

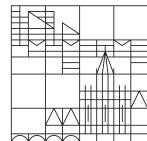
Non-linear methods for the quantification of cyclic motion

Dissertation zur Erlangung des akademischen Grades des Doktors der
Ingenieurwissenschaften an der Universitat Konstanz

vorgelegt von

Juan Carlos Quintana Duque

Universität
Konstanz



Mathematisch-Naturwissenschaftliche Sektion
Fachbereich Informatik und Informationswissenschaft

Tag der mündlichen Prüfung: 27.07.2016

Referenten:

Prof. Dr. Dietmar Saupe, Universität Konstanz
Prof. Dr. Sven Kosub, Universität Konstanz

Abstract

Traditional methods of human motion analysis assume that fluctuations in cycles (e.g. gait motion) and repetitions (e.g. tennis shots) arise solely from noise. However, the fluctuations may have enough information to describe the properties of motion. Recently, the fluctuations in motion have been analysed based on the concepts of variability and stability, but they are not used uniformly. On the one hand, these concepts are often mixed in the existing literature, while on the other hand, these concepts have been related to different methods for non-linear and chaotic time series analysis. After the clarification of these concepts, this dissertation presents the analysis of the evidence for chaos in cycling motion. Further, new algorithms are described for the analysis of the variability of the observed variables from the gait motion by means of a novel curve registration method. Finally, a novel approach for the estimation of torque variation from pedal motion during cycling ergometry is introduced and validated. This thesis includes three main parts.

In the first part, a case study of knee motion during cycling ergometry is presented using the state-of-the-art tools. Non-linear methods and hypothesis testing based on surrogate data were used to analyse whether the aperiodic behaviour of noisy data captured from knee motion data originating from a low-dimensional deterministic (possibly chaotic) process or whether it was governed by stochastic ones. The critical sensitivity of chaotic systems to both the initial conditions and perturbations may explain the irregular behaviour of pedalling motion. However, the time series recorded from knee motion did not show a clear scaling region typical in chaotic time series. Further, evidence of chaos based on surrogate data could not be found for all time series. In addition, an improvement of the statistical criterion for hypothesis testing based on surrogate data is discussed.

In the second part, a novel tool for curve registration that facilitates the analysis of cycles in cyclic motion is presented and validated using gait acceleration data. The equalized DBA (eDBA) method calculates the average of a set of cycles based on dynamic time warping (DTW) and a modification of DTW barycentric averaging (DBA). The eDBA algorithm allows the study of the kinematic variables in cyclic motion depending on the phase using the eDBA average cycle as the reference for phase registration. A novel quality definition is given which provides a criterion for the selection of the best phase angle for further analysis. Further, the effects of phase registration using the eDBA method on Self-Organizing Maps (SOMs) are described. The quality of the SOMs and the classification rate improved when the registration was applied in the preprocessing step.

In the third part, a novel approach for the estimation of torque variation from pedal motion in cycling ergometry is presented and validated. For an ergometer with almost constant brake torque, we may assume that variations in the net torque can be extracted from the pedal motion alone. The key problem is to reliably estimate the angular pedal acceleration from measurements of pedal motion, which can be made in the laboratory using motion-capturing with two light-emitting diodes (LEDs) or a plain commercial video camera with two light-emitting diodes (LEDs). The results from video were close to motion capturing data (MoCap) results when a novel method for the correction of the marker position was applied.

Zusammenfassung

Traditionelle Methoden in der menschlichen Bewegungsanalyse gehen davon aus, dass die Schwankungen in Zyklen (z.B. Gangbewegung) und Wiederholungen (z.B. Tennisschlag) ausschließlich aus Rauschen stammen. Die Schwankungen können jedoch genügend Informationen enthalten, die die Eigenschaften der zu analysierenden Bewegung beschreiben können. In der Literatur werden zyklische Bewegungen durch die Begriffe Variabilität und Stabilität analysiert. Dabei werden die oben genannten Konzepte jedoch nicht einheitlich verwendet, wodurch eine unmittelbare Vergleichbarkeit der vorhandenen Ergebnisse nicht immer möglich ist. Auf der einen Seite werden diese Begriffe oft in der bestehenden Literatur gemischt, während auf der anderen Seite haben diese Konzepte auf verschiedene Methoden zur nichtlinearen und chaotischen Zeitreihenanalyse in Verbindung gebracht. In dieser Dissertation werden nach der Erklärung der Begriffe sowohl nicht-lineare Analysen mit Hilfe der State-of-the-Art-Tools, als auch multidimensionale Analysen von beobachteten Variablen mit Hilfe einer neuartigen Kurvenregistrierungsmethode verwendet. Ferner wurde die Schätzung der Drehmomentvariation von Pedalbewegung während der Fahrradergometrie mit einem neuen Verfahren zur Korrektur der Pedalposition erreicht. Diese Arbeit umfasst drei Hauptteile.

Im ersten Teil wird eine Fallstudie über Pedalierbewegungen auf dem Fahrradergometer vorgestellt. Nicht-lineare Methoden und Hypothesentests mit Surrogate-Daten werden verwendet, um zu analysieren, ob das aperiodische Verhalten von Kniebewegungsdaten aus einem (möglicherweise chaotischen) niedrig-dimensionalen deterministischen Prozess stammt, oder ob es durch einen stochastischen Prozess erklärt werden kann. Die kritische Empfindlichkeit von chaotischen Systemen, sowohl bezüglich der Anfangsbedingungen als auch der Störungen, können das unregelmäßige Verhalten der Tretbewegung erklären. Allerdings zeigen die gemessenen Zeitreihen von Kniebewegungen keine deutliche Skalierungsregion, wie dies bei chaotischen Zeitreihen zu erwarten wäre. Ferner wurde kein ausreichender Beweis für die meisten Zeitreihen durch die Hypothesentests mit Surrogate-Daten gefunden. Darüber hinaus wird eine Verbesserung der statistischen Kriterien zur Überprüfung der Hypothese durch Surrogate-Daten diskutiert.

Im zweiten Teil wird ein neues Werkzeug für die Registrierung von Kurven vorgestellt, das die Analyse von Bewegungsdaten ermöglicht. Das Werkzeug wurde mit Beschleunigungsdaten von Gangbewegung validiert. Der eDBA Algorithmus berechnet den Durchschnitt einer Reihe von Zyklen durch dynamic time warping (DTW) und eine Modifikation der DTW barycentric averaging (DBA). Der eDBA Algorithmus ermöglicht die Untersuchung der Eigenschaften von kinematischen Variablen in zyklischen Bewegungen in Abhängigkeit von der Phase. Die Phasenregistrierung nutzt den eDBA durchschnittlichen Zyklus als Referenz. Ferner wurde ein Kriterium für die Qualität der Informationsgehalt in jedem Phasenwinkel vorgeschlagen, das die Auswahl der besten Phasenwinkel für weitere Analysen zur Verfügung stellt. Schließlich wird die Auswirkung der Phase-Registrierung mit dem eDBA Algorithmus auf Self-Organizing Maps (SOMs) beschrieben. Die Qualität der SOMs und die Klassifikationsrate haben sich verbessert, als die Phase-Registrierung

in der Vorverarbeitung der Daten angewendet wurde.

Im dritten Teil wird ein neuartiger Algorithmus zur Abschätzung der Drehmomentschwankung von Pedalbewegung bei Fahrten auf dem Radergometer vorgestellt und validiert. Die Variation der Winkelbeschleunigung ist proportional zu der Variation des aufgebrachten Drehmoments, wenn die Ergometerbremskraft bzw. -widerstand konstant ist. Wir schlagen vor, die Variation des Netto-Drehmoments aus Messungen der Pedalbewegung abzuleiten, die im Labor mit Bewegungserfassungsgeräten oder kommerziellen Videokameras aufgenommen werden können. Die Rekonstruktionen der Position und der Ausrichtung des Pedals wurden aus Videoaufnahmen von zwei Leuchtdioden und aus Bewegungserfassung mit zwei aktiven Infrarot-Markern durchgeführt. Die Ergebnisse aus Video waren ähnlich wie die Ergebnisse aus MoCap, wenn ein neues Verfahren zur Korrektur der Position angewendet wurde.

Acknowledgements

This thesis is dedicated to my parents and brother who supported me all the way from the beginning of my studies.

I would like to thank my advisor Prof. Dr. Dietmar Saupe for continual support throughout my Ph.D. He gave me the scientific freedom to choose and to pursue my topic and guided me by giving valuable advice when it was needed.

I would like to thank Thorsten Dahmen, Alexander Artiga Gonzalez, Raphael Bertschinger, and all my colleagues in the multimedia signal processing workgroup for fruitful collaboration and joyful research discussions.

I would like to thank Ingrid Baiker, Dr. Markus Herz, and Dr. Arno Weiershäuser for being always helpful, friendly and for taking care of me during my Ph.D.

A very special thanks to Lina María Sanabria for supporting me with her love during the last two years.

This work was partially supported by the DFG Research Training Group GRK-1042 “Explorative Analysis and Visualization of Large Information Spaces” and by the DFG project SA 449/12-1 “Powerbike: Model-based optimal control for cycling” ¹.

