 0$ and varying $u_0 = \gamma = [0.20, 0.28, 0.29, 0.37, 0.50, 0.65]$

The conservative component of the driving signal is obvious in the $u - v$ plane while the chaotic behavior is obvious in the $x - y$ plane. One can observe a large change in behavior occurring when the initial condition of the driving oscillator is slightly varied. This is typical for a chaotic system and it is impossible to determine what its trajectories for various initial conditions will look like. The state portrait of x and y can reveal some spectacular geometrical shapes. Two more examples for a quasi-periodic regime are shown in Fig. 1.7.

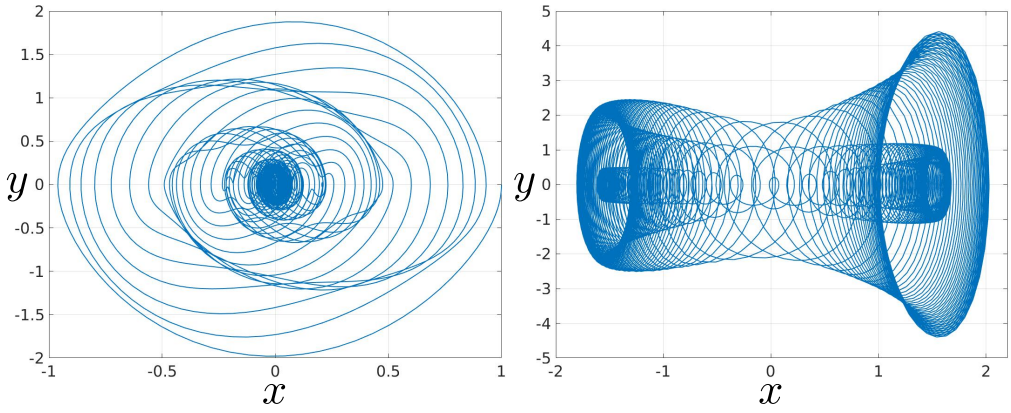

Figure 1.7: Quasi periodic state portraits of the system given by Eq. 1.23 for parameters $c = 0.05$, $\beta = -3$, $\alpha = .1$, $\omega_E = 5$ and initial conditions $x_0 = 1$, $y_0 = 0$, $v_0 = 0$ and varying $u_0 = \gamma = 0.9$ (left) and $c = 0.02$, $\beta = -10$, $\alpha = 9$, $\omega_E = 0.04$ and initial conditions $x_0 = 1$, $y_0 = 0$, $v_0 = 0$ and varying $u_0 = \gamma = 50$

A chaotic attractor, unlike its non chaotic counterparts, does not form a manifold embedded in euclidean space. Both mathematical concepts such as autocorrelation and fourier power sepctrums as well as geometric and numerical methods such as Lypanouv-exponents and fractal dimension can be used to classify chaotic attractors, although it should be noted here, that these common tools are typically unsuitable for studying natural physiological phenomena. Another example of such a nonlinear system exhibiting regimes of chaotic behaviour is the famous Lorenz system which is a simplified 3-dimensional model for Regleigh-Bérnard convection and is further introduced in[54].

The Poincaré map is a tool used to characterize the nature of the solution in the state space of a continuous n-dimensional system and was first described in 1881 by Hérni Poincaré[70]. It is the discretization of the state space in which an $(n-1)$-dimensional hyperplane, the Poincaré section, Σ is assumed, through which a returning trajectory s transversely crosses. For a system of differential equations of the form Eq. $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$, the right hand side of the equations determine the direction of the tangent of the trajectory s in the point $\mathbf{x}$[37]. In this way stability of a periodic orbit can be investi-

gated in the $(n-1)$-dimensional subspace and since the stability of one point of a periodic orbit holds for all other points of the orbit, it is sufficient to investigate only one arbitrary point.

1.1.2 Power Spectrum

Limit cycle and torus attractors can be detected by taking the Fourier transform of the respective periodic or quasiperiodic regimes of the time series $x(t)$. Assuming a periodic time series $x(t)$ with frequency ω_0, the resulting power spectrum is a series of equidistant pulses, with highest amplitudes at $\omega_0 = \frac{2\pi}{T}$. Assuming a quasiperiodic time series $x(t)$ with two incommensurable frequencies ω_1 and ω_2, the power spectrum will have the highest peaks at ω_1 and ω_2 as well as lesser peaks at all other linear combinations of the two.

The power spectrum of a chaotic system on the other hand, will not show distinct peaks at certain frequencies, but a continuous band of frequencies, with possible characteristic peaks[23]. Fig. 1.8 depicts this phenomenon for the chaotic Duffing system described by Eq. 1.23 for $c = 0.3$, $\beta = 1$, $\alpha = 1$, $\omega_E = 1.2$ and initial conditions $x_0 = 1$, $y_0 = 0$, $v_0 = 0$ and $u_0 = \gamma = 0.5$ which corresponds to the chaotic attractor in the state portraits of Fig. 1.6. The autocorrelation function shows correlation between a signal $x(t)$ and $x(t-\tau)$ which fades after long enough time. Stochastic white noise on the other hand, will show no autocorrelation and a constant Fourier spectrum dependent on the variance $P_i = \sigma^2$.

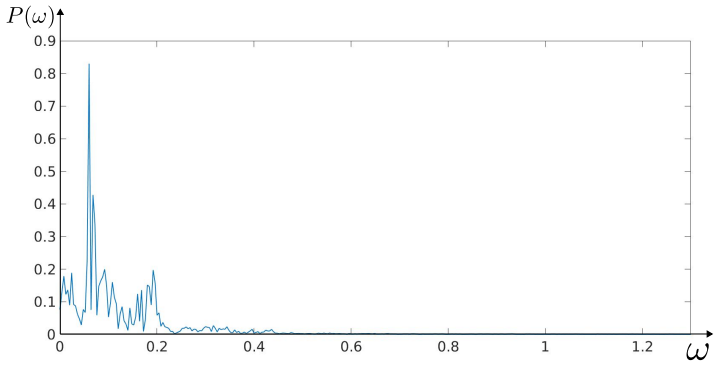

Figure 1.8: Power spectrum of Duffing system for $c = 0.3$, $\beta = 1$, $\alpha = 1$, $\omega_E = 1.2$ and initial conditions $x_0 = 1$, $y_0 = 0$, $v_0 = 0$ and $u_0 = \gamma = 0.5$.

Testing for periodic, quasi-periodic and chaotic behavior in an unknown time series can be difficult and there are better criteria for classifying chaotic behavior such as the Lypunov exponents.

A. Lyaounov published a detailed paper in 1892 in which he studied stability of systems having no closed form solutions by introducing the concept of certain characteristic exponents. Lypanouv exponents are used as convergence (or divergence) criteria of neighboring trajectories in the state space. The trajectories of point, limit cycle and torus attractors are stable and neighboring trajectories stay close to each other. Although the trajectories of strange attractors also experience an attractive force in the state space, initially close trajectories diverge exponentially[37].

1.2 Bifurcation theory

As previously discussed, small changes in the parameters of a systems can drastically influence the long term behavior of a system. The term bifurcation was first used in 1885 in Poincares paper[71] and is defined as

Definition 1 *a small change in one or more parameter of a system, inducing a sudden change in the overall topological behavior[37].*

If the change of a system $x(t)$, on account of a change in a parameter a of the system, occurs continuously, then we call the pair $\{a, x(t)\}$ a bifurcation

pair[82]. A bifurcation diagram depicts this change in the stability of the solutions with regards to the parameter being varied. With every bifurcation point come new equilibrium solutions which have drastic consequences on systems. The Poincaré map described in the previous section is a tool with which a bifurcation diagram can be derived and is described for the Lorenz system[57] in more detail in[54].

The main bifurcations of two dimension systems include saddle-node bifurcation, pitchfork bifurcation and hopf bifurcation which will be described in the following sections.

1.2.1 Saddle-Node Bifurcation

The saddle-node bifurcation is the simplest example and will be illustrated by the simple two-dimensional dissipative system

$$\begin{array}{rcl} \dot{x} & = & \alpha - x^2 \\ \dot{y} & = & -y \end{array} \tag{1.32}$$

with the parameter α. The two singular points are

$$\boldsymbol{S}_{\pm} = \left| \begin{array}{rcl} x_{\pm} & = & \pm\sqrt{\alpha} \\ y_{\pm} & = & 0 \end{array} \right. \tag{1.33}$$

and symmetric so that only one needs to be considered. The corresponding eigenvalues are

$$\left.\begin{array}{rcl} \lambda_{1\pm} & = & -2x_{\pm} \\ \lambda_{2\pm} & = & -1 \end{array}\right\} \qquad \begin{array}{rcl} \lambda_{1\pm} & = & -2(\pm\sqrt{\alpha}) \\ \lambda_{2\pm} & = & -1 \end{array} \tag{1.34}$$

Since $0 > \lambda_{1+} > \lambda_{2+}$, $\boldsymbol{S}_+$ is a stable node singular point, and since $\lambda_{2-} < 0 < \lambda_{1-}$, $\boldsymbol{S}_-$ is a saddle singular point.

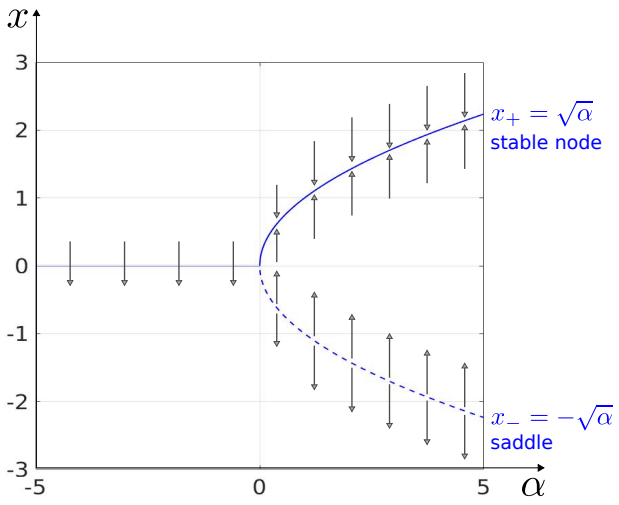

Figure 1.9: Saddle node bifurcation diagram of system described by Eq. 1.32.

Fig. 1.9 depicts the saddle node bifurcation of Eq. 1.32. By convention, dashed lines represent unstable singular points and bold lines stable singular points. There are no real solutions for $\alpha < 0$ so the system diverges to $-\infty$. At $\alpha = 0$ a bifurcation point occurs and the solutions collide to zero, two singular points $\boldsymbol{S}_{\pm}$ appear. When the number of solutions at a bifurcation point increases, it is said to be supercritical (in the opposite case it is subcritical) For $\alpha > 0$ the behavior depends on the initial condition of x. If $x < -\sqrt{\alpha}$, the system diverges to $-\infty$, if $x > -\sqrt{\alpha}$, the system converges to $\sqrt{\alpha}$.

1.2.2 Pitchfork Bifurcation

The next main bifurcation pattern is called a pitchfork bifurcation. For this we will now consider the cubic 2-dimensional nonlinear system of ordinary differential equations

$$\begin{aligned} \dot{x} &= x\alpha - x^3 \\ \dot{y} &= -y \end{aligned} \tag{1.35}$$

The three singular points are

$$\boldsymbol{S}_0 = \left| \begin{array}{rcl} x_0 &=& 0 \\ y_0 &=& 0 \end{array} \right. \qquad \boldsymbol{S}_\pm = \left| \begin{array}{rcl} x_\pm &=& \pm\sqrt{\alpha} \\ y_\pm &=& 0 \end{array} \right.$$

The corresponding eigenvalues explicitly depend on the x-value of the singular points

$$\left. \begin{array}{rcl} \lambda_{1_{\{0,\pm\}}} &=& -3x_i^2 + \alpha \\ \lambda_{2_{\{0,\pm\}}} &=& -1 \end{array} \right\} \quad \begin{array}{rcl} \lambda_{1_{\{0\}}} &=& \alpha \\ \lambda_{2_{\{0\}}} &=& -1 \end{array} \qquad \begin{array}{rcl} \lambda_{1_{\{\pm\}}} &=& -2\alpha \\ \lambda_{2_{\{\pm\}}} &=& -1 \end{array} \tag{1.36}$$

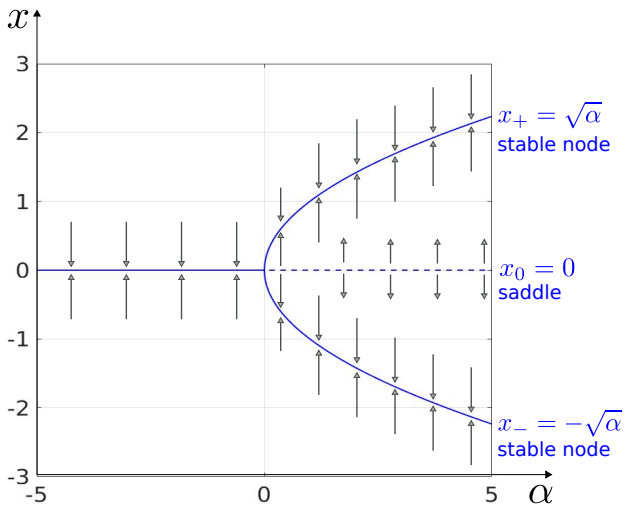

Figure 1.10: Pitchfork bifurcation diagram of system described by Eq. 1.35.

The singular point $\boldsymbol{S}_0$ exists for any value α. For $\alpha < 0$ we have two negative eigenvalues, $\boldsymbol{S}_0$ is a stable node. For $\alpha > 0$, $\lambda_{2_{\{0\}}} < 0 < \lambda_{1_{\{0\}}}$ and we get a saddle node. The singular points $\boldsymbol{S}_\pm$ have identical eigenvalues and exist for $\alpha \geq 0$. The two resulting eigenvalues are both negative and thus the singular points $\boldsymbol{S}_\pm$ are stable nodes.