¹<https://www.informatik.uni-konstanz.de/saupe/forschung/laufende-projekte/powerbike/>

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 1.1 | Thesis problems and contributions | 2 |
| 1.1.1 | Study of the evidence of chaos in cycling motion | 2 |
| 1.1.2 | Phase registration, cycle averaging and classification | 3 |
| 1.1.3 | Estimation of torque variation from noisy motion capturing or video data | 4 |
| 1.2 | My publications | 4 |
| 1.3 | Thesis outline | 5 |
| 2 | Concepts related to human movement | 7 |
| 2.1 | Motion capturing and human motion data | 7 |
| 2.1.1 | Properties of human motion data | 8 |
| 2.2 | Movement variability, stability and redundancy | 10 |
| 2.2.1 | Coordinative variability and end-point variability | 11 |
| 2.3 | Chaos and human motion as a dynamical system | 12 |
| 3 | Review of non-linear methods for time series | 15 |
| 3.1 | Dynamical systems and the reconstruction theorem | 16 |
| 3.2 | Embedding space | 20 |
| 3.2.1 | Calculation of embedding parameters from time series | 20 |
| 3.2.2 | Problems and limitations | 23 |
| 3.3 | Dynamic invariants | 25 |
| 3.3.1 | Maximal Lyapunov exponent | 26 |
| 3.3.2 | Fractal dimension and correlation dimension | 27 |
| 3.3.3 | Recurrence period density entropy (RPDE) | 30 |
| 3.4 | Definition of chaos | 30 |
| 3.4.1 | Which dynamical systems do exhibit chaos? | 31 |
| 4 | Analysis of evidence of chaos in cycling ergometry | 33 |
| 4.1 | Surrogate data hypothesis testing | 34 |
| 4.1.1 | Methods for surrogate data | 35 |
| 4.1.2 | Rejection of the null hypothesis | 40 |
| 4.2 | Analysis of evidence of chaos in motion data of cycling ergometry | 44 |
| 4.2.1 | Methods | 45 |
| 4.2.2 | Results | 48 |
| 4.3 | Conclusions | 51 |

| | |
|--|------------|
| 5 Parametrization of cyclic motion and self-organizing maps | 59 |
| 5.1 Analysis of multi-dimensional variables with transversal sections | 62 |
| 5.2 The equalized DBA (eDBA) method | 63 |
| 5.2.1 Data and assumptions | 63 |
| 5.2.2 Calculation of the average cycle based on eDBA | 63 |
| 5.3 Approach to phase registration using eDBA average cycle and transversal sections . | 68 |
| 5.3.1 Quality of transversal sections | 69 |
| 5.4 Self-organizing map (SOM) | 70 |
| 5.4.1 Quality of self-organizing maps | 71 |
| 5.5 Results | 72 |
| 5.5.1 Experimental Setup | 72 |
| 5.5.2 Data processing | 72 |
| 5.5.3 Parametrization and quality of transversal sections | 75 |
| 5.5.4 Parametrization and self-organizing maps | 75 |
| 5.6 Conclusions | 83 |
| 6 Estimation of torque variation from pedal motion in cycling ergometry | 85 |
| 6.1 Calculation of the second derivative from positional data | 87 |
| 6.2 Correction of the marker position | 87 |
| 6.3 Related concepts: Force and torque | 88 |
| 6.3.1 Deduction of torque equation for the pedal | 90 |
| 6.4 Torque validation | 92 |
| 6.5 Experimental setup | 93 |
| 6.5.1 Data acquisition and processing | 93 |
| 6.5.2 Test design | 95 |
| 6.6 Results and discussion | 96 |
| 6.7 Conclusions | 99 |
| 7 Conclusions and future research | 101 |
| 7.1 Conclusions | 101 |
| 7.2 Future research | 102 |
| 7.2.1 Non-uniform and multidimensional embedding | 102 |
| 7.2.2 Multifractality | 103 |
| 7.2.3 Chaos control theory | 103 |
| 7.2.4 Selection of a suitable Poincaré section in an embedding space | 104 |
| 7.2.5 Curve registration and curve averaging for continuos time | 104 |
| 7.2.6 Asymmetry indices for torque variation from pedal motion in cycling | 105 |
| 7.2.7 Compensation for the bike sway in videos of pedalling motion | 105 |
| 8 Appendix | 107 |
| 8.1 Types of dynamical systems | 107 |
| 8.2 Non-linear filtering | 108 |
| 8.3 Savitzky–Golay smoothing filter | 109 |
| 8.4 Relaxation method | 110 |
| 8.5 List of acronyms | 112 |
| Bibliography | 115 |

Chapter 1

Introduction

The quest for an adequate mathematical framework for describing motion behaviour has a long tradition in human movement science. This quest relies on an extensive integration of disciplines such as sport science, physiology, medicine, physics, computer science, and mathematics. With the growing market for lower-cost and portable sensors in all kinds of forms, new approaches and ideas for the analysis of human motion are being developed. Further, new technologies are introducing a change of paradigms in the way that cyclic motion and repetitive tasks are evaluated.

Quantitative studies of human motion have typically focused on properties of the average motion and their fluctuations were considered to be noise. However, recent empirical studies, focused on gait motion and its kinematic variables, have revealed fluctuations with interesting structures, even under constant environmental conditions. These fluctuations have been used to determine disease severity, medication utility, and fall risk (Harbourne & Stergiou, 2009; Davids, Bennett, & Newell, 2006). Thus, the alleged noise in kinematic variables may contain valuable information useful for characterizing mature motor skills and healthy states.

Movement variability, defined as the amount of variation, and movement stability, given by the structure of the variation, are omnipresent and unavoidable in all kinds of motions. For example, even Olympic athletes show fluctuations in their motion performance despite several years of intensive training in particular movements. Both movement variability and movement stability are quantifiable, with a lack or excess of one or both factors resulting in a deficiency of motor skills. However, the amount of variation deemed normal is dependent upon the place in which it is measured, the level of difficulty of the associated task, and the number of degrees of freedom necessary to complete it.

An increasing amount of research has recently been directed towards the analysis of long-term measurements of human motion data. However, there is no consensus on how to best define and evaluate the variation of kinematic variables of motion data. On the one hand, the concepts of variability and stability are often mixed in the existing literature, while on the other hand, these concepts have been related to the current state-of-the-art methods for non-linear time series analysis, which are still under development. Non-linear time series analysis is the study of time series with computational techniques sensitive to non-linearity under the assumption that the time series are noisy measurements of particular solutions of a deterministic dynamical system.

In this dissertation, I explore non-linear tools for the analysis of the evidence of chaos in time series and address the question of what other algorithms, methods of analysis, and data collection techniques can be implemented to assess movement variability and stability in cycling and gait motion. Furthermore, I have developed an approach to calculate the crank angular acceleration and the torque variation from videos of pedalling motion recorded with low-speed camcorders.

1.1 Thesis problems and contributions

Cyclic motion is at the core of many sports, such as running, swimming, or cycling. The study of the corresponding kinematic variables is fundamental for the evaluation of training routines and the assessment of performance. This dissertation is devoted to the analysis of cyclic motion and divided into three major parts. The first part presents the analysis of evidence of chaos in cycling ergometry motion under different experimental constraints. The second part presents a novel method for phase registration and cycle averaging, its potential application to the analysis of the dynamics of cyclic motion, and how phase registration improves the results of classification and clustering of cycles based on self-organizing map (SOM). The third part describes and validates a novel method for the estimation of the angular pedal acceleration and torque variation in cycling ergometry from motion capturing or video data. These main parts are explained in the following.

1.1.1 Study of the evidence of chaos in cycling motion

Humans move through the participation of multiple joints and limbs, and experience the real world by exploiting all their available senses. Human motion is the product of a very complex system which cannot explicitly be known. Even if it were, it would typically contain a myriad of variables defying any mathematical or even numerical analysis. However, since the discovery of Takens (1981) of a topologically faithful reconstruction of the state space of dynamical systems in a lower dimensional, so-called embedding space, it has become possible to estimate features from the dynamics using only noisy measurements of some function in the original (unknown) state space. The time series observed from motion might help to understand the human motor system that has generated them, even if the model is unknown or too complex.

The dynamic invariants are features calculated either from the original state space or from the reconstruction of the state space using embedding. However, the dynamic invariants calculated from a time series are meaningful only when there is indeed deterministic non-linearity (i.e. chaos) in the time series. A time series can be anywhere between purely random and strictly deterministic. To validate the evidence of chaos, the method of surrogate data hypothesis testing has become a central tool, which is explained in detail in Chapter 3. A dynamic invariant, e.g. the Lyapunov exponent or correlation dimension, is applied to both the original time series and the surrogate data sets. The dynamic invariant is calculated for different sets of parameters and their average value is used as the measure of comparison between the original and the surrogate data. Finally, a statistical criterion is used to determine whether the values of the dynamic invariants of the original and surrogate are sufficiently different. If they are, the null hypothesis that the original and the surrogate data come from the same process is rejected.

There is a lack of studies examining the dynamic structure of pedaling motion and its potential benefits for distinguishing subtle differences between pedaling motion patterns. My contributions are the following:

- The study of the evidence of chaos in pedaling motion calculated from the time series of the x -coordinate of the position of the knee joint with the surrogate data methodology and some dynamic invariants.
- An extension of the statistical criterion for comparison between the discriminating statistics of surrogate and original data. This extension includes the error of the calculation of dynamic invariants as discriminating statistic, to avoid wrong rejections of hypotheses.

1.1.2 Phase registration, cycle averaging and classification

The conventional kinematic analysis of human motion is performed manually by experts deriving characteristic features from the average of a few single cycles. However, their conclusions may not be identical because their interpretations might be influenced by different levels of expertise, the way that the average cycle is calculated, and the difficulty in comprehending large amounts of information. Thus, novel methods are needed which overcome these limitations and enable large quantities of data to be analysed and interpreted objectively.

Individual cycles in cyclic motion differ in duration and local speed because the timing of muscular, neurological, and respiratory systems of the human body varies according to environmental, biomechanical, and morphological constraints. Therefore, computing cross-sectional averages over a family of misaligned curves without registration can lead to the cancellation of critical shape characteristics (Kneip & Gasser, 1992), modifications of landmarks (Chau, Young, & Redekop, 2005), and to inflated amplitude variability estimates (Sadeghi, Mathieu, Sadeghi, & Labelle, 2003). Thus, the calculation of the prototypical average curve requires that the single-cycle curves must first be registered.

The calculation of an average cycle is one main step for phase registration, data analysis by traditional statistics, and novel methods based on transversal sections. The transversal section and the analysis of intersections are similar to some tools for the analysis of dynamical systems near a periodic solution. The transversal section corresponds to the so-called Poincaré sections and the intersections with one transversal plane correspond to the so-called Poincaré map.