1.2.3 Transcritical Bifurcation

In a transcritical bifurcation, the number of solutions does not change. The stability of the solution swaps, the stable node switch to a saddle node and vice versa. We consider the following set of 2-dimensional ordinary differen-

tial equations

$$\begin{aligned} \dot{x} &= x\alpha - x^2 \\ \dot{y} &= -y \end{aligned} \tag{1.37}$$

The two singular points are

$$\boldsymbol{S}_0 = \left| \begin{array}{rcl} x_0 &=& 0 \\ y_0 &=& 0 \end{array} \right. \qquad \boldsymbol{S}_* = \left| \begin{array}{rcl} x_* &=& \alpha \\ y_* &=& 0 \end{array} \right.$$

The corresponding eigenvalues are again explicitly dependent on the x-value of the singular points

$$\lambda_{1,2_{\{0,*\}}} = \underbrace{\frac{-(2x_i + 1 - \alpha) \pm \sqrt{4(x_i^2 - x_i\alpha - x_i) + \alpha^2 + 2\alpha + 1}}{2}} \tag{1.38}$$

$$\begin{aligned} \lambda_{1_{\{0\}}} &= \alpha \\ \lambda_{2_{\{0\}}} &= -1 \end{aligned} \qquad \begin{aligned} \lambda_{1_{\{*\}}} &= -1 \\ \lambda_{2_{\{*\}}} &= -\alpha \end{aligned} \tag{1.39}$$

Fig. 1.11 depicts this swapping of stability. For $\alpha < 0$ $\boldsymbol{S}_0$ is asymptotically stable and $\boldsymbol{S}_*$ is unstable and for $\alpha > 0$ $\boldsymbol{S}_0$ is unstable and $\boldsymbol{S}_*$ is stable.

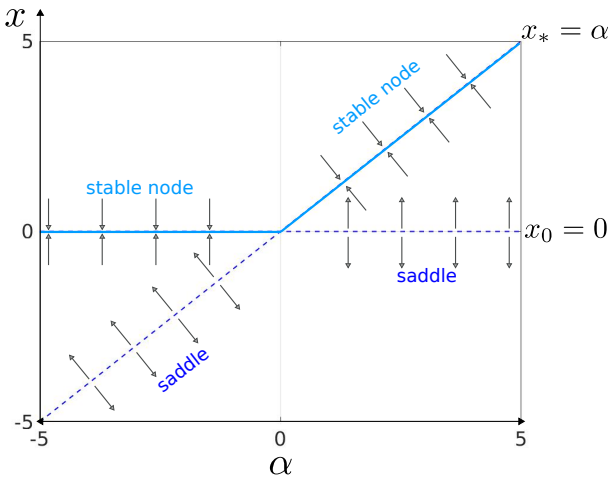

Figure 1.11: Transcritical bifurcation diagram of system described by Eq. 1.37.

1.2.4 Hopf Bifurcation

Up to now, the eigenvalues of the Jacobian matrix of the systems considered were real valued. The more interessting case are those systems exhibiting non real eigenvalues of the Jacobian. Eberhard Hopf studied the bifurcations of such n-dimensional systems[34]. Consider following 2-dimensional system of equations

$$\begin{aligned} \dot{x} &= -y + x\left(\alpha - (x^2 + y^2)\right) \\ \dot{y} &= x + y\left(\alpha - (x^2 + y^2)\right) \end{aligned} \tag{1.40}$$

which can be rewritten into polar coordinates with $x = r\cos\phi$ and $y = r\sin\phi$

$$\begin{aligned} \dot{r} &= \alpha r - r^3 \\ \dot{\phi} &= 1 \end{aligned} \tag{1.41}$$

The singular points are

$$r_0 = 0 \qquad r_* = \sqrt{\alpha} \quad \alpha > 0 \tag{1.42}$$

the latter only holds for positive α values and leads to a limit cycle in the state portrait of the xy-space or in other words a periodic orbit who's amplitude grows with $\sqrt{\alpha}$ (see Fig. 1.12).

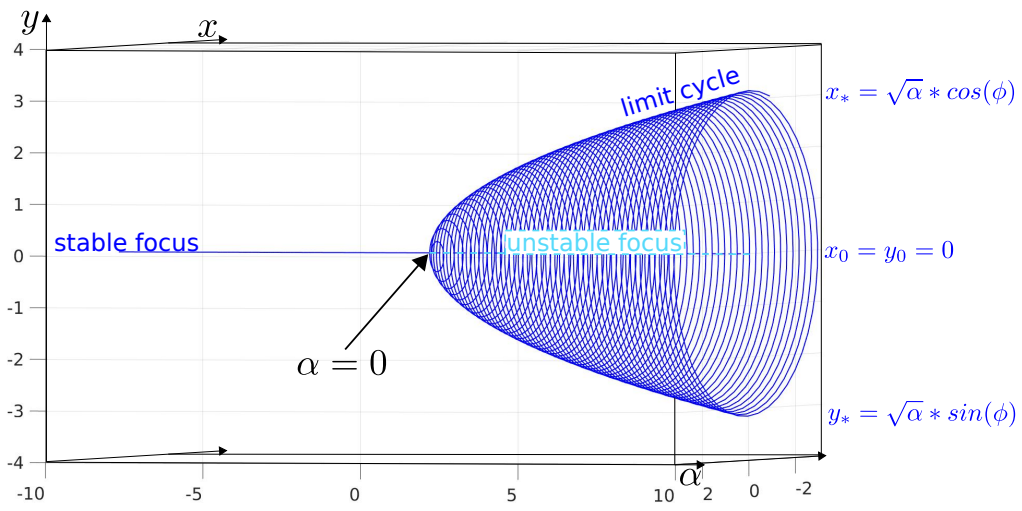

Figure 1.12: Hopf bifurcation diagram described by Eq. 1.40

The solution at r_* depends only on the radius; for $r > \sqrt{\alpha}$, $\dot{r} < 0$ and for $r < \sqrt{\alpha}$, $\dot{r} > 0$ and thus we have a stable limit cycle to which trajectories are pulled. For $\alpha < 0$, $r_* \in \mathbb{I}$, i.e. the trajectories surrounding the limit cycle spiral into a single fixed point r_0 and for $\alpha < 0$, $r_* \in \mathbb{R}$, i.e. the trajectories spiral away from the fixed point (see Fig. 1.13.)

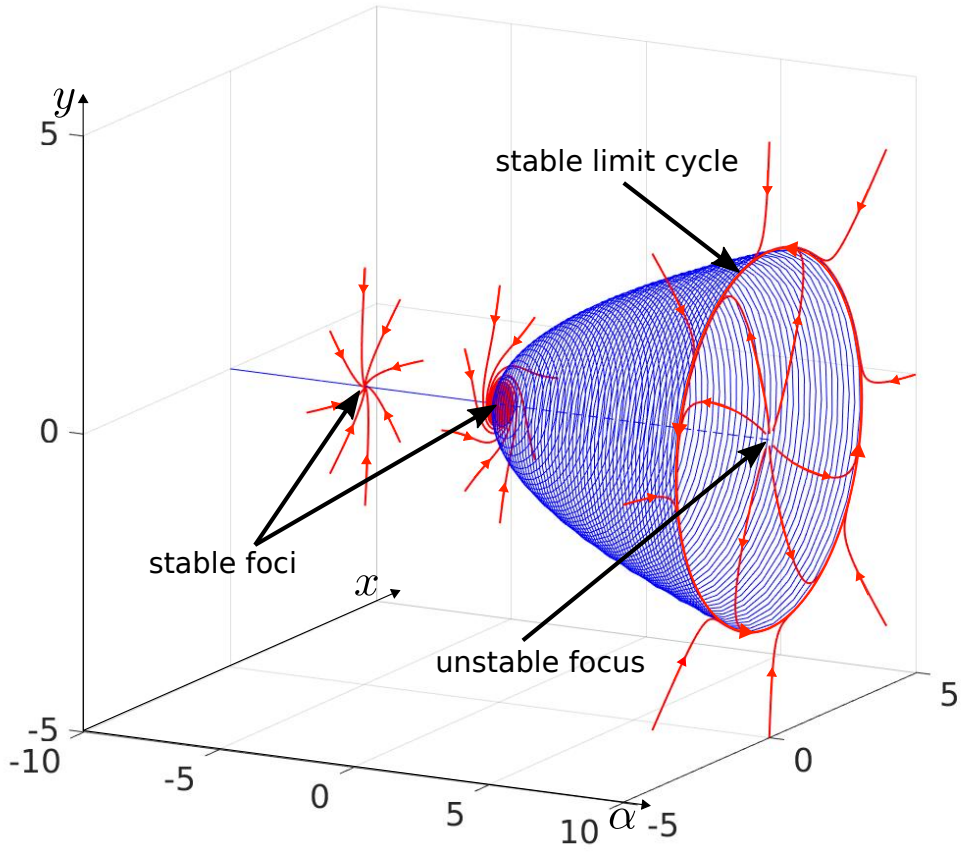

Figure 1.13: Hopf bifurcation and corresponding state portraits.

Chapter 2

Delay Differential Analysis

Many new techniques have emerged for analyzing experimental medical time series data, most of these rely on linear based algorithms. However, naturally occurring systems exhibit chaotic, deterministic behavior and thus, important aspects may be missed by linear analysis. Nonlinear methods are receiving more recognition, as they are able to detect structures in complex biological systems. One of the most common nonlinear approaches for revealing time series dynamics are delay and differential *embeddings*. An embedding converts a single time series into a multidimensional object in an embedding space revealing valuable information about the system, without having direct access to all the system's variables[33,66,83,77]. This concept is demonstrated in the next section by the embedding of the x variable of the set of first order differential equations (Eq. 1.24) of the Duffing oscillator (see Fig. 2.1). Delay diferential analysis (DDA) combines a derivative embedding with functional nonuniform delay embeddings to detect dynamical differences and nonstationarities in a given data set (Fig. 2.2). Functional, meaning that for the delay embeddings, polynomial models are used as an approximate of the system and nonuniform, meaning that the delays used in each of these approximations are independent of one another. A crucial difference to most time series analysis techniques, is that it is performed in the time domain, not in the spectral domain. DDA has been applied to the well known systems, the Lorenz system and the Rössler system, as a proof of concept[54,50].

DDA has shown successful applications in distinguishing, between heart conditions in electrocardiographic (ECG) recordings, between Parkinson electroencephalographic (EEG) and control EEG recordings, between cortical

states and epileptic events from high density electrocorticogram (ECoG) data, as well as in sifting out non linear features in information processing dysfunction in schizophrenics by revealing dynamical state changes[43,48,52,53]. In this thesis, DDA will be applied to single electrode EEG data to see if it is able to capture dynamical differences during various sleep stages as well as between patients with and without the sleep disorder obstructive sleep apnea.

2.1 Delay Differential Embedding

Embedding theory was already introduced in the first chapter. A general existence theorem for embeddings in an Euclidean space was given by Whitney (1936)[33]: A generic map from an n-dimensional manifold to a $2n+1$ dimensional Euclidean space is an embedding. Whitney's theorem implies that each state can be identified uniquely by a vector of $2n+1$ measurements, thereby reconstructing the state portrait. Takens embedding theorem (1981)[83] further states that this can also be achieved with $2n+1$ delay embeddings (time delayed versions) of one generic signal.

The concept of embeddings is made clear in Fig. 2.1 which depicts the differential, delay and delay-differential embeddings of the x variable of Eq. 1.24 of the Duffing oszillator. Some embeddings are only topologically equivalent to the original state portrait and caution must be taken concerning the reliability of the embedding. Due to the nature

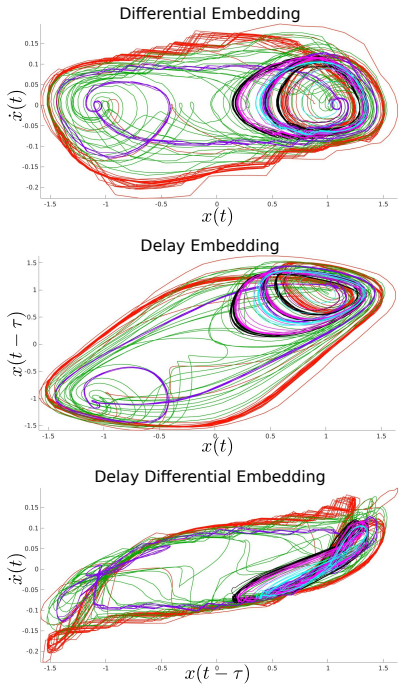

Figure 2.1: Embedding $x(t-\tau)$ with $\tau = 10$ of the variable $x(t)$ of the system given by Eq. 1.23 for parameters $c = 0.3$, $\beta = 1$, $\alpha = 1$, $\omega_E = 1.2$ and initial conditions $x_0 = 1$, $y_0 = 0$, $v_0 = 0$ and varying $u_0 = \gamma = [0.20, 0.28, 0.29, 0.37, 0.50, 0.65]$

of the variables in a nonlinear dynamical system being dependent of one another, each one will contain information about the rest of the system variables. An embedding represents a global diffeomorphism to the original system and thus we can obtain information about the dynamics of highly complex systems from mere single variable measurements. Fig. 2.2 depicts the relationship between DDA and embedding theory.