We proposed a novel method called equalized DBA (eDBA) to provide a tool for cycle registration that facilitates the analysis of cycles in cyclical motion data. The eDBA allows the study of features of the kinematic variables in cyclic motion depending on the phase using the eDBA average cycle as the reference for phase registration. Features can be extracted from the intersection points on each transversal section to characterize the overall motion. For example, the variance in one selected section may indicate the degree of regularity of the cyclic motion. In addition, the effects of phase registration using the eDBA method on self-organizing map (SOM) were studied. SOM have been used to reduce the complexity of joint kinematic and kinetic data in order to cluster, classify and visualize cyclic motion data.

My contributions are the following:

- An algorithm for calculating the average periodic cycle of a set of cycles based on dynamic time warping (DTW) and a modification of DTW barycentric averaging (DBA).
- A method for parametrizing cycles based on DTW and the distances to transversal sections placed on the average periodic cycle. The parametrization allows studying how the kinematic variables change with the phase angle.
- A definition of the quality of cycle intersections with transversal sections, providing a criterion for the right choice of the section.
- With acceleration data recorded from gait motion, it is shown that the quality of the SOM and the classification rate improved when the phase registration was applied in the preprocessing step.

1.1.3 Estimation of torque variation from noisy motion capturing or video data

The interaction between the cyclist and the bicycle is given by the body position and the pedaling technique. However, the latter has received less attention due to the misconception that the bike goes ahead regardless of how the force on the pedals is applied. A good pedaling technique allows the subject to minimize the energy needed to perform the motion and to avoid the risk of injuries. Applying correctly oriented forces to the pedal (i.e. pedalling technique) is a major component of skilled performance on the bicycle. Although there is no agreement on the characteristics of the optimal pedalling technique, any desired pedalling technique can be learned and trained by cycling ergometry.

A training session in the lab is based on the analysis and feedback of the distribution of the applied pedal forces during each revolution of the crank. In many cases, the torque variation is enough for the training of pedalling techniques and the correction of force asymmetry between the legs. See some examples in (Böhm, Siebert, & Walsh, 2008; Faria, 2009).

In cycling, the pedalling technique is determined mostly by variations in the torque applied to the pedals during crank rotation. The torque at the pedal is the sum of the torques needed to overcome all resistive forces and the torque required for any changes of angular momentum of the ergometer flywheel. This last torque is proportional to the angular acceleration of the crank. For an ergometer with almost constant brake torque, one may assume that variations in the pedal force can be extracted from the pedal motion alone. The key problem is to reliably estimate the angular crank acceleration from noisy motion or video data.

My contributions are the following:

- A novel method to compute these torque variations from the pedalling motion using video from a plain commercial camera, an ergometer and an automatic correction of the marker position. The results have shown that an expensive optical device is not necessary to estimate the torque variation applied to the cranks.
- The validation of the method using motion capturing and video, direct pedal torque measurement, ergometer brake forces ranging between 100 and 250 N, and cadences of 60, 80, and 100 rpm.

1.2 My publications

Parts of this thesis have been published in my following publications:

- Quintana-Duque, J. C. (2011). Kompression menschlicher Bewegungsdaten: Eine Übersicht. In D. Link & J. Wiemeyer (Eds.), *Sportinformatik trifft Sporttechnologie, 8. Symposium der dvs-Sektion Sportinformatik* (Vol. 217, pp. 157–162). Feldhaus Edition Czwalina.
- Quintana-Duque, J. C., & Saupe, D. (2012). Evidence of chaos in pedaling motion using nonlinear methods. In *Proceedings of World Congress of Performance Analysis of Sport IX* (pp. 30-1–11). Routledge.
- Quintana-Duque, J. C. (2012). Non-linear dynamic invariants based on embedding reconstruction of systems for pedaling motion. In *Sportinformatik 2012. 9. Symposium der dvs-Sektion Sportinformatik.* (pp. 28–38). Konstanz, Germany.

- Quintana-Duque, J. C., Saupe, D., & Vieten, M. (2014). Parametrization of cyclic motion and transversal sections. In *Sportinformatik X* (pp. 111–117).
- Quintana-Duque, J. C., & Saupe, D. (2015). Phase registration improves classification and clustering of cycles based on self-organizing maps. In *Proceedings of the 2nd International Workshop on Sensor-based Activity Recognition and Interaction* (pp. 5:1–5:6). New York: ACM.
- Quintana-Duque, J. C., Dahmen, T., & Saupe, D. (2015). Estimation of pedal torque variation from pedal motion in cycling. *International Journal of Computer Science in Sport*, 14(1), 34–50.

1.3 Thesis outline

This dissertation is organized as follows:

Chapter 2 introduces the concepts of motion capturing, movement variability, movement stability and redundancy. These concepts are often mixed and have been described using more than one mathematical tool. They are clarified to have an idea about that which is related the content of this dissertation.

Chapter 3 reviews some background on non-linear methods.

Chapter 4 covers the surrogate data hypothesis testing for the analysis of the evidence of chaos in cyclical motion data. Here, it is shown that motion in cycling ergometry does not strongly evidence chaotic properties.

Chapter 5 describes the new method of parametrization of cyclic motion called eDBA and the analysis of cycles based on transversal sections. Further, the improvement of the classification and clustering of cycles based on SOM and phase registration is shown.

Chapter 6 shows the novel approach for the estimation of torque variation from pedal motion in cycling. The pedalling technique can be trained using a plain commercial camera, an ergometer, and our method.

Lastly, Chapter 7 concludes this dissertation by summarizing its contributions, discussing its limitations, and outlining future research.

Chapter 2

Concepts related to human movement

The human body is an articulated multi-joint system moving through the three-dimensional world and offers a large number of combinations of movements at different joints. Human movements are not simply preprogrammed (Rosenhahn, Klette, & Metaxas, 2007) and are the result of the interaction between the brain, the biomechanical system, and the environment. Based on environmental, biomechanical and morphological constraints, the human body must generate both stable and flexible motion output in response to changing intentions and dynamical environmental constraints.

Until now, the available technologies have had access neither to the intention behind the movement nor to the high-dimensional body signals controlling the biomechanical and neuromuscular systems. Thus, a deep understanding of the production of movement is still not possible. However, features of human movements have been related to mathematical concepts in order to describe the motor behaviour of the human body.

Many publications have been devoted to the analysis of multiple repetitions of a task based on two fundamental concepts: variability and stability. However, these concepts are often mixed and described using more than one mathematical tool. These problems have lessened the possibility of following related results in the literature and of reaching a definite conclusion.

This chapter clarifies some concepts related with human movement which are necessary to have an idea about what is related to the content of this dissertation. This chapter is divided into three sections: Section 2.1 introduces motion capturing, motion data, and its properties. Section 2.2 presents the concepts of variation, stability, end-point and coordinative variability. Finally, Section 2.3 describes the reasons to analyse human motion with tools from dynamical system theory.

2.1 Motion capturing and human motion data

Motion capture is the process of recording the movement of objects or people and is useful to display, understand, and quantify human motion. Motion capture technologies are becoming more popular every day because the range of methods and applications for processing motion information is increasing and the costs of those technologies are getting even lower. Motion data have been used successfully in recent years in many different fields, such as, 3D special effects in movies, motion analysis, modeling of human movement, motion simulation, and biofeedback to improve performance in sport.

Traditional marker-based optical motion capture systems (e.g. LUKOtronic¹) and modern

¹LUKOtronic is a multicamera system with infrared markers. For more details, see <http://www.steinbichler.de/en/>.

markerless motion capture technologies (e.g. Kinect sensor²) are of many kinds and can be classified as optical, magnetic, mechanical, or acoustic. With respect to the nature of the sensor, they can be active, i.e. the sensors are measuring and transmitting the position and/or orientation or passive, i.e. the sensors are detected by an external component (Preda, Jovanova, Arsov, & Preeux, 2007). See (Moeslund, Hilton, & Kruger, 2006) for a survey of advances in vision-based human motion capture. In addition to motion capture systems, a variety of technologies has been used to complete the information about motion, e.g. the inertial measurement unit (IMU), which is an electronic device that measures velocity, orientation, and gravitational forces, using a combination of accelerometers, gyroscopes, and magnetometers.

Human motion capture techniques depend on the intended degree of abstraction imposed on the virtual counterpart of the human actor. In order to capture a full skeleton motion sequence, a minimal set of 40–50 markers are required with sampling rates commonly between 120 and 240 Hz. The high sampling rate is necessary to capture the high-frequency details created due to environmental contacts such as foot strikes (Arikan, 2006). In some cases, capturing subtle human movements necessitates as many as over 300 markers (Liu & McMillan, 2006).

Motion data is defined as a set of samples which represent the spatial position (and in some devices the orientation) of all attached markers of the pose of a real object at uniformly spaced instants of time. Human motion data is often acquired from a single actor, recording the marker position over time of strategically placed markers on any anatomical landmark, e.g. the centers of the rotation joints or the extremes of limb segments. A sequence of movements is represented in motion capture data as an ordered sequence of frames.

The number of independent quantities which must be specified in order to define uniquely the position and orientation in space of any (bio)mechanical system is called the number of degrees of freedom (DOF) (Ivancevic & Djukic, 2009). A marker is represented with from 1 to 6 DOF. Thus, a full representation contains the coordinates of the spatial position (forward/backward, up/down, left/right) and the rotation about three perpendicular axes (pitch, yaw, and roll).

For many applications, motion data is captured using a large number of markers with a high sample rate, which needs a lot of disk space to be recorded and large bandwidth to be transmitted. However, bandwidth and storage space are limited resources. A direct use of this data is impractical, time consuming, and its efficient use for indexing and retrieval is impossible. In order to make the information compact and quickly accessible, data compression is necessary. A survey of compression techniques for motion data is presented in my work in (Quintana-Duque, 2011).