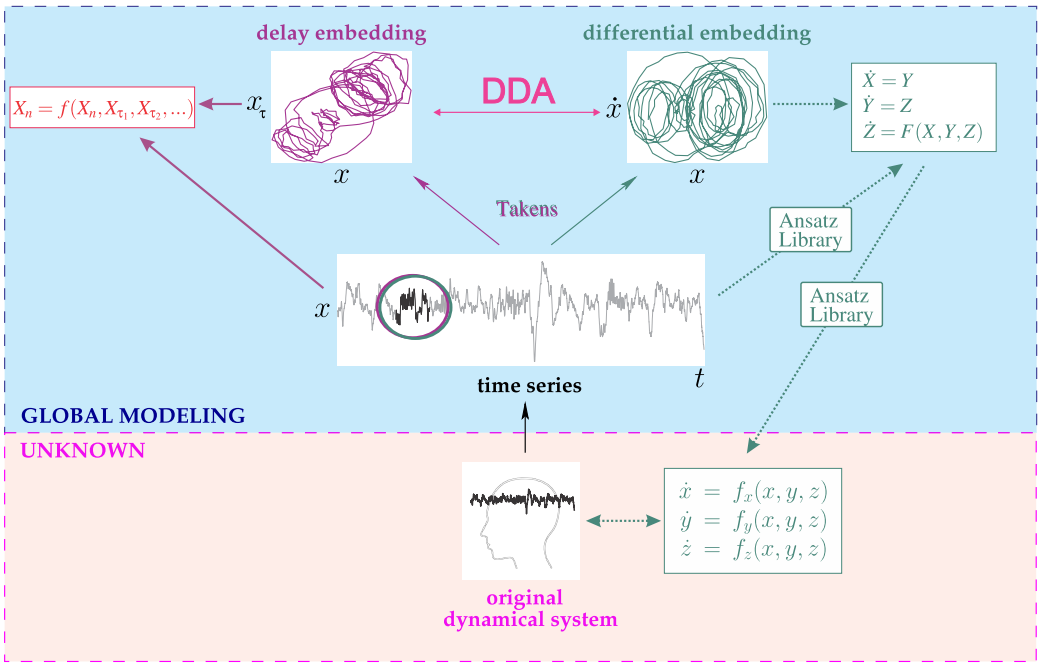

Figure 2.2: DDA and global modeling. By taking the time series of one variable $y(t)$ of a highly complex many dimensional system, and embedding this variable in various approximating models, we are able to capture valuable information about the dynamics of original system.

The *Ansatz Library* is a tool to reconstruct a general system from a noise free single measurement containing the original dynamical system as a subset[41,42,46,45]. The additional subsets include the original system in various coordinate representations as well as systems which are topologically equivalent to the original system. It shows that measurements from the same dynamical system on different timescales can be identified as belonging to the same dynamical system and can be seen as the theoretical limit for information contained in a single time measurement variable of a nonlinear

deterministic system. For real world noisy data, such global modeling is not possible and focus is put towards *detecting* and *quantifying* dynamical differences between data classes[44].

The general DDA model, a delay differential equation (DDE), is a non-uniform functional embedding of an unknown signal $x(t)$, which gives the derivative of the signal $x(t)$ in terms of the values of the function at previous times such that

$$\dot{x}(t) = \sum_{i=1}^{I} a_i \prod_{n=1}^{N} x(t)_{\tau_n}^{m_{n,i}} \tag{2.1}$$

where I is the number of monomials in the model, N is the number of delays and τ_n, $m_{n,i} \in \mathbb{N}_0$ with $x = x(t)$ and $x_{\tau_n} = x(t - \tau_n)$[44]. When applied to real data, we limit the number of terms on the right side of Eq. 2.1 and search for a low-dimensional DDE capable of detecting dynamical features in the data. For EEG time series data, the following DDA model has been proven to be successful in capturing underlying dynamical information and is used in the rest of this thesis[51,47,52,76]

$$\dot{x} = a_1 x_{\tau_1} + a_2 x_{\tau_2} + a_3 x_{\tau_1}^2 \tag{2.2}$$

with $x_{\tau_n} = x(t - \tau_n)$ and $\tau_1 \neq \tau_2$. The numerical derivative on the left hand side is done numerically using the center derivative[63]

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{2M} \sum_{n=1}^{M} \frac{x(t+m) - x(t-m)}{m}\,, \tag{2.3}$$

where M is the number of forward and backwards steps in computing the derivative. Data are noise contaminated, therefore, the dependence between error an M is not straightforward. In this thesis $M = 4$ was chosen.

For the analysis in this thesis, τ_1 and τ_2 were each set to $n = \{1, \ldots, 50\}\delta t$ where $1\delta t = \frac{1}{f_s}$ and f_s is the sampling rate, resulting in 2450 delay pairs (without the main diagonal) which is stored in a feature matrix $\hat{A}$ of dimension 50×50, and elements $A_{k,l} = (a_1, a_2, a_3, \rho)_{\tau_1,\tau_2}$.

The coefficients $\vec{\alpha} = a_i$ $(i = 1:3)$ as well as the least squares error ρ are used as features to distinguish dynamical differences in the time series data. This feature set reflects both linear and nonlinear properties of time signals and a connection between linear DDEs and tradiontal frequency analysis and non linear DDEs and higher order statistics has been made by Lainscsek and Sejnowski (2015)[44]. DDA is a nonlinear data anlysis framework that (1) uses

raw data, (2) uses sparse models that match the macroscopic architecture of the underlying dynamical system (3) disregards amplitude information to concentrate on the dynamical aspects of the data. Previously, it has been demonstrated that DDA captures essential features of data to produce exceptional classification performance.[49].

The coefficients $\vec{\alpha} = a_i$ are estimated numerically from the over determined system of equations using a singular value decomposition algorithm[72] as follows

$$\dot{\mathbf{x}} = \hat{M}\,\vec{\alpha}$$

$$\begin{pmatrix} \dot{x}(t) \\ \dot{x}(t+1) \\ \dot{x}(t+2) \\ \vdots \\ \dot{x}(t+L) \end{pmatrix} = \underbrace{\begin{pmatrix} x(t-\tau_1) & x(t-\tau_2) & (x(t-\tau_1))^2 \\ x(t+1-\tau_1) & x(t+1-\tau_2) & (x(t+1-\tau_1))^2 \\ x(t+2-\tau_1) & x(t+2-\tau_2) & (x(t+2-\tau_1))^2 \\ \vdots & & \\ x(t+L-\tau_1) & x(t+L-\tau_2) & (x(t+L-\tau_1))^2 \end{pmatrix}}_{\hat{M}} \underbrace{\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}}_{\vec{\alpha}} \tag{2.4}$$

where L is the number of time points used. Eq. 2.4 was solved using supervised singular value decomposition.

The errors ρ of Eq. 2.2 were calculated for each delay pair with mean-squared error estimation:

$$\rho = \frac{1}{L}\sqrt{\left(\dot{x} - \sum_{i=1}^{I} a_i \prod_{n=1}^{N} x(t)_{\tau_n}^{m_{n,i}}\right)^2} \tag{2.5}$$

Fig. 2.3 depicts this process. One needs merely the signal, its square and its numerical derivative. The signal is then delayed by τ_1 and τ_2 respectively and and a window length (WL) is chosen to compute the first set of features $A_{k,l} = (a_1, a_2, a_3, \rho)_{\tau_1,\tau_2}$ for this window. The process is then repeated by sliding over the data with a certain window shift (WS). The choice of WL depends mainly on the number of pseudoperiods in one data window in ST DDA.

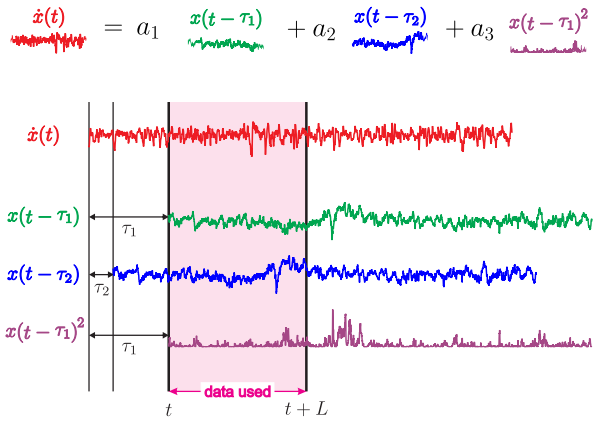

Figure 2.3: Single trial DDA: the features a_i (Eq. (2.2)) are estimated for each data window and time series.

DDA can be applied in a single trial (ST) or cross trial (CT) manner to time series which can be assumed to have similar overall dynamical content such as multiple trials of the same recording or EEG recordings of two channels. In ST-DDA, the features $A_{k,l} = (a_1, a_2, a_3, \rho)_{\tau_1,\tau_2}$ from Eq. (2.2) are estimated for each data window and time series separately as opposed to CT-DDA where they are estimated for multiple time series simultaneously (see Fig. 2.4). In CT DDA the WL can be reduced to increase the temporal resolution in the analysis. This implies that the number of rows in Eq. 2.4 will have n times more rows in the CT case

$$\begin{pmatrix} \dot{\mathbf{x}}_\mathbf{1} \\ \dot{\mathbf{x}}_\mathbf{2} \\ \vdots \\ \dot{\mathbf{x}}_\mathbf{n} \end{pmatrix} = \begin{pmatrix} \mathbf{M_1} \\ \mathbf{M_2} \\ \vdots \\ \mathbf{M_n} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \tag{2.6}$$

where n is the number of dynamically similar time series.

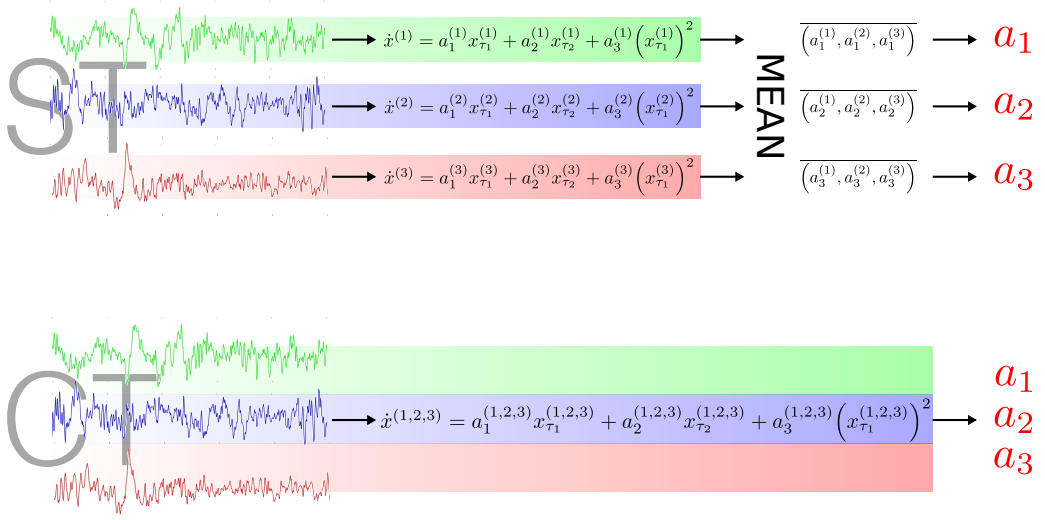

Figure 2.4: Comparison of single trial and cross trial DDA.

Boltzmann's ergodic hypothesis states, that in a sufficiently long measurement, the time average of a mesurand must equal to the ensemble average of the system[15]. CT-DDA uses this hypothesis as an Ansatz and combines multiple time series before computing the coefficient matrix arising from Eq. (2.2). Comparing ST and CT DDA outputs can thus be used as dynamical coherency test

$$\overline{ST} \stackrel{!}{=} CT \tag{2.7}$$

2.2 Repeated Random Subsampling k-fold Cross Validation

Cross-validation (CV) is a supervised learning algorithm which allows us to asses how generalizable results of a statistical analysis will be to an independent data set. The basic idea is to split the data into a (known) training set and a first seen testing set, either exhaustively, where all possible ways to divide the original sample are taken into account, or non-exhaustively.

The non-exhaustive methods include k-fold CV in which the data is randomly divided into k sets of roughly equal size. Here the data is split randomly into 3 subsets and training is carried out on two thirds of the data

and then then tested on the remaining third. The predictions of the model is then summarized into some type of performance measure, here the area under the receiver operating characteristic (ROC) curve. This process was repeated one hundred times in order to provide a better Markov estimate for the accuracy of the model[38].

Using this supervised structure selection method, we can determine which delay pairs give the features that best fit, or rather classify, the data. For each delay pair, the classes which are to be distinguished are split into two feature vectors containing the signals from subjects from each class. For the first delay pair

$$\boldsymbol{\tau}_1 = (\tau_1, \tau_2) = (1, 2)$$

which corresponds to the matrix element $A_{1,2} = (a_1, a_2, a_3, \rho)_{\tau_1,\tau_2}$, we get two feature matrices $\hat{A}_1$ and $\hat{A}_2$

$$\hat{A}_1 = \begin{pmatrix} A^1_{1,2} \\ A^2_{1,2} \\ \vdots \\ A^s_{1,2} \end{pmatrix} \; ; \quad \hat{A}_2 = \begin{pmatrix} A^1_{1,2} \\ A^2_{1,2} \\ \vdots \\ A^s_{1,2} \end{pmatrix} \tag{2.8}$$

with $s = 1$:(number of subjects). The feature vectors $\hat{A}_1$ and $\hat{A}_2$ are then randomly split into a training and testing group (2 : 1 ratio) $\hat{A}_1^{train}$, $\hat{A}_1^{test}$ and $\hat{A}_2^{train}$, $\hat{A}_2^{test}$. Two label vectors **L1** and **L2** are created in the same length as the training vectors where one class is labeled with zeros and the other with ones. In this way, the distance from the hyperplane will be measured at 0.5 (see Fig. 2.5).

The weight vector $\mathbf{W}$ is computed using the training feature vectors as well as an a fore placed $\mathbf{1}$ vector which prevents the separating hyperplane from passing through the origin

$$\mathbf{W} = \begin{bmatrix} \mathbf{1} & \hat{A}_1^{train} \\ \mathbf{1} & \hat{A}_2^{train} \end{bmatrix} \setminus \mathbf{L} \tag{2.9}$$

The last step is to test the weights on the testing vectors in order to produce the 1-dim parameter $\mathbf{D}$ from a 4-dim hyperplane spanned by the feature space

$$\mathbf{D} = \mathbf{W} \cdot \begin{bmatrix} \mathbf{1} & \hat{A}_1^{train} \\ \mathbf{1} & \hat{A}_2^{train} \end{bmatrix} \tag{2.10}$$

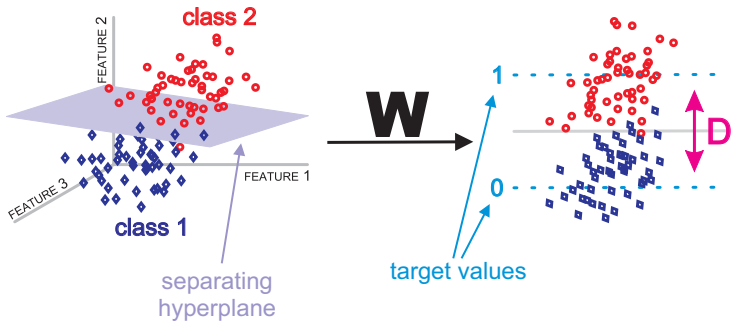

Figure 2.5: Construction of the weight matrix: The classes 1 and 2 correspond to the training matrices R_1 and R_2. The weights W then find the best separating hyperplane (best separation of the 2 classes) that maps the feature points of the 4-D feature space (constant, a_i, a_j, and error F) to a distance from the hyperplane D. All $D < 0.5$ correspond then to class 1 (R_1) and all $D \geq 0.5$ correspond then to class 2 (R_2)[54].

In a perfect system, the distances from the hyperplane D of the first class would be > 0.5 and those of the second class would be < 0.5.

Fig. 2.6 shows the effect of the additional **1** vector in Eq. 2.9 where the first row is the calculation without the **1** vector and the second row with it.

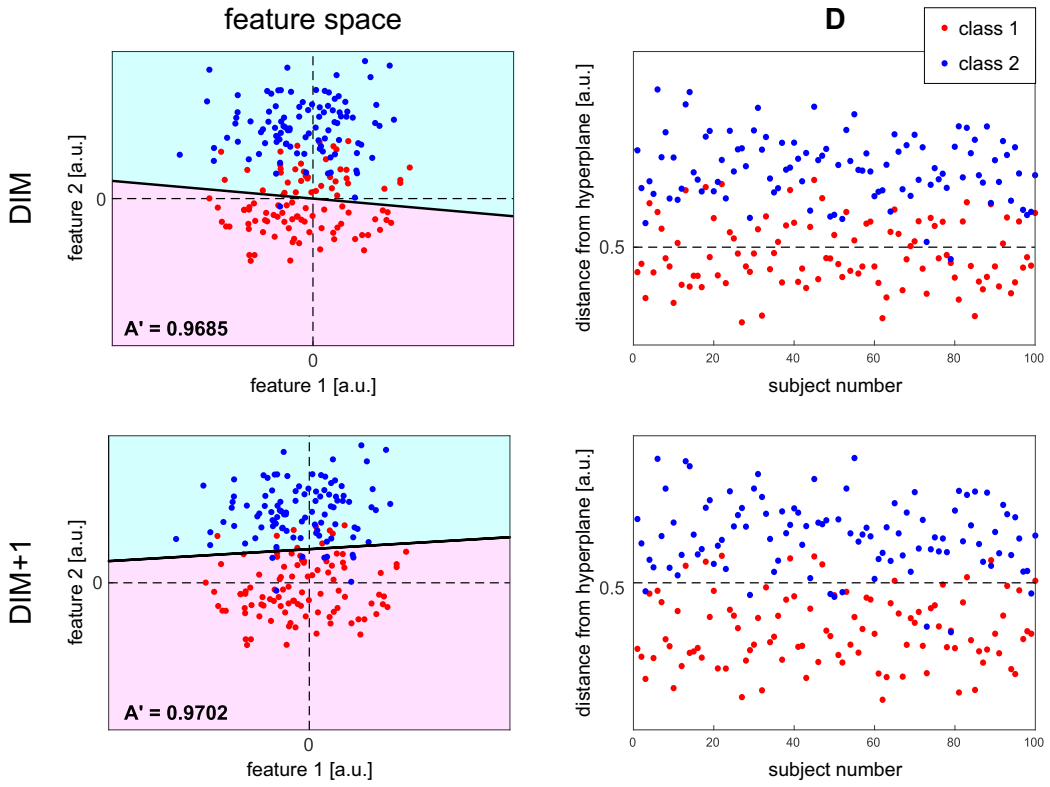

Figure 2.6: Impact of the additional **1** vector in the calculation of **D** in Eq. 2.9. The top row corresponds to the calculation without and the the bottom row with the additional **1** vector.

2.2.1 Receiver Operating Characteristic Curve

The area under receiver operating characteristic (ROC) curve, is used to test diagnostic performance, or to test the ability to correctly classify subjects into clinically relevant subgroups[62]. Fig. 2.7 shows the construction of the ROC curve for two arbitrary, not 100% distinguishable, classes (red and blue). The true positive rate (TPR) is the measure of the proportion of actual positives that are correctly identified as such (eg. the proportion of sleep apnea patients which have been correctly classified as having apnea) and the false positive rate (FPR) is the proportion of actual negatives that are correctly identified as such (eg. the proportion of non sleep apnea patients which have been correctly classified as healthy). The ROC curve is then derived by plotting the TPR against the FPR for each threshold (cyan dashed lines). The area under this curve A' is then used as a separation measure to compare the DDA model's with varying delay pairs. A' ranges from 0 to 1, where 1 is perfect separation and 0.5 means no separation of the 2 classes. An A' of zero means perfect separation of the two classes with inverted labels.[54]

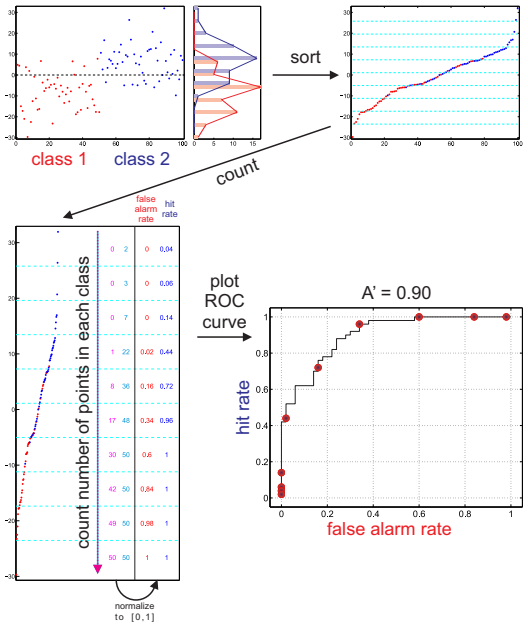

Figure 2.7: ROC curve for two arbitrary classes red and blue[54].

Chapter 3

Automatic Sleep Scoring

3.1 The Sleep Cycle

When we talk about leading a healthy lifestyle, focus is laid on two things, namely diet and exercise. Sleep is equally important in maintaining a healthy brain and body. Sleep deprivation can have severe effects on cognitive performance such as impairing attention and working memory as well as long-term memory, decision making and vigilance[4,58]. Further physiologic consequences include adverse affects on endocrine functions and immune responses[6]. The length of sleep needed varies immensely between individuals but is on average 7 to 8.5 hours per day[19] and we cycle through states of wakefulness, rapid-eye-movement and non-rapid-eye-movement sleep. The so called Rechtsschaffen and Kales (R & K) sleep scoring manual of 1968[73] divides sleep into five stages: S1 S2 S3 S4 and REM sleep. Since 2007, the American Academy of Sleep Medicine updated this scoring manual combining S3 and S4 into a single slow wave sleep stage leaving the stages: N1 N2 N3 and REM. This heuristic discretization of sleep was devised as a standardized scoring system to allow uniformity between sleep laboratories and should rather be viewed as a gradual transition of a waveform[84].

EEG activity is divided into four bands based on the frequency and amplitude of the waveform whos frequencies depend on the reference used. Conventionally there are four bands

1. beta band: $> 13Hz$
2. alpha band: $> 8Hz \ \& < 13Hz$

3. theta band: $> 4Hz$ & $< 8Hz$

4. delta band: $< 4Hz$

Preceding a nights sleep is typically several minutes of wakefulness (W). The EEG will show mixed beta and slpha activities and W is recorded when $>$ 50% of a 30 second epoch shows scorable alpha activity. As drowsiness kicks in and the eyes are closed, the EEG will show predominant alpha activity and the EMG activity will reduce. Typically, sleep stage N1 follows W, and often a person will reawaken after having spent only a few epochs in N1. These transition states are often difficult to score[84].

3.1.1 N1 NREM Sleep (R & K S1)

Sleep stage N1 or light sleep is a transition sleep stage characterized by low-voltage and fast EEG activity. It is defined as an epoch consisting of $> 50\%$ theta activity whilst the alpha activity drops to $< 50\%$. The patterns can be quite fluctuating and thus difficult to interpret. During drowsiness and N1 sleep, the eyes begin to slowly roll (SREM), breathing becomes shallow, the heart rate becomes regular, blood pressure falls and there is little or no body movement. Thoughts will begin to drift, the person may experience sensations of floating and one is easily awakened from N1 sleep stage. In general the time spent in N1 increases with age.[84]

3.1.2 N2 NREM Sleep (R & K S2)

Sleep stage N2 are characterized by the so called K-complexes and sleep spindles[12]. Sleep spindles are burts of brain activity, around $11 - 16$ Hz generated in thalamic circuits as a consequence of cortical firing and typically last between $0.3 - 3$ seconds and reoccurring every $5 - 15$ seconds[16]. They get their name from the fact that their EEG pattern resemble yarn spindles. In the recent years, attention has been put on researching sleep spindles and their potential roles in memory consolidation and other cognitive functions as well as in psychiatric and neurological disorders[79,78,29,26,69,39]. Excessive sleep spindles may also be due to medications such as benzodiazepines[84]. DDA has shown to be very successful in automatic sleep spindle detection as the second fastest, yet providing the highest agreement with human expert scoring[76]. K-complexes are characterized by a sharp negative EEG wave deflection, followed by a slower positive component. They arise in the central vertex

and must persist for at least 0.5 seconds. They may occur with or without external stimuli which indicates that the brain is still minimally responsive that these k-complexes could represent a form of cortical evoked potential. Sleep spindles often immediately follow k-complexes. Sleep stage N2 accounts for the largest amount of time spent in any sleep stage especially in adults. Next to the bursts of activity, is characterized by predominant theta activity and delta activity is only allowed to occur for less than 20% of each epoch; this is the threshold for slow wave sleep. There is a significant decrease in physiological bodily functions such as blood pressure, brain metabolism, gastrointestinal secretions and cardiac activity. A person will feel farther and farther away from a consiouss state and will be more difficult to arouse.[84]

3.1.3 N3 NREM Sleep (R & K S3 and S4)

The N3 sleep stage, or slow wave sleep (SWS), includes the old R & K stages S3 and S4 partly to overcome interrater variability between scorers[7,21] and partly because experts believe such a distinction to have no clear clinical significance[84]. N3 is characterized by high amplitude low frequency waves and both sleep spindles as well as k-complaxes may occur. There exist no clear criteria for EOG and EMG, but generally muscle tone is further decreased. SWS is thought to be play an key factor in cerebral restoration and recovery, however its exact nature and role is still not clearly understood[35,84,75]. Dreaming has been found to be associated with local decreases in slow wave activity (SWA) in posterior brain regions and many parasomnias such as sleep terrors and sleep walking may manifest themselves during SWS[80,84]. The N3 sleep stage has also been linked to a peak in growth hormone secretion. Arousal from N3 is far more difficult and waking up from this stage may be accompanied by confusion and disorientation. The amount of time spent in SWS typically decreases during the night to avoid sleep inertia in the morning.[84]

3.1.4 REM Sleep

Rapid eye movement sleep first occurs about 90 to 120 minutes after sleep onset in adults. About $20 - 25\%$ of sleep is spent in REM and subsequent REM periods typically become progressively longer and more robust. It is characterized by low-amplitude irregular pattern EEGs of mixed frequencies of the theta and alpha bands. As its name already implies, pronounced

bursts of eye activity are visible in the EOGs. Physiological activity tends to be higher than in the NREM sleep stages and some very interesting effects occur. Blood pressure and pulse rate may dramatically increase or fluctuate. Respiratory functions are affected so that breathing becomes irregular and the brains consumption of O_2 increases. Both men and women experience episodes of vascular engorgement of the genitalia. More than 90% of REM periods in the male are associated with full or partial penile erection. Women experience the same frequency of episodes, however differ in distribution, in greater frequency during NREM sleep, in duration, and in their less tight correspondence to REM sleep phases.[84,28]

Infants spend a majority of their early lives in a REM like sleep state similar to the REM state of adults (the differences lie in the cortical EEGs of adult and infant REM sleep) implying that this sleep stage is important for brain development. This reduction of time spent in REM with age, arises the question as to if REM sleep serves the same functions across a life span[14,60]. There is research that suggests that the muscle twitches occurring during REM may function to aid sensori motor system development. Multiple brain regions which are not active during the wake state are activated by muscle twitches during REM sleep.[67]

REM sleep is regulated by an intricate network of brain circuits and a countless number of neuromodulators. Hippocampal neural activity during REM sleep is involved in memory consolidation, however its exact functional role remains disputed. The big question as to what exact biological an neuropsychiatric function dreams have for us is still a mystery.[67,17]

3.2 ZOMA Health Inc.

This thesis was done in cooperation with the start up company ZOMA Health Inc., situated in California, whos goal is to implement DDA to quickly and accurately assess EEG sleep data. There are commercially available devices that claim to classify sleep stages according to heart rate and accelerometer data, however there accuracy is at the most decent and sleep stages are reduced to 4 stages namely: deep, light, REM and awake. Studies comparing such wearable wristband sleep trackers amongst each other as well as with research grade devices typically show high sensitivity (ability to detect sleep) and low specificity (ability to detect wake)[55,59,22].

A home monitoring device which utilizes single channel EEG would greatly

improve reliability of sleep tracking. ZOMA's vision is to create such a device in cooperation with the group of Dr. Dott. Greco at the TU Graz working on long lasting wearable tattoo electrodes[27]. With the help of DDA, ZOMA hopes to offer hospital grade sleep scoring from the comfort of a patients home. This would enable people suffering from sleep conditions to continuously monitor their sleep each night to track treatment effects at a fraction of the cost and time of a traditional PSG. Monitoring sleep using only single electrode EEG would get rid of the elaborate amount of recording equipment required for traditional PSG. Furthermore, a sleep lab environment often affects the normal sleep patterns of patients and testing sleep from home would give more reliable incite into an individuals typical sleep architecture and how lifestyle choices affect it. The data will be sent to a smartphone app which is able to create a hypnogram within a matter of seconds greatly surpassing the computational time that other traditional machine learning techniques require. In contrast to such algorithms, DDA is an approach based on physical fundamentals, namely nonlinear dynamics. This reduces the variability between sleep scoring among people with various factors such as age, sex, race, ethnicity, fitness and possible existing conditions.