2.1.1 Properties of human motion data

Reconstructing 3D human motion data in a virtual environment raises the question about how the information should be analysed. The analysis of human motion data can be classified according to the size of the groups of DOF in three hierarchical levels. See the examples in Fig. 2.1. The lowest level of the hierarchy consists of each joint and its data, forming an individual feature group. The joint positions are often described using a root marker as the reference (e.g. the white circle in Fig. 2.1a). Traditionally in biomechanics, the data from joints (e.g. displacement, rotation angle, velocity, acceleration, force, etc.) are presented as a function of time. The middle level of the hierarchy consists of sets of joints grouped as limbs (see Fig. 2.1b) from which features, such as

²Kinect sensor is a horizontal bar connected to a small base with a motorized pivot and is designed to be positioned lengthwise above or below the video display. For more information, see <https://www.microsoft.com/en-us/kinectforwindows/develop/>.

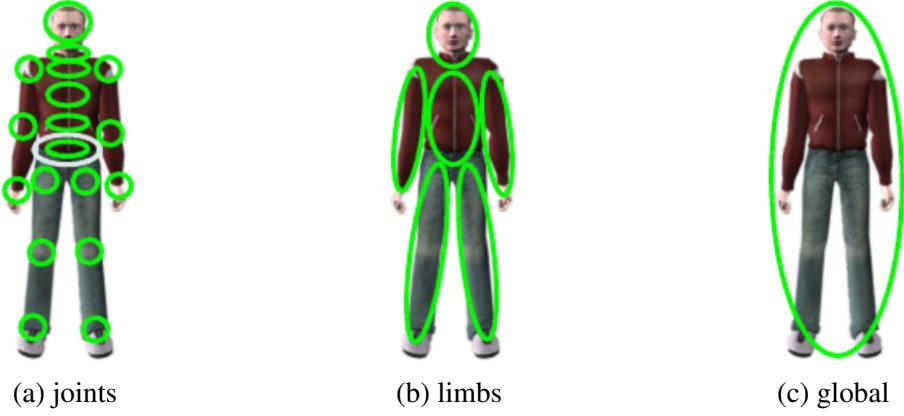


Figure 2.1: Examples of possible feature groups for motion analysis. Each feature group is illustrated as a green circle capturing dependencies between joints at different scales. The white circle is the root marker used as the reference for the other markers. Figure modified from (Ren, 2006).

the moment of inertia and the coordination between them can be analysed. The highest level of the hierarchy consists of all joints combined into one feature group (see Fig. 2.1c).

Taking the hierarchical levels into account, human MoCap has the following properties that are useful for the purpose of the analysis of the data:

- Although the absolute coordinates of the markers on the body can be highly variant, their coordinates relative to the root (e.g. the white circle in Fig. 2.1a) typically exhibit repetitive patterns.
- Positions and orientations of human body parts vary smoothly and continuously over time. When viewed at an appropriately small time scale, they can therefore be modeled by splines, wavelets, or even local linear dynamical systems (Beaudoin, Poulin, & van de Panne, 2007).
- Temporal coherence and joint correlation are present in MoCap data because all DOF are sampled simultaneously from the same continuous physical motion and because of the strong correlations among temporally adjacent frames (Shimada & Uehara, 2000; Li, Okuda, & Takahashi, 2008; Liu & McMillan, 2006; Lin & McCool, 2007).
- Humans move in diverse, but also highly structured ways. Certain motions, such as running or walking, have a repetitive structure, others represent cognitive routines of various levels of complexity, e.g. crossing the street by checking for cars on the way (Sminchisescu, 2008).
- Human movements are highly coordinated among body moving parts by nature. For example, while walking, the body balances itself with its arms and legs almost periodically.
- Although many of our daily movements are cyclical, we never repeat the same movement. For example, as a person walks, his or her footprints never repeat exactly, reflecting variability from step to step in a continuous cycle of movement (Harbourne & Stergiou, 2009; Stergiou, Harbourne, & Cavanaugh, 2006).
- In the context of the animation and visualization of human motion, some DOF are more important perceptually than others. Therefore, it is possible to remove redundancy between DOF using dimensionality reduction techniques and also to omit some uninfluential DOF without significantly affecting the perceptual accuracy (Lin & McCool, 2007).

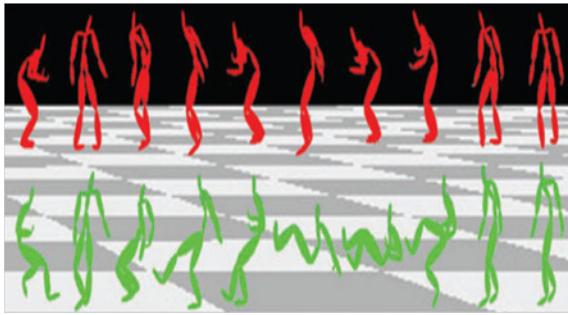


Figure 2.2: Two different motion sequences (red and green) with high local similarity at two arms compared in each column. Figure modified from (Gu et al., 2009).

- For many cyclical motions, the kinematics (angles, distances, and speeds) of the limbs are highly consistent. The kinetics (muscle forces and torque), however, can be quite variable in creating the exact same kinematics or pattern of motion. See an example of gait motion in (Winter, 1984).
- In a large database, there are many different copies of similar looking motions, as is shown in Fig. 2.2. If we only focus on the motions of a certain body part, the probability of finding similar local motion patterns in different motion sequences is significantly increased (Gu, Peng, & Deng, 2009).

2.2 Movement variability, stability and redundancy

One of the main characteristics of human motion is that although many of our daily movements are repeated almost cyclically under same environmental constraints (e.g. as in gait and cycling motion), we never repeat the same movement twice. This characteristic is attributed to movement redundancy. Movement redundancy refers to the biomechanical DOF that exceed the minimum number of DOF required to accomplish a given motor task, resulting in the existence of multiple motor solutions (Bernstein, 1996).

Due to movement redundancy, some motion behaviours which appear to be highly stable are performed in variable ways. This is especially evident when we observe elite sports players or musicians performing (Stergiou et al., 2006). Not only is their performance more consistent than that of less capable individuals, but they also seem to have developed a large number of ways of performing a task (Stergiou & Decker, 2011) despite several years of intensive training of specific movements (Davidson, Glazier, Araujo, & Bartlett, 2003).

The way that the human motor control system manages movement redundancy is described by the concepts of movement variability and movement stability. Both concepts describe different properties within the motion control process. Variability refers to the ability of the motor system to reliably perform a task under a variety of constraints. Stability refers to the dynamic ability to offset an external perturbation and how a motor behaviour emerges in time, often across a range of time scales (Stergiou & Decker, 2011). It has been studied with non-linear dynamic stability analyses, for example, in the form of Lyapunov exponents³ (Dingwell & Cusumano, 2000). However, the

³Lyapunov exponents quantify the average exponential rate of divergence of neighbouring trajectories in state space, and thus provide a direct measure of the sensitivity of the system to infinitesimal, i.e. ‘local’ perturbations (Dingwell & Cusumano, 2000). This dynamic invariant is explained more in the next chapter.

concepts of movement variability and movement stability are often mixed in the literature.

Traditionally, the methodology for the study of individual human motion behaviour (known in the literature as single-subject analysis) consists of making repeated measurements of responses (trials) over a period of time and the average of several motion curves. The deviation around the mean of a dependent variable, which is measured over repeated trials, is considered random noise. Thus, movement variability has been equated with noise, is considered detrimental to movement performance, and is typically eliminated for data analysis.

Recent studies (Hausdorff, Peng, Ladin, Wei, & Goldberger, 1995; Slifkin & Newell, 1998; Glass, 2001) have found evidence that the amount and the nature of the alleged noise contain important information useful for characterizing mature motor skills and healthy states. Methods for the detection of chaos⁴ applied to motion data have revealed a chaotic structure and fractal properties (Hausdorff, 2007) in healthy subjects, suggesting that the nature of the alleged noise is not entirely random. Thus, the idea that variability is noise is not only a default proposition based on untested assumptions but is also erroneous (Oullier, Marin, Stoffregen, Bootsma, & Bardy, 2006). The inherent and unavoidable noisiness of the motor system seems related to the adaptability of the patterns of coordination needed to secure stable outcomes in dynamic performance (K. Newell, Deutsch, Sosnoff, & Mayer-Kress, 2006).

Much research on motion focuses now on the amount of motion variability (Bartlett, Wheat, & Robins, 2007; Glass & Mackey, 1988; K. M. Newell & Corcos, 1993; Davids et al., 2006) and movement (in)stability (Lipsitz, 2002; Stergiou & Decker, 2011) that should be considered optimal for a specific motion. Deviations from the optimal values (too much or too little variability) result in motion systems that are less adaptable to perturbations. Lack or excess of variability and stability are now understood as deficiencies in motor skills. For example, in the absence of skillfulness, a performer essentially becomes rigid and invariant, thus freezing many DOF. In contrast, a skilled performer exploits many DOF, making the performance more graceful and adaptable (Vereijken, van Emmerik, Whiting, & Newell, 1992).

Understanding movement variability and stability can enhance applications related to motion as varied as motion synthesis for special effects in movies (Barbivc et al., 2004), motion classification (Glazier, Wheat, Pease, & Bartlett, 2006), sport training (Bini, Diefenthäler, & Mota, 2010), monitoring of human motion in real time (Pollard, Heiderscheit, van Emmerik, & Hamill, 2005), therapy evaluation (Hamill, van Emmerik, Heiderscheit, & Li, 1999), diagnosis of disease severity (Hamill, Palmer, & Van Emmerik, 2012), medication utility, and fall risk (Davids et al., 2003).

2.2.1 Coordinative variability and end-point variability

Recently, there has been a change in the way that human motion variability is evaluated. The idea that variability decreases with skill acquisition in one context (motor learning paradigm) and increases with skill acquisition in another context (the development of a behavioural repertoire) is readily explained by the way in which variability is measured, (Stergiou & Decker, 2011).