A long term goal would be to create a wearable device to be able to also continuously track other conditions such as epilepsy.

3.3 Polysomnography

A polysomnography various physiological changes that occur during sleep including brain activity (EEG), eye movements (EOG), muscle activity (EMG), heart rhythm (ECG), breathing functions, respiratory airflow, respiratory effort indicators and peripheral pulse oximetry, in order to determine the amount of time spent in each of the six sleep stages and to make diagnostic evaluations about various sleep illnesses. Polysomnography is best performed by certified sleep technicians in sleep laboratories, however often inexperienced and persons who lack knowledge in this area are allowed to perform the polysomnography.

Trained sleep specialists then visually assess 30sec epochs of data at a time into wake or rapid eye movement (REM) and non rapid eye movement (NREM: S1, S2, S3, S4) sleep stages according to the Rechtsschaffen and Kales (R & K) rules[73]. These scores are then depicted into whats known as a hypnogram. An example is given in Fig. 3.1.

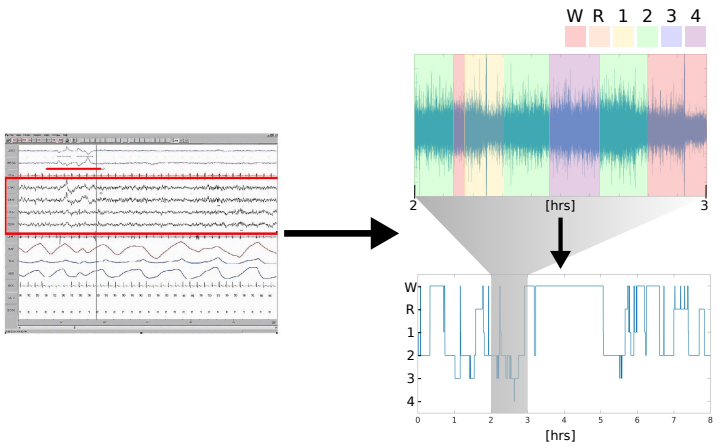

Figure 3.1: Human scored hypnogram.

3.4 DDA for automatic sleep scoring

Until 2007, the Rechtsschaffen and Kales sleep scoring manual from 1968 was used which comprises the 6 sleep stages described above[73]. Since then, the American Academy of Sleep Medicine updated these scoring rules to overcome high inter-rater variability[12]. Sleep stages S3 and S4 are notoriously difficult to distinguish and so they were combined to a single slow wave sleep stage N3. Sleep scoring according to the R & K as well as the ASSM rules have been tested for inter-rater agreement between scorers from various sleep labs and despite the merging of S3 and S4 to N3, the inter-rater agreement was only slightly over 72%[7,21]. A discrete scaling of sleep stages ignores the actual gradual transitioning of wave forms and therefore, an automatic sleep stage classification algorithm would benefit in overall reliability as well as substantially reduce the time and cost of traditional polysomnography. Furthermore, achieving a hypnogram via merely one EEG electrode could enable commercially available home continuous sleep monitoring devices like the one described in Sec. 3.2.

3.4.1 DDA outputs as marker for sleep stage classification

During sleep, brain activity changes during various sleep stages which can be seen as the parameter fluctuations of a dynamical system. In 2014 Lainscsek

et al.[51] showed that DDA features are sensitive to these state changes with a correlation with manual scoring of $\overline{r} = (0.86 \pm 0.1)$[51]. An genetic algorithm was implemented to search for the DDE model best able to classify the data in 30sec epochs. This was performed on 35 single electrode EEG recordings (C3/A2) that were recorded at the sleep laboratory at the medical university hospital Intensive Care Unit in Rouen. The resulting model was

$$\dot{x} = a_1 x_{\tau_1} + a_2 x_{\tau_2} + a_3 x_{\tau_1}^2 x_{\tau_2} \tag{3.1}$$

with $\tau_1 = 1$ and $\tau_2 = 3$ (see Fig. 1 in[51]). The coefficient a_2 was found to have the best correlation with the manually scored hypnogram. Since a_2 has arbitrary units and classical hypnograms scale with $6:1$, the coefficient had to be scaled accordingly.

Here, the DDA will be compared and retested to the data sets coming from healthy individuals of the Physionet database and sleep apnea patients from the SHHS dataset. It is to be noted that the algorithm used to score the EEG patterns is a slightly modified proprietary version of the above, however it uses the same basic idea of using the DDA features to classify the various sleep stages. There were minimal differences between the performances of the models Eq. 3.1 and Eq. 2.2. Any model with two nonlinear and one linear term has been proven useful for detecting dynamics in EEG data.

3.5 Datasets

The EEG data used in this thesis were taken from the public databases *Physionet* as well as *The Sleep Heart Health Study (SHHS)* for which ZOMA Inc. received permission to use. The Physionet database contains 197 whole night polysomnographies as well as manually scored hypnograms according to the R & K manual[73] based on Fpz-Cz/Pz-Oz (see Fig. 3.2) EEGs. This data set was comprised of 153 recordings from 1987-1991 of healthy Caucasians from the ages of 25-101. The recordings were performed in the subjects homes and they were taking no sleep related medication. Sampling rates of the EEG signals were 100Hz. Of these 147 subjects were analyzed.

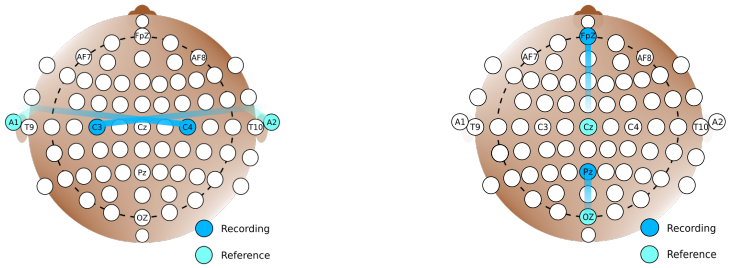

Figure 3.2: Electrode placement of SHHS (left) and Physionet (right) data sets.

The SHHS data set was clinically collected from 6441 individuals between 1995 and 1998 and 2001 by the National Lung and Blood Institute to determine the role sleep-disordered breathing plays in the risk of coronary heart disease, stroke, all cause mortality, and hypertension. The inclusion criteria was met by subjects who were aged 40 or older, no previous history of treatment of sleep apnea, no tracheostomy and no current home oxygen therapy. The EEG recordings were taken at C3/A2 and C4/A1 (see Fig. 3.2 and sampled at 125Hz. In this thesis, 5792 of the 1449 subjects were considered for automatic sleep scoring. Patients who met only had five sleep stages instead of six were excluded to ensure that all had been scored by the R & K rules.

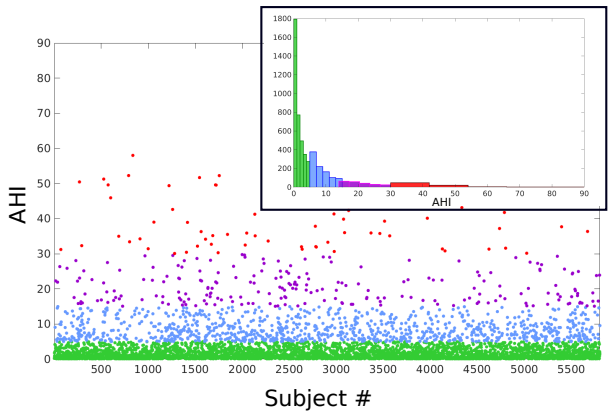

Figure 3.3: SHHS dataset of OSA subjects according to their AHI.

3.5.1 MUSE EEG Headband

In addition to these professional grade data sets, the commercially available EEG headband from MUSE^{TM} was tried to be used as a preliminary source of home EEG recordings as a proof of concept. This device has a sampling rate of 256Hz and the conducting material is located on the forehead region as well as on two rubber strips behind the ear. The electrodes used are AF7/FpZ, AF8/FpZ, T9/FpZ and T10/FpZ (see Fig. 3.4).

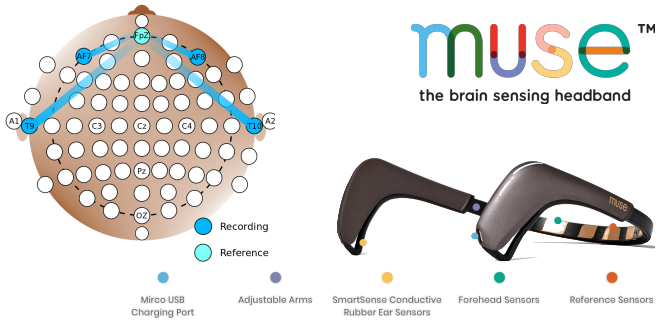

Figure 3.4: Electrode placement of MUSE data.

It is meant as a mediation assistant and unsuitable for measuring EEG while sleeping as it is very sensitive to movement. Keeping the device in place was nearly impossible even during an upright nap and a proper looking hypnogram could not be obtained using DDA on the recorded data. Besides the fact that the data was of extremely poor quality (see Fig. 3.5), the DDE model used for analysis was trained on the C3/A2 electrode. The model search may have to be repeated for data coming from electrodes of different placements.

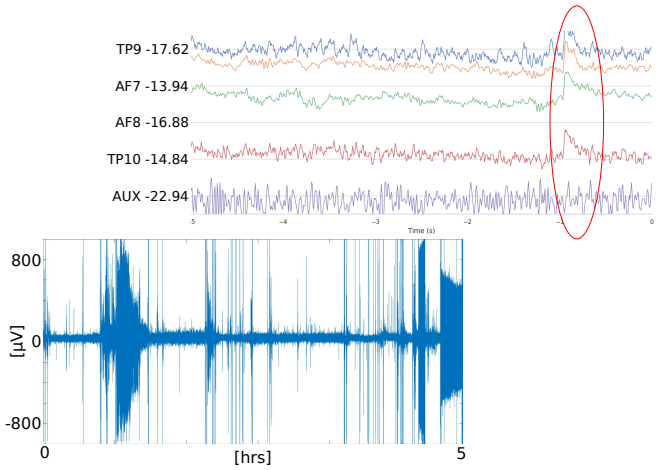

Figure 3.5: Sample EEG data from MUSE headband.

3.6 Results

As a classification performance parameter, the correlation coefficients between the DDA outputs (D) and the manually scored hypnogram (H) were calculated using the MATLAB function $CC = corrcoef(H, D)$. Further statistics will be required in order to reliably compare with literature. Nevertheless, most alternative techniques require more of the available data from the PSG than merely the EEG signals.

3.6.1 SHHS

Fig. 3.6 (left) depicts a sample nights hypnogram who correlated very well ($CC = 0.90$) with the DDA features. DDA seems to be able to capture the gradual transitioning from one sleep stage to another very well. In all, 1449 recording of the data set were taken into account and the results are given in a histogram in Fig. 3.6 (right). The mean correlation was 0.72 which is significant for preliminary results and when considering the fact that inter-rater agreement between sleep scorers in Basner et al. [7].

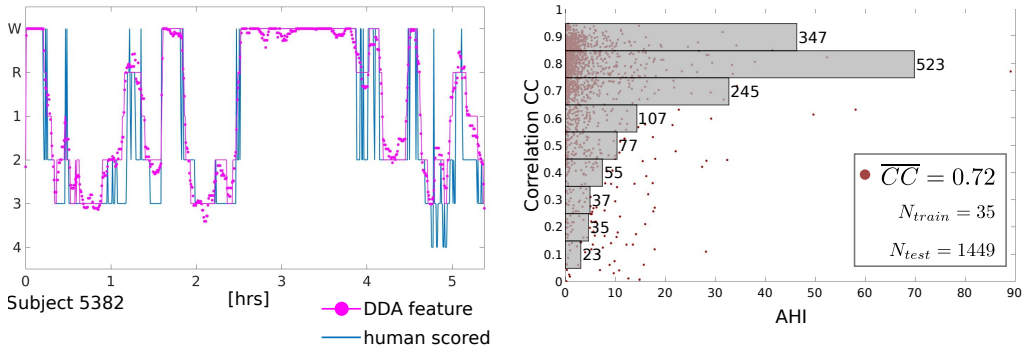

Figure 3.6: Sample DDA features over human scored hypnogram (left) and histogram of correlations of DDA outputs to manually scored polysomnography (right).

Concerning the patients who had very low correlations (about $N(CC \leq .5) \approx 15\%$ of the entire data), further investigation is required as to why the analysis performed poorly. This could be the result of various factors such as faulty connection between scalp and electrode, scoring patients with conditions can be challenging for some sleep technicians and furthermore, the way the information of the DDA features was extracted, is just one of many possible approaches and will be subject of further investigation. Fig. 3.7 depicts three sample subjects who had $CC < 0.20$.

Figure 3.7: Three sample human scored hypnograms (left) and corresponding DDA features (right) for subjects who scored a correlation of $CC < 0.20$.

Karimzadeh et al. selected 140 patients form the SHHS data bank and performed a distributed classification procedure to automatically detect sleep stages according to the AASM annotations (W, N1, N2, N3, REM) . To enhance EEG signal quality a 8th order Butterworth filter was used for preprocessing. They used frequency band entropies of the EEG phase and envelope for each 30 second epoch as features. Then the 140 subjects were divided into training and validation sets and a decision tree was estimated. Accuracy, specificity and sensitivity were used as performance measures and achieved $\overline{ACC} = 83.17\%$, $\overline{SP} = 86.99\%$ and $\overline{SE} = 69.35$ respectively. An overall accuracy according to Fig. 7 in was only 0.73.

Biswal et al.[13] achieved an overall accuracy of 78% using deep neural networks when training on 9000 PSGs from the Massachusetts General Hospital Sleep Laboratory and testing on the SHHS data set. They combined deep recurrent and convolutional neural networks (RCNN) for supervised learning of the AASM sleep stage labels. This approach involved no preprocessing of data and they claim that their method is robust to physiologic variability between patients because the algorithm was trained on 8 yeas of clinical sleep data (MGH).