In the literature, two types of variability can be distinguished: end-point variability and coordinative variability. The traditional view of variability is based on the concept of end-point variability. From this perspective, the variability of a measurement of movement (e.g. stride length, stride time, etc.) should be less in a healthy individual and expert performers, and greater in a less healthy individual and novices (Hamill et al., 2012). Coordinative variability is the variability of the interaction

⁴Chaos is defined as a non-periodic long-term behaviour in a deterministic system that exhibits a sensitivity to initial conditions. For more details, see Section 3.4.

between segments or joints and is viewed as more functional, since a consistent outcome can be achieved by different patterns of joint relations owing to the dynamics of the joint biomechanical degrees of freedom (Davids et al., 2003).

As skilled motor performance is often characterized by a low variability of the outcome measures, it follows that skilled motor performance is also characterized by highly consistent patterns of movement. Therefore, movement variability has been considered as a problem in the sensorimotor system that should be minimized or eliminated. However, under different task constraints, such as those encountered in tasks requiring the use of multiple biomechanical degrees of freedom, it has been found that the variability of outcome measures does not necessarily relate to the variability of the accompanying patterns of movement (Glazier et al., 2006).

Coordinative variability has the opposite interpretation to end-point variability, but these two concepts of variability must be integrated into any functional movement analysis. In gait dynamics, the goal is not a discrete spatial location, but the maintenance of segmental relations (co-ordination) over many cycles that define the locomotor pattern itself (Hamill et al., 2012).

The best ways to quantify end-point and coordinative variability are still under debate, as well as the amount of end-point and coordination variability that should be considered optimal for specific motions. The end-point variability is usually described by the variance, the standard deviation, and the coefficient of variation. The analysis of coordinative variability is usually based on the assumption that the two oscillating segments under scrutiny are of a one to one frequency ratio and exhibit a sinusoidal time history. However, the body segments used in cyclic motion do not always meet these assumptions. See (Hamill et al., 2012; Stergiou et al., 2006) for techniques to describe coordination in motion, e.g. the discrete relative phase (DRP) and continuous relative phase (CRP).

2.3 Chaos and human motion as a dynamical system

Chaos theory has inspired a new set of tools useful for time series analysis and provides a new language to formulate time series problems and to find their solutions (Schreiber, 1999). Chaos provides an alluring explanation for the erratic behaviour of time series because it can be exhibited by systems that are both low-dimensional and deterministic (Theiler, 1995). Especially alluring is the hypothesis that the irregular fluctuations of the human motion data reflect an underlying chaotic determinism.

Motion variability is an essential feature of the human motor system attributed to its ability to coordinate many different physiological systems over many different timescales. The introduction of non-linear dynamics and chaos theory to the study of human motion has revealed a chaotic structure in its variation. For this reason, the human motor control behaviour has been related to chaotic dynamical systems. However, we must bear in mind that deterministic chaos is only one possible source of irregular behaviour in real-life systems. The other source is the ever present noise, which even if combined with linear dynamics can also produce irregularly appearing behaviour (Perc, 2005).

Some dynamical systems are chaotic everywhere, but in many cases, the chaotic behaviour is found only in a subset of the state space. A chaotic system is characterized by its sensitive dependence on initial conditions. The evolution of the output in time of such systems can be completely different even for very close initial conditions, making the predictability of these systems tough and even impossible in the long term. Furthermore, since experimental initial conditions are never known entirely, these systems are intrinsically unpredictable. For more details, see Section 3.4.

Chaotic systems can have very irregular output with purely deterministic dynamics and may

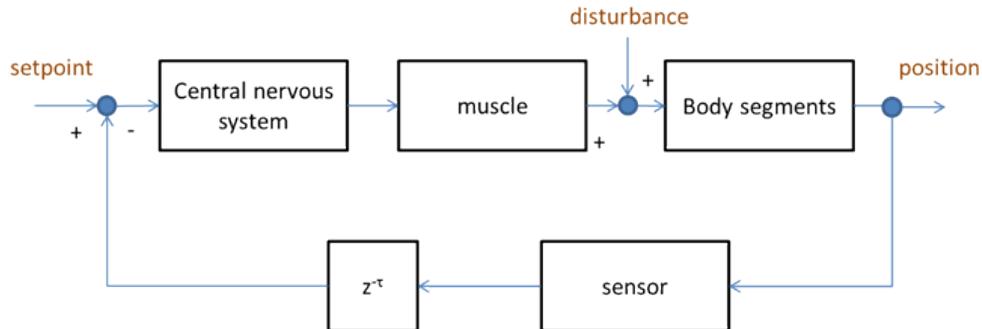


Figure 2.3: A model of a human motor system. The human motor system can be considered as a feedback controller with external perturbations. The desired posture is compared with the actual posture and the error is used to control and stabilize the musculoskeletal system. Additional loops with large time delays can be included to expand the conceptual scheme with the ‘next-cycle’ adaptation phenomenon. Adapted from (van der Kooij et al., 2008).

drive fluctuations in system output that apparently seem random. In a similar way, motor behaviour is sensitive to the initial conditions of the motion activity. For example, the way that we perform a movement can change between consecutive days or hours although we perform the same routine. Thus, human motion cannot be predicted over the long term with a simple model.

The ‘next-cycle’ adaptation explains the modifications of motion behaviour of current acts based on antecedent ones, e.g. the desired posture is compared with the actual posture and the error is used to control and stabilize the musculoskeletal system. It has been hypothesized that sensory feedback plays a role in adjusting step-to-step limb trajectories to maintain balance during locomotion (Nashner, 1980), and in smoothing unintended irregularities that occur during unperturbed movements (Gandevia & Burke, 1992; Glass, Beuter, & Larocque, 1988). The latter is used to model the human motor system as a dynamical system, e.g. the human motor system can be considered as a feedback controller with external perturbations (e.g. see Fig. 2.3) with a linkage system (body segments), actuators (muscles), sensors (e.g. tactile sensors, visual and vestibular system), controller (Central Nervous System) and τ representing the time delay caused by transport and processing in the nervous system (van der Kooij, Koopman, & van der Helm, 2008). If the system is chaotic, one is able to produce an infinite number of desired dynamical behaviours (either periodic or aperiodic) using the same chaotic system, with the help of only tiny perturbations chosen properly (Boccaletti, Grebogi, Lai, Mancini, & Maza, 2000).

The challenge in modeling human motion is to explain coordinated movements as the orderly products of a hybrid biological organization involving a very large number of different subsystems (e.g. vascular, neural, muscular, skeletal) (Beek, Peper, & Stegeman, 1995). However, the assumed model of human motion control system influences the interpretation of the results using tools of dynamical systems. The next chapter provides an overview of non-linear methods for the analysis of time series and their application to motion data recorded from cycling ergometry.

Chapter 3

Review of non-linear methods for time series

Human motion is the product of the organization of a vast number of different subsystems, e.g. vascular, neural, muscular, and skeletal subsystems. The collection of all those subsystems is called motor system. The knowledge about the motor system is limited because the unfeasibility to know all variables of all subsystems. Some variables of those subsystems are even technically impossible to measure.

For example, a part of the human motor system is the central nervous system (CNS) which coordinates the activity of all parts of the body. The CNS contains about 10^{11} nerve cells (or neurons). With each neuron having an average of 10^4 connections, a total of 10^{15} connections are present. Given this huge complexity, any model attempting to understand the function of the CNS is only very simple (van der Kooij et al., 2008). However, the complexity of the human motor system may be reflected by the variables that can be observed from the motion.

Even if the model of a system is unknown or too complex, it is possible to analyse the time series of experimental measures in order to understand the system that generated them. Recently, many efforts have been done to explain the complex behaviour of the human motor system using non-linear features instead of the usual statistical models. Non-linear time series analysis is the study of time series with computational techniques sensitive to non-linearity in the data, under the assumption that the time series are noisy measurements of particular solutions of a deterministic dynamical system.

A deterministic dynamical system is determined by a set of functions which describes the time dependence of a point in state space. State space refers to the space whose axes in a coordinate system are the state variables, wherein the state of the corresponding deterministic dynamical system is represented as a vector. The path traced by the system state as it evolves over time is called trajectory or orbit. If the system is chaotic, the trajectory converges after a transient time to an attractor which is a subset of state space to which a dynamic system evolves.

The analysis of human motion with non-linear methods assumes an underlying dynamical system which is not known explicitly. Even if it were, it would typically contain a myriad of variables defying any mathematical or even numerical analysis. However, since the discovery of Takens (1981) of a topologically faithful reconstruction of the state space of a dynamical system in a lower dimensional space, the so-called embedding space, it became possible to estimate and to study the dynamics from noisy measurements of only a few state variables of the underlying motion.

The reconstruction theorem was obtained independently by Aeyels (1981) and Takens (1981) but it is usually called Takens' theorem. The reconstruction theorem gives the conditions under which a chaotic dynamical system can be reconstructed from a sequence of observations of the state of a dynamical system $x[n]$, i.e. there is a map between the original state space and a ‘reconstructed’

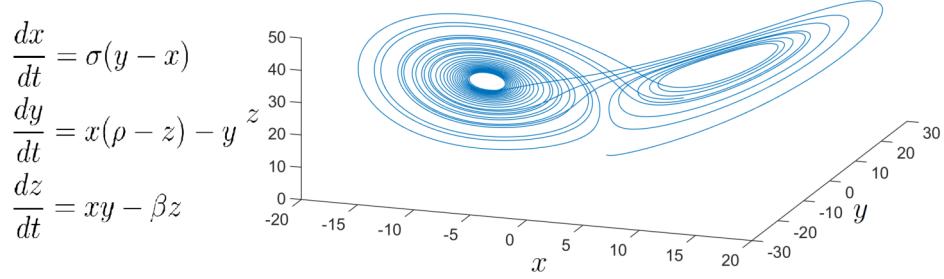


Figure 3.1: Example of a continuous dynamical system: The Lorenz attractor with parameters $\sigma = 10$, $\rho = 28$ and $\beta = 8/3$. A dynamical system is represented as a set of functions which describe the time dependence of the variables in state space. After a transient time, the system converges to an attractor, which is a subset of the state space to which a dynamic system evolves. The states of the system trace a path as they evolve over time, which is called a trajectory or orbit. In this example, the trajectory is clearly visible (blue line) and is constrained to a subset of the state space.

state space that preserves the dynamical properties. The underlying idea of Takens' theorem is that all the variables of a dynamical system influence one another. Thus, the observations $x[n]$ of even one single variable are the result of the combination of the influences of all other variables. Takens' theorem guarantees that the observations $x[n]$ of one single variable will be enough to reconstruct the attractor of the dynamical system in an embedding space with dimension d_e if the observations are infinite and without noise. In practice, a sequence of observations at a sufficiently high sampling rate is enough to approximate the attractor when the data is sufficiently large and contains a low level of noise.

Dynamic invariants characterize the state space of the corresponding unknown dynamical system and its reconstruction in embedding space. Only in rare cases, when a very simple dynamical system is given explicitly, is it possible to rigorously compute features about the dynamics. Thus, in practice only a numerical approximation is feasible. However, the dynamic invariants calculated from a time series are meaningful only when there is indeed a deterministic dynamical system with a non-linearity underlying the time series.

The aim of this chapter is to review some non-linear methods for time series. Section 3.1 describes the reconstruction theorem (i.e. Takens' theorem) for dynamical systems. Section 3.2 reviews the concept of an embedding space and the state-of-the-art techniques for the calculation of the embedding parameters for an uniform embedding. Section 3.3 describes the calculation of three dynamic invariants from embedding space. Finally, Section 3.4 defines chaos and describes which dynamical systems can exhibit chaos.

3.1 Dynamical systems and the reconstruction theorem

An autonomous dynamical system consists of a state space and its dynamics¹. The state space is the collection of all possible world-states of the system in question. Each world-state represents a complete snapshot of the system at some moment in time, and is specified by a vector $\mathbf{x} \in \mathbb{R}^Q$ with

¹The term dynamics has two very different meanings in dynamic systems and biomechanics. In dynamic systems theory, dynamics refers to the time evolution of a system at any level of description, whereas in biomechanics, dynamics refers to the forces and torques that physically cause the motion (Beek et al., 1995).

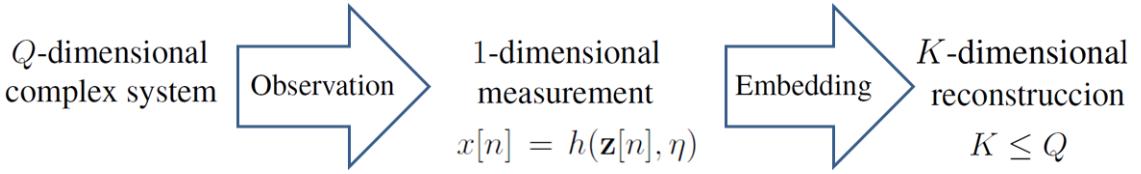


Figure 3.2: Embedding reconstruction of a multidimensional attractor. The $h(\cdot)$ is the observation function and η is an observational noise parameter. Topological properties of Q -dimensional complex system are preserved in K -dimensional embedding space by the delay embedding method.

Q being the number of state variables. The dynamics can be described either by an Q -dimensional map (discrete dynamics) or by a system of Q first-order differential equations (continuous dynamics). Notice here that sampling the continuous system with sampling time Δt gets you a discrete time system, but this is not necessarily a 1-D map.

In continuous time, we don't specify the evolution of the world-state directly, but rather the rates of change of the components of the state $\frac{dx}{dt} = F(x, t)$. In the discrete-time case, the dynamics is an equation that transforms one point in the state space (world-state), representing the state of the system now, into another point (a second world-state), representing the state of the system one time unit later. This equation is given by a function $x[n + 1] = f(x[n])$.

For both continuous and discrete time it is valid that once an initial world-state is chosen, the dynamics determine the world-state at all future times, forming a trajectory in the state space. Given a period of time long enough, the trajectory may eventually settle inside a bounded set, called a stable attractor, which is a portion of the state space to which trajectories converge asymptotically. The path traced by the system's states as they evolve over time is called a trajectory or orbit. See Fig. 3.1 for an example of a continuous system, a dynamical system known as the Lorenz attractor.

Since the discovery of Takens (1981) of a topologically faithful reconstruction of the state space of dynamical systems in a lower dimensional, so-called embedding space, it became possible to estimate features from the dynamics only from noisy measurements of some function in the original (unknown) state space, i.e. based on the embedding reconstruction a time series can be analyzed from the point of view of nonlinear dynamics. For a complete mathematical defintion and explanation of the reconstruction theorem, see (Broer & Takens, 2010; Broer, Takens, & Hasselblatt, 2010).

In the following, we focus only on the discrete time case and on uniform sampling. Let $z \in \mathfrak{M} \subseteq \mathbb{R}^S$ be a S -dimensional state vector, and let \mathfrak{M} be the attractor to which the dynamics evolve. Usually, it is not possible to directly observe all components of the state vector $z \in \mathbb{R}^S$ of a high-dimensional dynamical system but it is possible to observe the system obtaining 1D-measurements $x = h(z, \eta)$ providing a projection of the state space, where $h : \mathfrak{M} \rightarrow \mathbb{R}$ is an observation function and η is an observational noise parameter. The observation function $h(\cdot)$ provides us with a way to measure the current state of the system $h(z[n], \eta)$. Since $h(\cdot)$ gives us only a scalar value, it cannot offer a complete description of the system. But, observing $x[n] = h(z[n], \eta)$ at many successive uniformly separated times will.

Given a sequence of state space points $z[n], n = 0, 1, \dots$, from the dynamics of the underlying system, the sequence of discrete observations $x[n] = h(z[n], \eta)$ from an uniform sampling can be used to create an K -dimensional embedding space for the state space reconstruction. If the sequence of observations is long enough and contains little noise, the topological properties of an S -dimensional chaotic system can be reconstructed in the embedding space. See Fig. 3.2 for an explanation.

The accuracy of the numerical estimation of dynamic features from experimental data depends on the quantity and quality of the time series as well as on the quality of the reconstruction of the complexity of the dynamical system using an embedding space. Embedding allows us to study systems for which the state space variables and the governing equations are unknown. The basic idea behind embedding is that the past and future of a time series contain information about unobserved state variables that can be used to define a state at the present time (Casdagli, Eubank, Farmer, & Gibson, 1991). The past and future information contained in the time series can be encapsulated in the vector defined by

$$\mathbf{v}[n] = (x[n + \tau'_{m_f}], \dots, x[n + \tau'_2], x[n + \tau'_1], x[n], x[n - \tau_1], x[n - \tau_2], \dots, x[n - \tau_{m_p}]). \quad (3.1)$$

The dimension of this delay vector is $d = m_f + m_p + 1$, which preserves the properties of the dynamical system when the dimension d is large enough, i.e. $d \geq d_b + 1$, where d_b is the box-counting dimension of the attractor of the system. The number of samples taken from the past is m_p , and the number from the future is m_f . If $m_f = 0$ then the reconstruction is predictive; otherwise it is mixed. The time separation between coordinates can be uniform or non-uniform and is given by $\tau_1 < \tau_2 < \dots < \tau_{m_p}$ and $\tau'_1 < \tau'_2 < \dots < \tau_{m_f}$.

The delay embedding reconstruction (i.e. $m_f = 0$), for discrete time series and uniform sampling, is often used in the literature. The uniform delay embedding is the classical method and determined by two integer parameters: The time delay τ and the embedding dimension d_e . This embedding is defined by Eq. 3.2,

$$\mathbf{v}[n] = (x[n], x[n - \tau], x[n - 2\tau], \dots, x[n - (d_e - 1)\tau]). \quad (3.2)$$

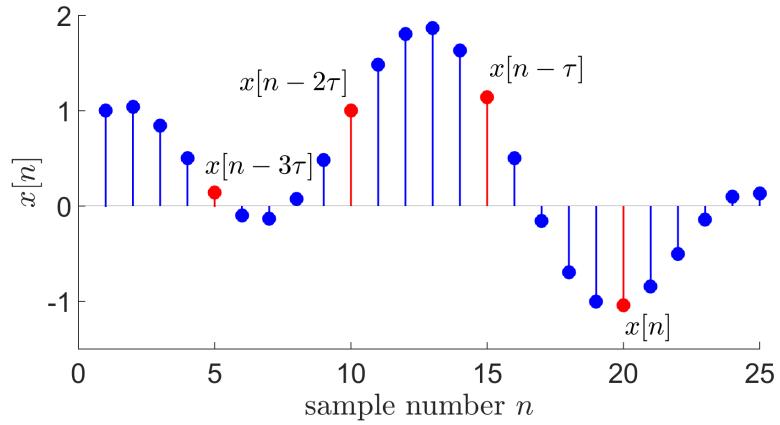
The non-uniform delay embedding with not equidistant consecutive delays $\tau_1 < \tau_2 < \dots < \tau_{d_e-1}$ is defined by Eq. 3.3,

$$\mathbf{v}[n] = (x[n], x[n - \tau_1], x[n - \tau_2], \dots, x[n - \tau_{d_e-1}]), \quad (3.3)$$

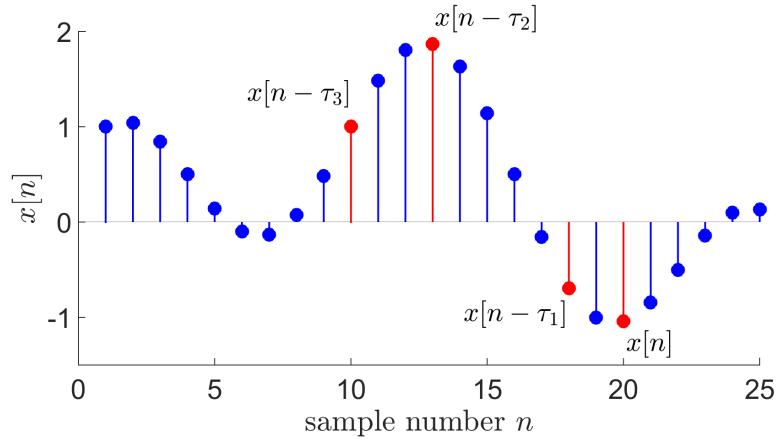
and is considered as an alternative to uniform embedding techniques. See Fig. 3.3a and Fig. 3.3b for examples of each embedding.