3.6.2 Physionet

The physionet data base includes two sets of recordings: CH1 Fpz/Cz and CH2 Pz/Oz (see Fig. 3.2). The resulting correlations score on average $\overline{CC}_{CH1} = 0.60$ for CH1 and $\overline{CC}_{CH2} = 0.73$ for CH2. Due to the fact that the DDE model used was trained on the 35 patient recordings from the C3/A2 electrode[51], it is not surprising that there is a large performance difference between CH1 and CH2 in the physionet data. The CH2 electrode Pz/Oz is topographically much nearer to the training electrode C3/A2.

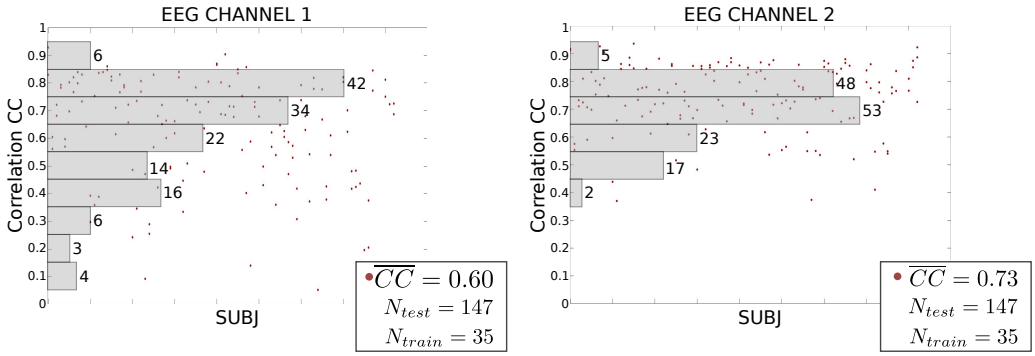

Figure 3.8: Histograms of correlations of DDA outputs to manually scored polysomnography for CH1 (Fpz/Cz) and CH2 (Pz/Oz).

Imtiaz and Villegas'[36] approch involved a combination of state machines and decision trees to create an automatic sleep scoring algorithm using only EEG data. They acheived an overall accuracy of 82% and 79% accuracy for training and testing respectively. Prakash and Roy[65] extracted features with the help of Ensemble Empirical Mode Decmposition, Hjorth parameter and zero-crossing rate and achieved an accuracy of $\approx 92\%$ when classifying into 6 sleep stages.

Other attempts often rely on more than just the EEG signal (such as EMG and ECG data) to extract features for automatic sleep detection[25,24,87]. DDA and the method presented in[13] are unique to other algorithms in that they do not require preprocessing of the time signals. In addition, the computational effort of DDA remains extremely low compared with other techniques and the results presented in this thesis are promising.

Comparing with other literature, there is much room for improvement with the proposed method of using DDA features. Since the training set used,

arose from data coming from topographically different electrodes, retraining the DDA features would be a suitable first step.

Chapter 4

Sleep Apnea Classification

Sleep apnea is a condition in which a persons breathing ceases during sleep which can have devastating consequences for those affected. We distinguish between central sleep apnea (CSA) and obstructive sleep apnea (OSA). CSA is a condition in which the brain fails to send signals to breath during sleep. Escessive daytime sleepiness frequently results from the fragmentation of sleep due to OSA. Sleep apnea can also be cause for cognitive impairment and psychological problems like anxiety and depression[20]. The focus in this thesis is put on the OSA, although it is important to note that investigation of CSA in the same way could have promising results.

4.1 Apnea Hypopnea Index

Hypopnea is defined as periods of abnormal or shallow breathing while apnea means a full blockage of breathing of at least 10 seconds or more. The apnea-hypopnea-index (AHI) is the combined average number of apneas and hypopneas that occur per hours of sleep. The American Academy of Sleep Medicine (AASM) categorizes the severity of sleep apnea into the following[12]:

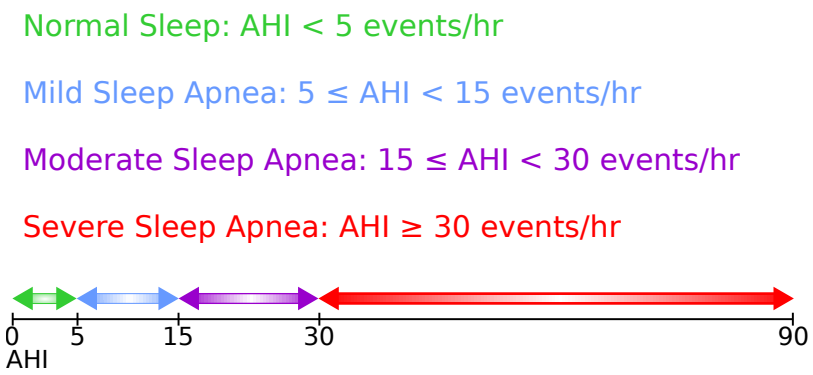

Figure 4.1: Apnea Hypopnea Index

This classification of apnea has its drawbacks, as it does not consider the time spent in each apnea-hypopnea event, which can be the cause for more severe symptoms and conditions. The gold standard for OSA diagnoses and severity classification to this day remains to be attended polysomnography. This method is expensive and time consuming and an automatic standardized way of detecting and classifying OSA using only one EEG electrode, could enable patients a continuous monitoring of this disease from home. In this way, tracking progression as well as treatment responses could be closely monitered. Furthermore, some patients show considerable night-to-night variability in the AHI, so treatments need to be personalized[81].

4.2 Obstructive Sleep Apnea

OSA has been found to be highly prevalent in the population, however there are discrepensies between the results of epidemiological studies about the prevalence of OSA most likely due to various diagnostic techniques, definitions for hypopnea scoring, study design and characteristics of included subjects[30]. Additionally, the abnormality in a patients sleep cycle patterns varies with severity[10]. Scoring abnormal respiration remains to be a controversial challenge in the field of sleep medicine and automated EEG based computer methods could be key in understanding this class of conditions as well as probing new treatment options for patients.

The gold standard for OSA treatment is a continous airway positive pressure (CPAP) ventillation machine which administers mild air pressure to keep

airways open. Apart from the negative psychological affect CPAP machines have on many patients, long term use can cause upper air symptoms such as nasal congestion, dry nose, sore throat, rhinorrhea, sneezing and a bleeding nose[68].

Patients with mild to moderate AHI are more likely to have REM related OSA. There are controversial findings between the relationship of AHI-REM and excessive daytime sleepiness. Most severe hypopnea and apnea events typically occur during REM sleep in the second part of the night. Respiratory events during REM sleep tend to be of longer duration and associated with more significant arterial oxygen desaturation[40,11].

REM sleep is composed of a phasic and tonic component. The phasic component is a state driven by the sympathetic nervous system and characterized by rapid eye movements, muscle twitches, and respiratory variability. The tonic REM component is a driven by the parasympathetic nervous system with no eye movements in which the body is essentially in a paralytic state. The eye movements bursts of phasic REM are markers of brainstem activity affecting respiration. Periods of REM sleep are longest and more frequent in the morning hours, so it is of no surprise that patients with OSA and other abnormal respiratory illnesses experience the greatest changes in ventilation[12,1,5]. Patients suffering from OSA are more susceptible to phasic changes in REM sleep and the muscles of respiration other than the diaphragm are less active due to the generalized state of hypotonia. However, periodic decrements in diaphragmic activity also do occur during bursts of eye movement.[8]

Robert J. Thomas et. al[85] bring up some important considerations about the scoring of sleep apnea events some of which are listed below:

- The smallest degree of oxygen desaturation which can be visually detected is $2\% - 3\%$. In clinical practice, nondesaturating hypopneas have clinical consequences and should thus be taken into account.

- Sleep cycle patterns modulate respiration and there are periods of stability intrinsic to NREM sleep which can determine the presence of apnea.

- The AASM scoring of respiratory events includes the 3-second arousal rule[12] which has no biological basis. Alpha and beta intrusions, the K-complex and delta-bursts of less than 3 seconds may be considered

arousal events. Thus computerized arousal detection via EEG could differentiate between significant vs. non significant respiratory events.

- Obstructive apneas should always be scored, however nonobstructive apneas (not leading to an arousal or oxygen desaturation, especially during REM), isolated postarousal apneas and apneas during sleep wake transition periods should not be scored.

- There is also no biological basis for scoring only apneas which last for 10 seconds $\geq$ 10sec.

These items as well as a few other points made in their book, give evidence to support the fact that the AHI is still an insufficient parameter of the severity of sleep apnea.

In this thesis, only the REM EEG segments were considered of the SHHS data set. This 1) greatly reduced the amount of data when performing DDE model search, 2) could give a small insight into the clinical dispute among sleep experts and research between OSA events during REM sleep and 3) reduced variability in the data as the interrater agreement for sleep stage REM is the hightest and thus the most reliable among the sleep stages[7].

4.3 Results

For this analysis, the SHHS data base described in Sec. 3.6.1 was used. Only REM EEG segments of patients scored after the 1968 published Rechtschaffen and Kales manual[73] were considered to reduce variability between sleep scoring. The REM peaks of each subject were pinpointed using the manually scored hypnogram, any subsequent periods of REM which were within 3 minutes of each other were combined. This is depicted in Fig. 4.2 for one subject.

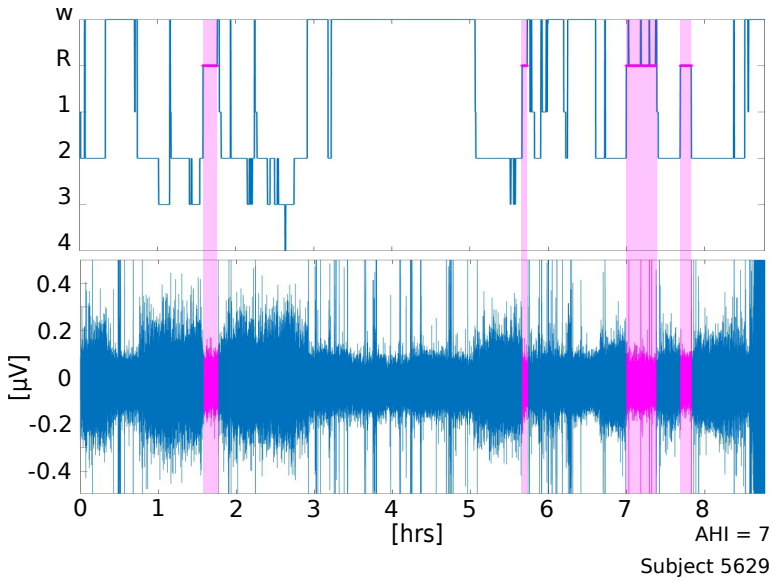

Figure 4.2: Visualization of pinpointing REM EEG segments for a sample subject.

The analysis applying DDA to these EEG time series signals described in Ch. 2 was carried out both single trial and cross trial (Fig. 2.4) for a window length of $WL = 1$ min and a window shift of $WS = 30$ sec to match typical human scoring of 30 sec epochs. The mean, standard deviation and median was taken of the resulting DDA features over the windows for each of the 2450 delay pairs. This leaves following feature matrices:

$$dim(\hat{A}_{CT}) = 12 \times 2450$$
$$dim(\hat{A}_{ST}) = 12 \times 6 \times 2450$$

The goal was now to see if the DDA features could classify the subjects categorized according to their apnea-hypopnea-index (AHI) (see Fig. 4.1). Additionally one marker was taken at $AHI = 10$.

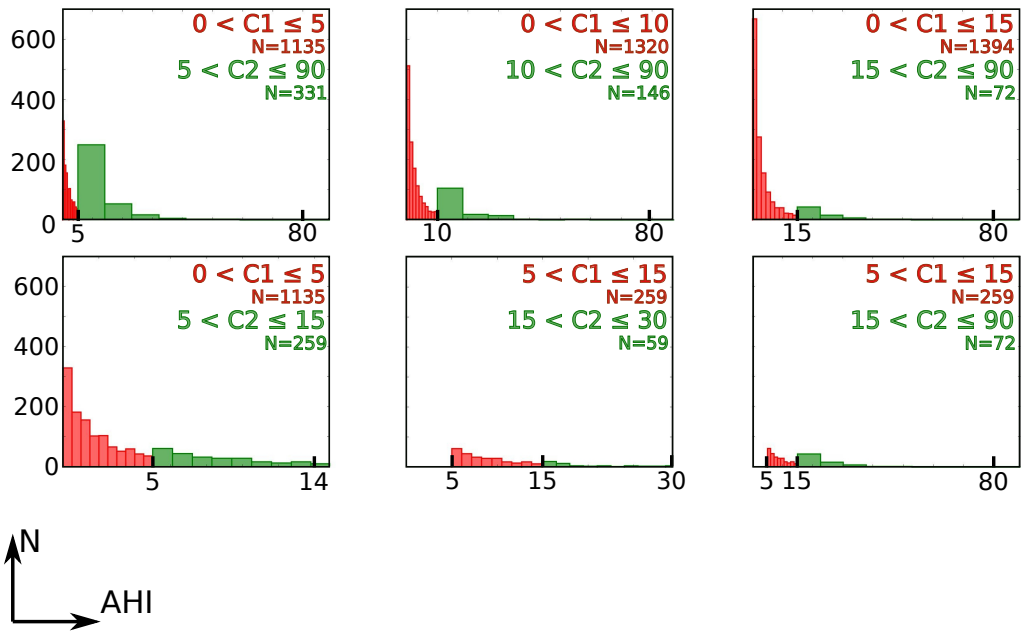

Figure 4.3: Subject groups to be distinguished.

Fig. 4.3 shows that the sample size of patients with high AHI indices (green) was much smaller than the ones with lower AHI indices. In order to keep the sample sizes of C_1 and C_2 similar, C_1 was chosen to be $N_{C_1} = N_{C_2} + 10$.

Cross validation as described in chapter 2 Sec. 2.2 was performed one hundred times for each of the two classes depicted in Fig. 4.3 and the area under the ROC curve A' was calculated for each delay pair as the performance measure. The mean was taken over the 100 A's ($\bar{A}'$).