Takens' theorem gives a rigorous justification for state space reconstruction. However, it gives little guidance on reconstructing state spaces from a finite amount of noisy data. Both noise and finiteness will prevent us from having access to infinitesimal length scales, so that the structure we want to exploit should persist up to the largest possible length scales (Schreiber, 1999). For example, the measurements in Takens' proof are arbitrarily precise, resulting in arbitrarily precise states. This makes the specific value of the lag time τ arbitrary so that any reconstruction is as good as any other. However in practice, the presence of noise in the data blurs states and makes picking a good lag time critical (Casdagli et al., 1991).

The original version of Takens' theorem assumes that the underlying system is autonomous and noise free. Unfortunately this is not the case for many real systems. However, Takens' theorem has been extended to non-autonomous systems, where f is also a function of some other variable $\mathbf{y}[n]$, i.e. $\mathbf{x}[n + 1] = f(\mathbf{x}[n], \mathbf{y}[n])$, where $\mathbf{y}[n]$ can be: (1) a deterministic system, so that $\mathbf{y}[n + 1] = \kappa(\mathbf{y}[n])$ for some mapping κ , (2) arbitrary (3) or stochastic. See (Stark, Broomhead, Davies, & Huke, 2003; Casdagli, 1992; Monroig, 2009) for the corresponding theories.



(a) Uniform embedding with $\tau = 5$.



(b) Non-uniform embedding with $\tau_1 = 2$, $\tau_2 = 7$ and $\tau_3 = 10$.

Figure 3.3: Examples of four-dimensional embeddings. Top: Uniform embedding is defined with embedding delays $\tau < 2\tau < 3\tau$ as $\mathbf{v}[n] = (x[n], x[n - \tau], x[n - 2\tau], x[n - 3\tau])$. Bottom: Non-uniform embedding is defined with embedding delays $\tau_1 < \tau_2 < \tau_3$ as $\mathbf{v}[n] = (x[n], x[n - \tau_1], x[n - \tau_2], x[n - \tau_3])$.

3.2 Embedding space

Essentially, three methods for reconstructing the state space from scalar time series are available: delay coordinates (Packard, Crutchfield, Farmer, & Shaw, 1980), derivative coordinates (Parlitz, 1998) and principal coordinates (Broomhead & King, 1986). All three methods are related by linear transformations (Gibson, Farmer, Casdagli, & Eubank, 1992). However, the method of delays is widely used to reconstruct a state space from an observed time series.

The reconstruction using delay coordinates is based on the embedding dimension d_e and the embedding delays as defined in Eq. 3.2 and Eq. 3.3. The embedding preserves the properties of the dynamical system that generated the time series under study when the embedding dimension d_e is large enough, i.e. $d_e \geq 2d_b + 1$, where d_b is the box-counting dimension² of the attractor of the dynamical system (Sauer, Yorke, & Casdagli, 1991). However, in most of the cases we don't know the box-counting dimension d_b of the attractor of a real dynamical system. Then, a suitable embedding dimension d_e must be estimated from the time series $x[n]$ with $n = 1, \dots, N$.

3.2.1 Calculation of embedding parameters from time series

From the theoretical point of view, the selection of the embedding dimension d_e larger than $2d_b + 1$ should not be a problem, but from a practical point of view, it is counterproductive (Signorini & Ferrario, 2011; Kantz & Schreiber, 2004). A precise knowledge of d_e is desirable since a large value of d_e will add redundancy. The additional dimensions due to redundancy are not populated by the dynamics of the system but entirely by the noise present in the data. Further, the redundancy increases the number of operations, e.g. to calculate dynamic invariants and non-linear predictions, which are proportional to d_e . Furthermore, for a d_e too large, practical problems occur due to the limited number of points, which constitute thinner sets in \mathbb{R}^{d_e} when d_e is increased. Therefore, to provide a proper reconstruction, d_e should be chosen as small as possible.

The selection of consecutive delays $\tau_1, \tau_2, \dots, \tau_{d_e-1}$ with $\tau_1 \leq \tau_2 \leq \dots \leq \tau_{d_e-1}$, although irrelevant in Takens' formal derivation, becomes important for experimental time series due to their finite length and measurement precision (Uzal, Grinblat, & Verdes, 2011). Thus, different time delays lead to reconstructions of the attractor that are diffeomorphically equivalent but geometrically different (Parlitz, 1998). When the time delay τ_{d_e-1} is too small, the coordinates of each reconstructed state do not significantly differ from other and therefore the reconstructed trajectories are too close and overlapping in the presence of noise, i.e. a possible fractal structure cannot be analysed. When the delay time τ_{d_e-1} is increased, then the attractor unfolds more and more, and its structure becomes visible. However, a too large value of τ_{d_e-1} may lead to additional folding that yields intersections of the attractor where the reconstruction is no longer one-to-one, even without noise. Additional folds in the reconstruction are unwanted because they bring states close together in the reconstructed state space that are not close in the original state space.

The object of empirical methods for the selection of the embedding delay τ is to make the components of the reconstructed vectors independent, as far as possible, yet not so far as to lose the information about the dynamic properties of the embedded time series (Ragulskis & Lukoseviciute, 2009). The choice of the time delay is usually performed using computational techniques based on the first zero of the autocorrelation function (or other specified values instead of zero, such as $1/e$ or 0.5) (Schreiber, 1999), the first minimum of the mutual information (Cao, 1997), or geometric considerations where the optimality is based on the principle of spreading out the attractor in the

²For introductory information on the various dimensions used in dynamical systems, see (Peitgen, Jürgens, & Saupe, 2006).

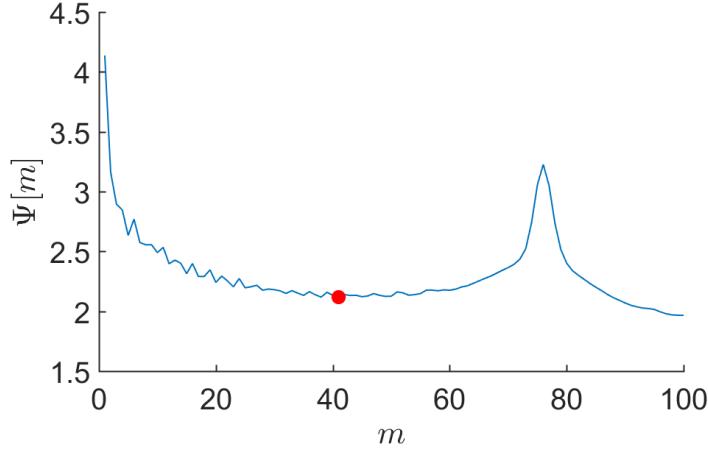


Figure 3.4: Calculation of optimal embedding lag using mutual information for the noisy positional data described in Section 4.2. The embedding delay is set to the minimum of $\Psi[m]$ of the interval between the first and the second largest peaks (red point in this example).

embedded space (Buzug & Pfister, 1992). When the data exhibits strong periodicity, a value of τ that is one-quarter of the length of the average cycle generally gives a good embedding. This time delay is approximately the same as the time of the first zero of the autocorrelation function (Small, 2005).

The algorithms used in this dissertation for the reconstruction of the state space were selected due to their robustness against noise. For the uniform embedding, the embedding parameters are calculated separately. First, the embedding delay τ is calculated, and then the embedding dimension d_e is based on τ as the parameter. These methods are presented in the following.

Calculation of the embedding lag for a uniform embedding: The mutual information method

A common criterion for the calculation of the embedding lag for a uniform embedding relies on the mutual information (Cao, 1997). This algorithm quantifies the information about a measurement at one time from a measurement taken at another time. For this reason, it is supposed that there should be a probability distribution associated with each system that regulates its observations (Signorini & Ferrario, 2011). In the context of a scalar time series, the mutual information $\Psi[m]$ can be defined by

$$\Psi[m] = \sum_{n=1}^{N-m} P(x[n], x[n+m]) \log_2 \frac{P(x[n], x[n+m])}{P(x[n])P(x[n+m])}, \quad (3.4)$$

where $P(x[n], x[n+m])$ is the probability of observing $x[n]$ and $x[n+m]$ and $P(x[n])$ is the probability of observing $x[n]$. $\Psi[m]$ is the amount of information we have about $x[n]$, by observing $x[n+m]$, and so one sets τ to be the local minimum of $\Psi[m]$ at which the observations are somewhat uncorrelated but not statistically independent (Fraser & Swinney, 1986). The primary difficulty with estimating the mutual information is that one must first estimate the probability distribution of the system state (and, moreover, of pairs of states). However, with an appropriate selection of histogram binning, the mutual information will usually provide a good guide to the appropriate choice of embedding lag (Small, 2005). See an example in Fig. 3.4.

Calculation of the embedding dimension: The false nearest neighbours method (FNN)

A classical method for the calculation of the minimal sufficient embedding dimension from a scalar time series $x[n]$, $n = 1, \dots, N$, called the false nearest neighbor (FNN), was proposed by Kennel, Brown, and Abarbanel (1992). This method tries to guess a suitable value for the embedding dimension d_e for a given embedding delay τ by successively embedding $x[n]$ in higher dimensions and looking at the consistency of the results and the unambiguity of the system state (Sauer et al., 1991).