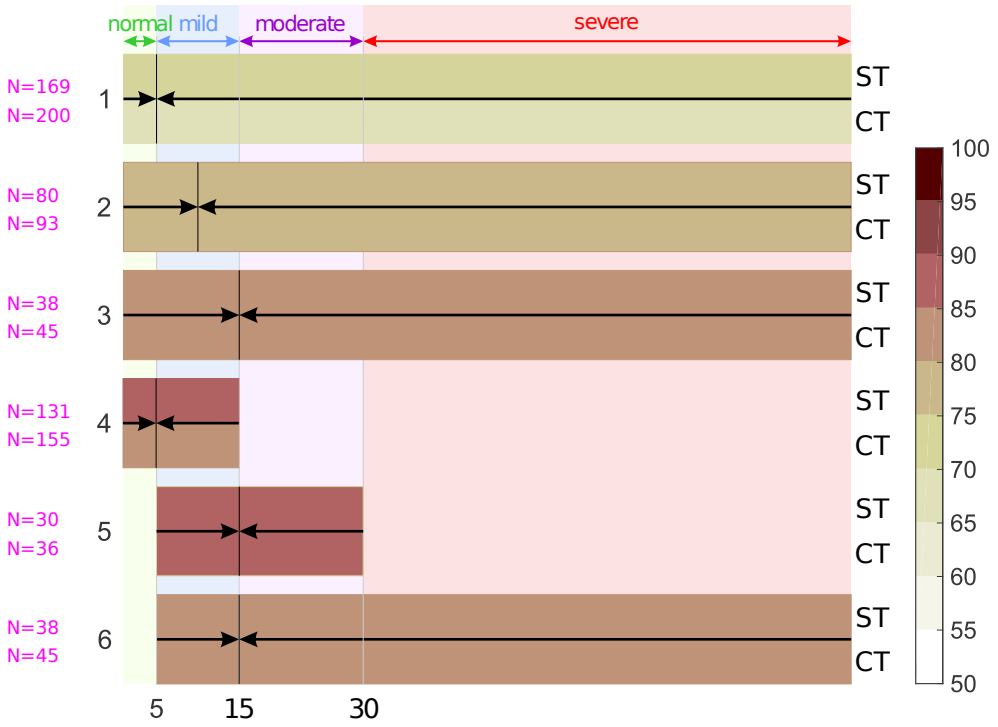

Figure 4.4: Separation performance $\bar{A}'$ for each of the 6 groups of Fig. 4.3 containing C_1 and C_2.

Fig. 4.4 shows the average separating values of $\bar{A}'$ as a color scale using randomly selected subjects of the class C_1 and all the subjects of class C_2 in Fig. 4.3 (10 repetitions each with a newly selected pool of subjects in C_1). Darker colors represent better classification and vice versa. The left hand side values indicate which of the 6 groups according to Fig. 4.3 were chosen. N indicates the sample sizes and the bottom scale represents the corresponding AHI indices. We can see that almost all class combinations achieve satisfactory separations of $\bar{A}' > 0.7$ and that cross trial and single trial performance is nearly identical.

DDA was able to detect dynamical differences in single electrode EEG data of different OSA severities. As previously described in chapter 2 (see Fig. 2.4 and Eq. 2.7), comparing ST and CT outputs serves as a test for dynamical coherency. Since the degree of a subjects apnea should not change in one night, i.e. the dynamics should stay the same, the delay pairs chosen

to be best for separation of each of the groups in Fig. 4.3 should be similar for the ST and CT case. This hypothesis was tested by visualizing which delay pairs were chosen according to the value of $\bar{A}'$, for each of the 6 groups containing C_1 and C_2 in a 2-dim histogram (100 runs). Since the subjects in C_1 were different each run, the best chosen delay pairs slightly differed each run and we can see the largest dispersion of delay pairs in Fig. 4.5 in groups 2 and 3. This is logical, because the pool of C_1 is the largest in these groups. It is obvious that ST and CT delay pairs coincide very nicely which endorses the previous analysis.

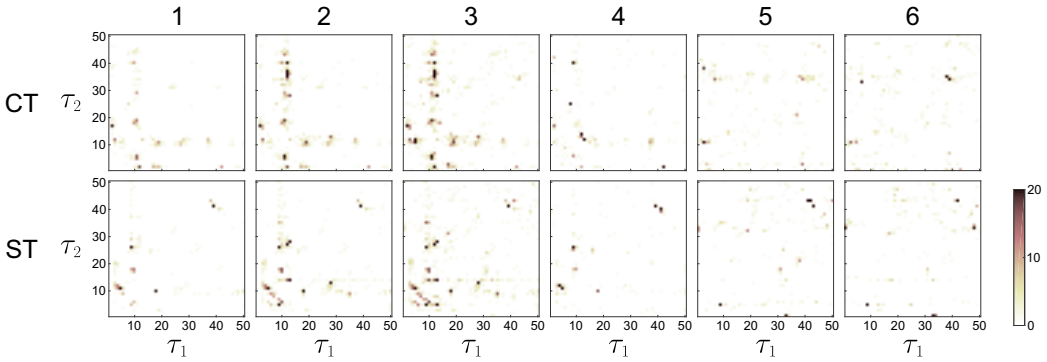

Figure 4.5: Histogram of best chosen delay pairs.

There have been other studies on the SHHS data set deriving an automatic apnea detection technique. Most of these utilize the respiratory data of the PSG recordings although some show promising results when regarding the EEG data.

Al-Angari and Sahakin[3] used 50 control and 50 OSA subjects and a support vector machine classifier with linear and second order polynomial kernels. They evaluated features from the magnitude and phase of the thoracic and abdominal respiratory effort signals for OSA detection. For subject classification an accuracy of 95% was achieved.

Biswal et al.[13] used a deep neural network approach and attained 80.2% (Table 1[13]) accuracy using a dataset of PSGs performed at the Massachusettes General Hospital Sleep Laboratory as a training set and the SHHS as a testing set. 9000 subjects were used for training and 1000 for testing. They re-validated the generalization of the neural network by further testing on 1000 randomly selected patients from the SHHS data set. They achieved a correlation of 0.77 to expert labeling and an accuracy of 80.2%.

The results presented in this thesis regarding automatic OSA classification barely scratch the tip of the iceberg and more time is required to delve deeper into this issue. The preliminary results presented here, only separate two distinct groups based on the apnea severity (eg. normal vs. mild OSA). In order to compare with [13], a model able to separate one AHI severity group from all the other groups, defined by the AASM (see Fig 4.1), must be determined.

Chapter 5

Conclusion

The results presented in the previous chapter show promise, that delay differential analysis (DDA) is a suitable classifier for sleep staging and sleep apnea. Using a nonlinear method for studying biological times series data can be highly beneficial, as nature is nonlinear. DDA is a noise insensitive technique inspired by nonlinear dynamics, in particular embedding theory. This thesis represents the preliminary results on its application to single electrode sleep EEG data and is the first step in achieving a reliable commercial continuous sleep monitoring device.

Sleep is often an overlooked factor in the well-being of humans, partially because of the effort and cost required in scoring sleep. We cycle through various stages of sleep during the night, each one characterized by different dynamics. This research applied DDA to two different sets of polysomnography data (SHHS $N = 1449$ and Physionet $N = 147$) and was able to capture dynamical differences of the 6 sleep stages with a mean correlation to human scorers of $\approx 72\%$ and $\approx 73\%$ for the two datasets.

DDA was also applied single-trial (ST) and cross-trial (CT) to the SHHS data set to try and distinguish between patients suffering from the sleep breathing disorder obstructive sleep apnea. Data was split into 6 different groups of two classes according to the apnea-hypopnea-index of the patients. The performing measure used was the area under the receiving operating characteristic (ROC) curve. ST had a mean separation of the two classes of 0.82 and CT of 0.79.

The results of this thesis must be reevaluated with other performance measures in order to reliably compare with literature. The statistical performance measures typically used for sleep scoring and sleep apnea classification

is the Cohen's-kappa index and accuracy determined by the confusion matrix.

5.1 Outlook

The results presented in this thesis show promising results, however this is just the tip of the iceberg and there is an abundant amount still to be researched in the application of DDA on sleep EEG data.

Concerning the automated sleep scoring algorithm, it would be beneficial to (1) review the data of the subjects who's hypnograms badly correlated with the DDA outputs with a sleep expert (2) because the algorithm was trained on only 35 patients, retrain the model to make sure the best delay pairs are being used for the electrode C3/A2 and then train as well for topographically different electrodes (3) include patients scored by the AASM scoring system and (4) utilize and score the output DDA features in other ways to compare this method with literature and make it fit for clinical use.

The results show interesting results as it seams that there are distinguishing features present in the EEG signals of OSA patients during REM sleep. Further investigation is needed such as (1) repetition for other sleep stages (2) testing on other data sets (3) include patients scored by the AASM scoring system (4) reapplication to CSA patients could give superior results as CSA events occur predominantly during REM sleep (5) see if there is a direct correlation between the AHI index (or event duration) and the DDA outputs and (6) ultimately test for apnea event detection.

Regarding the implementation a portable EEG device with the tattoo electrodes, a device needs to be built, that is compatable with the tattoo electrodes presented in Ferrari et. al.[27] an suited for sleep. Then the algorithm needs to be first tested on a short nap for the C3/A2 electrode. Afterwards, it can be tested and retrained for data arising from the forehead as well as behind the ears. The final step would be to include the tattoo electrode data in a clinical PSG as a proof of concept for obtaining hypnograms.

Bibliography

[1] G. A. Gould, M. Gugger, J. Molloy, V. Tsara, C. Shapiro, and N. J. Douglas. Breathing pattern and eye movement density during rem sleep in humans. *The American review of respiratory disease*, 138:874–877, 1988.

[2] D. Aeyels. Generic observability of differentiable systems. *SIAM Journal for Control and Optimization*, 19:595, 1981.

[3] H. M. Al-Angari and A. V. Sahakin. Automated recognition of obstructive sleep apnea syndrome using support vector machine classifier. *IEEE Trans Inf Technol Biomed.*, 16(3):463–468, 2015.

[4] P. Alhoa and P. Polo-Kantola. Sleep deprivation: Impact on cognitive performance. *Neurophsychiatr Dis Treat*, 3(5):553–567, 2007.

[5] Neilly J. B., Gaipa E. A., Maislin G., and Pack A. Ventilation during early and late rapid-eye-movement sleep in normal humans. *Journal of Applied Physiology*, 71: 1201–1232, 1991.

[6] S. Banks and D. F. Dinges. Behavioral and physiological consequences of sleep restriction. *Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine*, 3(5):519–28, 2007.

[7] M. Basner, B. Griefahn, and T. Penzel. Inter-rater agreement in sleep stage classification between centers with different backgrounds. *Somnologie - Schlafforschung und Schlafmedizin*, 12(1):75–84, 2008.

[8] H. F. Becker, A. J. Piper, W. E. Flynn, S. G. McNamara, R. R. Grunstein, J. H. Peter, , and C. E. Sullivan. Breathing during sleep in patients with nocturnal desaturation. *American Journal of Respiratory and Critical Care Medicine*, 159: 112–118, 1999.

[9] I. Bendixson. Sur les courbes définies par des équations différentielles. *Acta Mathematica*, 24:1–88, 1901.

[10] R. B. Berry. Chapter 15 - obstructive sleep apnea syndromes: Definitions, epidemiology, diagnosis, and variants. In R. B. Berry, editor, *Fundamentals of Sleep Medicine*, pages 237 – 261. W.B. Saunders, Saint Louis, 2012. ISBN 978-1-4377-0326-9.

[11] R. B. Berry. Chapter 16 - pathophysiology of obstructive sleep apnea. In R. B. Berry, editor, *Fundamentals of Sleep Medicine*, pages 263 – 279. W.B. Saunders, Saint Louis, 2012. ISBN 978-1-4377-0326-9.

[12] R. B. Berry, R. Brooks, C. E. Gamaldo, S. M. Harding, R. M. Lloyd, C. L. Marcus, and B. V. Vaughn. *The AASM Manual for the Scoring of Sleep and Associated Events: RULES, TERMINOLOGY AND TECHNICAL SPECIFICATIONS*. American Academy of Sleep Medicine, 1 edition, 2007.

[13] S. Biswal, H. Sun, B. Goparaju, M. B. Westover, J. Sun, and M. T. Bianchi. Expert-level sleep scoring with deep neural networks. *J Am Med Inform Assoc.*, 25(12): 1643–1650, 2018.

[14] M. S. Blumberg and A. M. H. Seelke. The form and function of infant sleep: From muscle to neocortex. In M. S. Blumberg, J. H. Freeman, and S. R. Robinson, editors, *Oxford Handbook of Developmental Behavioural Neuroscience*, pages 391–423. Oxford University Press, New York, 2010.

[15] Ludwig Boltzmann. *Vorlesungen über Gastheorie*. Leipzig, J.A. Barth, 1898. ("Ergoden" in chapter III, §32).

[16] Maxime Bonjean, Tanya Baker, Maxim Bazhenov, Sydney Cash, Eric Halgren, and Terrence Sejnowski. Interactions between core and matrix thalamocortical projections in human sleep spindle synchronization. *Journal of Neuroscience*, 32(15):5250–5263, 2012.

[17] R. Boyce, S. D. Glasgow, S. Williams, and A. Adamantidis. Causal evidence for the role of rem sleep theta rhythm in contextual memory consolidation. *Science*, 352 (6287):812–816, 2016.

[18] Letillier C. Nonlinear dynamical systems, course notes, 2019. Master IBIOM, Normandie University (Rouen).

[19] M. Carskadon, W. Dement, M. Kryger, T. Roth, and T. Roehrs. Normal human sleep: An overview. *Principles and Practice of Sleep Medicine*, 2, 2005.

[20] K. Cheshire, H. M. Engleman, I. J. Deary, C. Shapiro, and N. J. Douglas. Factors impairing daytime performance in patients with sleep apnea/hypopnea syndrome. *Archives of internal medicine*, 152:538–41, 04 1992.

[21] H. Danker-Hopfe, P. Anderer, J. Zeitlhofer, M. Boeck, H. Dorn, G. Gruber, E. Heller, E. Loretz, D. Moser, S. Parapatics, B. Saletu, A. Schmidt, and G. Dorffner. Interrater reliability for sleep scoring according to the rechtschaffen & kales and the new aasm standard. *Journal of Sleep Research*, 18(1):74–84, 2009.

[22] M. de Zambotti, S. Claudatos, S. Inkelis, I. Colrain, and F. Baker. Evaluation of a consumer fitness-tracking device to assess sleep in adults. *Chronobiology International*, 2015.

[23] D. Farmer, J. Crutchfield, H. Froehling, N. Packard, and R. Shaw. Power spectra and mixing properties of strange attractors. *Annals of the New York Academy of Sciences*, 357(1):453–471, 1980.

[24] E. Fehrmann. *Automated sleep classification using the new sleepstage standards*. PhD thesis, Rochester Institute of Technology, 2013.

[25] I. Fernández-Varela, E. Hernández-Pereira, D. Álvarez-Estévez, and V. Moret-Bonillo. A convolutional network for sleep stages classification. *Proceedings*, 2, 2019.

[26] F. Ferrarelli, R. Huber, M. J. Peterson, Marcello M., M. Murphy, B. A. Riedner, A. Watson, P. Bria, and G. Tononi. Reduced sleep spindle activity in schizophrenia patients. *American Journal of Psychiatry*, 164(3):483–492, 2007.

[27] L. M. Ferrari, S. Sudha, S. Tarantino, R. Esposti, F. Bolzoni, P. Cavallari, C. Cipriani, V. Mattoli, and F. Greco. Ultraconformable temporary tattoo electrodes for electrophysiology. *Advanced Science*, 5(3), 2018.

[28] C. Fisher, H. D. Cohen, R. C. Schiavi, D. Davis, B. Furman, K. Ward, A. Edwards, and J. Cunningham. Patterns of female sexual arousal during sleep and waking: Vaginal thermo-conductance studies. *Archives of Sexual Behavior*, 12(2):97–122, Apr 1983. ISSN 1573-2800. doi: 10.1007/BF01541556. URL `https://doi.org/10.1007/BF01541556`.

[29] S. Fogel, R. Nader, K.A. Cote, and C. Smith. Sleep spindles and learning potential. *Behavioral neuroscience*, 121:1–10, 2007.

[30] K. A. Franklin and E. Lindberg. Obstructive sleep apnea is a common disorder in the population—a review on the epidemiology of sleep apnea. *Journal of Thoracic Disease*, 7(8):1311–1322, 2015.

[31] C. Grebogi, E. Ott, S. Pelikan, and J. A. Yorke. Strange attractors that are not chaotic. *Physica D: Nonlinear Phenomena*, 13(1):261 – 268, 1984.

[32] Hamel. Georg duffing, ingenieur: Erzwungene schwingungen bei veränderlicher eigenfrequenz und ihre technische bedeutung. sammlung vieweg. heft 41/42, braunschweig 1918. vi+134 s. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik*, 1(1):72–73, 1921.

[33] W. Hassler. Differentiable manifolds. *Ann. Math.*, 37:645–680, 1936.

[34] E. Hopf. Abzweigung einer periodischen Lösung von einer stationären Lösung eines Differentialsystems. *Akad. Wiss. Leipzig*, 94:1–22, 1942.

[35] J. Horne. Human slow wave sleep: A review and appraisal of recent findings, with implications for sleep functions, and psychiatric illness. *Experientia*, 48(10):941–954, 1992.

[36] S. A. Imtiaz and E. Rodriguez-Villegas. Automatic sleep staging using state machine-controlled decision trees. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 378–381, 2015.

[37] Argyris J., Faust G., Haase M., and Friedrich R. *Die Erforschung des Chaos-Dynamische Systeme*, volume 3. Springer Vieweg, Heidelberger Platz 3, 14197 Berlin, Germany, 2017.

[38] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2*, IJCAI'95, pages 1137–1143. Morgan Kaufmann Publishers Inc., 1995.

[39] P. Y. Ktonas, S. Golemati, P. Xanthopoulos, V. Sakkalis, M. D. Ortigueira, H. Tsekou, M. Zervakis, T. Paparrigopoulos, and C. R. Soldatos. Potential dementia biomarkers based on the time-varying microstructure of sleep eeg spindles. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2464–2467, 2007.

[40] M.D. L. J. Findley, S. C. Wilhoit, M.E., and P. M. Suratt. Apnea duration and hypoxemia during rem sleep in patients with obstructive sleep apnea. *CHEST*, 87(4): 432–436, 1985.

[41] C. Lainscsek. *Identification and Global Modeling of Nonlinear Dynamical Systems*. PhD thesis, Technische Universität Graz, 1999.

[42] C. Lainscsek. Nonuniqueness of global modeling and time scaling. *Phys. Rev. E*, 84: 046205, 2011.

[43] C. Lainscsek and T.J. Sejnowski. Electrocardiogram classification using delay differential equations. *Chaos*, 23(2):023132, 2013.

[44] C. Lainscsek and T.J. Sejnowski. Delay differential analysis of time series. *Neural Computation*, 27(3):594–614, 2015.

[45] C. Lainscsek, C. Letellier, and F. Schürrer. Ansatz library for global modeling using a structure selection. *Physical Review E*, 64:016206/1–15, 2001.

[46] C. Lainscsek, C. Letellier, and I. Gorodnitsky. Global modeling of the Rössler system from the z-variable. *Physics Letters A*, 314:409–127, 2003.

[47] C. Lainscsek, M. E. Hernandez, J. Weyhenmeyer, T. J. Sejnowski, and H. Poizner. Non-linear dynamical analysis of EEG time series distinguishes patients with Parkinson's disease from healthy individuals. *Frontiers in Neurology*, 4(200), 2013.

[48] C. Lainscsek, M. E. Hernandez, J. Weyhenmeyer, T. J. Sejnowski, and H. Poizner. Non-linear dynamical analysis of EEG time series distinguishes patients with Parkinson's disease from healthy individuals. *Frontiers in Neurology*, 4(200), 2013.

[49] C. Lainscsek, J. Weyhenmeyer, M.E. Hernandez, H. Poizner, and T.J. Sejnowski. Non-linear dynamical classification of short time series of the Rössler system in high noise regimes. *Frontiers in Neurology*, 4(182), 2013.

[50] C. Lainscsek, J. Weyhenmeyer, M.E. Hernandez, H. Poizner, and T.J. Sejnowski. Non-linear dynamical classification of short time series of the Rössler system in high noise regimes. *Frontiers in Neurology*, 4(182), 2013.

[51] C. Lainscsek, V. Messager, A. Portman, T. J. Sejnowski, and C. Letellier. Automatic sleep scoring from a single electrode using delay differential equations. In J. Awrejcewicz, editor, *Applied Non-Linear Dynamical Systems, Springer Proceedings in Mathematics & Statistics*, volume 93, pages 371–382. Springer, 2014.

[52] C. Lainscsek, J. Weyhenmeyer, S. S. Cash, and T. J. Sejnowski. Delay differential analysis of seizures in multi-channel electrocorticography data. *Naural Comutation*, 29, 2017.

[53] C. Lainscsek, A. L. Sampson, R. Kim, M. L. Thomas, K. Man, X. Lainscsek, N. R. Swerdlow, D. L. Braff, T. J. Sejnowski, and G. A. Light. Nonlinear dynamics underlying sensory processing dysfunction in schizophrenia. *Proceedings of the National Academy of Sciences*, 116(9):3847–3852, 2019.

[54] X. Lainscsek. *Delay Differential Analysis on the Lorenz System.* Bachelor's thesis, Graz University of Technology, 2014.

[55] J.-M. Lee, W. Byun, A. Keill, D. Dinkel, and Y. Seo. Comparison of wearable trackers' ability to estimate sleep. *International Journal ofEnvironmental Researchand Public Health*, 15(6):1265, 2018.

[56] C. Letellier and J-M. Malasoma. Architecture of chaotic attractors for flows in the absence of any singular point. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 26:063115, 2016.

[57] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20:130 – 141, 1963.

[58] P. Maquet. The role of sleep in learning and memory. *Science*, 294:1048–52, 2001.

[59] M. Marino, Y. Li, M. N. Rueschman, J. W. Winkelman, J. M. Ellenbogen, J. M. Solet, H. Dulin, Berkman L. F., and O. M. Buxton. Measuring sleep: Accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep*, 36(11):1747–1755, 2013.

[60] G. A. Marks, J. P. Shaffery, A. Oksenberg, S. G. Speciale, and H. P. Roffwarg. A functional role for rem sleep in brain maturation. *Behavioural Brain Research*, 69(1): 1 – 11, 1995.

[61] O. Ménard, C. Letellier, J. Maquet, L. LE Sceller, and G. Gouesbet. Analysis of a nonsynchronized sinusoidally driven dynamical system. *International Journal of Bifurcation and Chaos*, 10(07):1759–1772, 2000.

[62] C. E. Metz. Basic principles of roc analysis. *Seminars in Nuclear Medicine*, 8(4):283 – 298, 1978.

[63] E. Miletics and G. Molnárka. Taylor series method with numerical derivatives for initial value problems. *J. Comp. Methods in Sci. and Eng.*, 4(1,2):105–114, 2004. URL `http://dl.acm.org/citation.cfm?id=1411367.1411380`.

[64] J. Milnor. On the concept of attractor. *Communications in Mathematical Physics*, 99(2):177–195, 1985.

[65] Anand P. and Vandana R. An automatic detection of sleep using different statistical parameters of single channel eeg signals. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 9:335–344, 2016.

[66] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw. Geometry from a time series. *Phys. Rev. Lett.*, 45:712, 1980.

[67] J. Peever and P. M. Fuller. The biology of rem sleep. *Current Biology*, 27:R1237–R1248, 11 2017.

[68] J. L. Pépin, P. Leger, D. Veale, B. Langevin, D. Robert, and P. Lévy. Side effects of nasal continuous positive airway pressure in sleep apnea syndrome: Study of 193 patients in two french sleep centers. *Chest*, 107(2):375 – 381, 1995.

[69] D. Petit, J.-F. Gagnon, M. L. Fantini, L. Ferini-Strambi, and J. Montplaisir. Sleep and quantitative eeg in neurodegenerative disorders. *Journal of Psychosomatic Research*, 56(5):487 – 496, 2004.

[70] H. Poincaré. Mémoire sur les courbes définies par une équation différentielle (i,ii,iii,iv). *Journal de Mathématiques Pures et Appliquées*, 7,8,1,2:157–217, 1881,1882,1885,1886.

[71] H. Poincaré. Sur l'équilibre d'une masse fluide animée d'un mouvement de rotation. *Acta Math.*, 7:259–380, 1885.

[72] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, New York, NY, USA, 1990.

[73] A. Rechtschaffen and A.Kales. *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects.* University of California Los Angeles Brain Information Service, NINDB Neurological Information Network,, 1968.

[74] T. Roque. Stability of trajectories from poincaré to birkhoff: approaching a qualitative definition. *Archive for History of Exact Sciences*, 65(3):295–342, 2011.

[75] T. Roth. Slow wave sleep: Does it matter? *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine*, 5(2):S4–S5, 2009.

[76] A. Sampson, C. Lainscsek, C. Gonzalez, I. Ulbert, O. Devinsky, D. Fabó, E. Halgren, S. S. Cash, and T. J. Sejnowski. Delay differential analysis for dynamical sleep spindle detection. *Journal of Neuroscience Methods*, 2019.

[77] Tim Sauer, James A. Yorke, and Martin Casdagli. Embedology. *Journal of Statistical Physics*, 65:579, 1991.

[78] M. Schabus, G. Gruber, S. Parapatics, C. Sauter, G. Klösch, P. Anderer, W. Klimesch, B. Saletu, and J. Zeitlhofer. Sleep Spindles and Their Significance for Declarative Memory Consolidation. *Sleep*, 27(8):1479–1485, 2004.

[79] T. J. Sejnowski and A. Destexhe. Why do we sleep? *Brain Research*, 886(1):208 – 223, 2000.

[80] F. Siclari, G. Bernardi, J. Cataldi, and G. Tononi. Dreaming in nrem sleep: A high-density eeg study of slow waves and spindles. *The Journal of Neuroscience*, 38: 9175–9185, 2018.

[81] A. S. Stöberl, E. I. Schwarz, S. R. Haile, C. D. Turnbull, V. A. Rossi, J. R. Stradling, and M. Kohler. Night-to-night variability of obstructive sleep apnea. *Journal of Sleep Research*, 26(6):782–788, 2017.

[82] R. Stoop and W.H. Steeb. *Berechenbares Chaos in dynamischen Systemen*, volume 1. Birkhäuser Basel, 2006.

[83] F. Takens. Detecting strange attractors in turbulence. In D. A. Rand and L.-S. Young, editors, *Dynamical Systems and Turbulence, Warwick 1980*, volume 898 of *Lecture Notes in Mathematics*, pages 366–381. Springer Berlin/Heidelberg, 1981.

[84] R. J. Thomas, S. Chokroverty, and S. Bhat. Chapter 3 - sleep stages and scoring technique. In S. Chokroverty and R. J. Thomas, editors, *Atlas of Sleep Medicine (Second Edition)*, pages 77 – 99. W.B. Saunders, St. Louis, second edition edition, 2014. ISBN 978-1-4557-1267-0.

[85] R. J. Thomas, S. Chokroverty, and S. Bhat. Chapter 4 - sleep-disordered breathing and scoring. In S. Chokroverty and R. J. Thomas, editors, *Atlas of Sleep Medicine (Second Edition)*, pages 100 – 118. W.B. Saunders, St. Louis, second edition edition, 2014. ISBN 978-1-4557-1267-0.

[86] Zhouchao Wei. Dynamical behaviors of a chaotic system with no equilibria. *Physics Letters A*, 376(2):102 – 108, 2011.

[87] L. Zhang, D. Fabbri, R Upender, and D Kent. Automated sleep stage scoring using deep learning. *Sleep*, 41:A118–A118, 04 2018.