For each embedding dimension k each embedding vector has a nearest neighbor with nearness in the sense of some distance function, e.g. the Euclidean distance. If a small embedding dimension k is not enough to unfold the original attractor in the reconstructed embedding space, false crossings of orbits of the projected attractor with itself are possible. Further, with increasing dimension, these neighbouring points are projected into neighbourhoods of other embedding points in the embedding space, i.e. the false neighbours will no longer be neighbours. However, in an appropriate embedding dimension or higher, the neighbours in the original state space are neighbours in the reconstructed embedding space. Due to the assumed smoothness of the dynamics and the preservation of its topological properties, if the current embedding with dimension k is sufficient to resolve the dynamics, an embedding with an extra dimension $k + 1$ will do it as well, i.e. the distance ratio between neighbour points in both embedding with k and $k + 1$ dimensions does not change too much. Therefore, by examining how the number of neighbours changes as a function of the dimension, an appropriate embedding can be determined.

The false nearest neighbor (FNN) algorithm is the following: Define the embedding vector with dimension k as

$$\mathbf{v}_k[n] = (x[n], x[n - \tau], x[n - 2\tau], \dots, x[n - (k - 1)\tau]), \quad (3.5)$$

and its r th nearest neighbours as $\mathbf{v}_k^r[n_r]$ with $n + k\tau < n_r \leq N$. For each point $\mathbf{v}_k[n]$ in the time series, one looks for its r th nearest neighbour $\mathbf{v}_k^r[n_r]$ in the k -dimensional space and computes the distance between these points. One repeats the same step for $k + 1$ dimensions. If the ratio of the distances in $k + 1$ and k dimensions is larger than a threshold γ , the neighbour is false. The threshold γ must be large enough to allow for the exponential divergence due to deterministic chaos. Denoting by σ the standard deviation of the data, $\|\cdot\|$ Euclidean distance, and Θ Heaviside step function, the fraction of the false neighbours is given by

$$\text{Fraction}_{\text{FNN}}(k, \gamma) = \frac{\sum_{r=1}^R \sum_{n=1}^{N-k-1} \Theta\left(\frac{\|\mathbf{v}_{k+1}[n] - \mathbf{v}_{k+1}^r[n_r]\|}{\|\mathbf{v}_k[n] - \mathbf{v}_k^r[n_r]\|} - \gamma\right) \Theta\left(\frac{\sigma}{\gamma} - \|\mathbf{v}_k[n] - \mathbf{v}_k^r[n_r]\|\right)}{\sum_{r=1}^R \sum_{n=1}^{N-m-1} \Theta\left(\frac{\sigma}{\gamma} - \|\mathbf{v}_k[n] - \mathbf{v}_k^r[n_r]\|\right)}. \quad (3.6)$$

The first step function in the numerator is equal to one if the r th neighbour is false, i.e. if the distance increases by a factor of more than γ (e.g. $\gamma = 5$) when the dimension k is increased to $k + 1$. The second step function suppresses all those pairs whose initial distance was already larger than σ/γ . Pairs whose initial distance is larger than σ/γ by definition cannot be false neighbours, since, on average, there is not enough space to depart farther than σ . These invalid candidates should not be counted, which is also reflected in the normalization (Kantz & Schreiber, 2004). Normally, one explores only the distance to the first nearest neighbour, i.e. $r = 1$, but with $r = R$ larger than one, the method is more robust against noise. The minimal dimension k is selected as the embedding dimension d_e when a second threshold β characterizing a negligible fraction of false nearest neighbours is achieved or when no changes are detected. See an example in Table 3.1.

Table 3.1: Example of selection of embedding dimension using FNN method with $\tau = 38$ and $\gamma = 5$ as parameters for the noisy positional data described in Section 4.2. With a negligible fraction of false nearest equal to 0.1, the selected embedding dimension is 6.

| τ | d | Fraction FNN |
|--------|---|--------------|
| 38 | 2 | 40.1% |
| 38 | 3 | 10.7% |
| 38 | 4 | 1.9% |
| 38 | 5 | 0.4% |
| 38 | 6 | 0% |

3.2.2 Problems and limitations

In the case of the human motor system, given the large number of interacting subsystems (e.g. vascular, neural, muscular, and skeletal subsystems), the human motion is likely to be high dimensional. The human motor system can be considered as a feedback controller with external perturbations. If the human motor system is chaotic, then this system may be able to produce a large number of desired dynamical behaviours of motion (either periodic or aperiodic) using the human body and tiny perturbations (e.g. from the brain) chosen properly. For details, see Section 2.3, and for some similarities with the theory of control of chaos, see (Boccaletti et al., 2000). Assuming that the human motion data is enough to reconstruct the dynamics of the motor system, the following problems and limitations of numerical calculation might affect the results.

Reconstruction of the embedding space from noisy time series

The application of dynamical systems theory to the study of a particular phenomenon depends on the identification of one (or a few) appropriate variable(s) that capture(s) the order (i.e. the attractor states) of the system. The choice of measured variable(s) is crucial for proper dynamic reconstruction (Tykierko, 2007). For example, the choice of the embedding delay is driven by the properties of the dynamic time series.

The selection of the parameters defining the embedding space is mostly based on one observed variable and under the assumption that the ‘presumed’ dynamical system generating the time series data does not change. In the ideal world, in the limit of an infinite amount of noise-free data, the numerical approaches to the calculation of dynamic invariants of well-known systems yield their theoretical values. One variable is as good as another for the embedding reconstruction. In the real world, the finitude of the experimental data inevitably contains noise due to environmental fluctuations and limited experimental resolution, which hinders the reconstruction of the embedding space and the calculation of its dynamic invariants. Casdagli et al. (1991) classified errors into three categories:

- Observational noise: The measuring instruments are noisy; what we actually observe is $x = h(\mathbf{z}) + \eta$, where $h : \mathfrak{M} \rightarrow \mathbb{R}$ is an observation function and η is an observational noise.
- Dynamical noise: External influences perturb the variables of the state space so that from the point of view of the system under study the evolution is not deterministic. This kind of system is called a stochastic dynamical system.
- Estimation error: We can estimate the dynamics in the reconstructed state space, but with a finite amount of data and limited data resolution the approximation is never perfect.

A robust embedding may reduce the effects of noise and data resolution. Dynamic invariants in robust embedding spaces for the analysis of noisy data are being explored but they are not sufficiently formalized yet. One option is to observe more than one variable of the ‘presumed’ dynamical system and create a single embedding space joining the embedding spaces of the variables. See (Barnard, Aldrich, & Gerber, 2001) for an example. Another option, proposed in (Vlachos & Kugiumtzis, 2010), consists of multidimensional embedding combined with non-uniform embedding. More research must be done about the advantages of these methods and their usability for analysing motion data.

Stationarity of the presumed dynamical system

Many studies mentioned in Chapter 2 used dynamic invariants for the quantification of motion variability and stability in human motion data. These studies assumed that the ‘presumed’ dynamical system behind the time series does not change during the data recording, i.e. the system is stationary. However, there are many variables in the human body which can neither be controlled nor even measured, e.g. the change of motion intention, the level of concentration during motion activity, and unintentional synchronization between humans and between human motion and metronomes reported by Nessler, De Leone, and Gilliland (2009).

If a long time series is available with a high sampling rate, it is recommended that one analyse different sections of the time series and compare their resulting invariants to confirm that the system did not change. Notice that, it is always possible to create several realizations out of that single set by chopping up the data. However, the numerical algorithms of the dynamic invariants are compromised by shortened data sets, reducing their power to reject a null hypothesis (Theiler, Eubank, Longtin, Galdrikian, & Farmer, 1992). Different ways of using invariants to detect changes in a dynamical system with chaos have been proposed in (Tykierko, 2007; Monroig, 2009). Further, specific tests for nonstationarity in a non-linear context have been proposed in the literature based on the theory of non-linear prediction. For a discussion of the limitations of those techniques, see (Schreiber, 1999).

Influence of filtering on the reconstructed embedding spaces

The accuracy of the numerical estimation of invariants from experimental data depends on the quantity and quality of the data as well as on the quality of the reconstruction of the complex dynamical system using an embedding space. Therefore, before filtering the data, one should always try to investigate the raw data.

In order to improve the signal-to-noise ratio (SNR) of the data, traditional linear filters like finite response filters (FIRs) and infinite impulse response filters (IIRs) are used frequently. They are based on the assumption that the signal and noise components can be distinguished in the spectrum. For sampled signals from non-linear systems, this poses a problem since the signal itself can have a broadband spectrum. Filtering with these methods can influence the reconstruction of the embedding spaces and their characteristics. For a review of the effects of filtering on attractors, see (Rosenstein & Collins, 1994; Broomhead, Huke, & Muldoon, 1992; Sauer et al., 1991).

Broomhead et al. (1992) demonstrated that a finite order non-recursive filter will not change the dimension. However, the cut-off frequency of the filter can influence the reconstruction of the embedding space. Under filtering, some complications are caused by the existence of periodic cycles (Sauer et al., 1991) when the cut-off frequency of the filter is not large enough. The influence of FIR on the reconstructed embedding space is twofold (Parlitz, 1998):

1. Due to the filter, elements of the time series are connected that are not close neighbours in time. This leads to (over) folding of the attractor (i.e. something similar to a choice of too large a delay in time).
2. On the other hand, FIR filters average the values of the time series and thus the components of the subsequently generated state vectors are more or less similar (i.e. an effect similar to the choice of too small a time delay).

The influence of an IIR filter on the reconstructed embedding space may also destroy the diffeomorphic equivalence between the original state space and the reconstructed state space. For example, the measurements of the fractal dimension can become dependent on the contraction rates associated with the filter dynamics (Badii et al., 1988; Mitschke, Möller, & Lange, 1988). Further, an IIR can induce a distortion of the geometry of the attractor, so that the observed system acquires greater complexity and, therefore, increased dimensionality (Rosenstein & Collins, 1994).

3.3 Dynamic invariants

If a dynamical system converges to an attractor in an M -dimensional state space, Takens' theorem says that it is possible to reconstruct the topological properties of this M -dimensional attractor using a projection to an N -dimensional embedding space created from observations of the system.

Given an appropriate reconstruction via embedding, one can investigate many aspects of the dynamics of the system that generated the time series, with the help of the dynamic invariants, e.g. to distinguish chaotic behaviour from noisy behaviour, to determine how many variables are needed to model the dynamics of the system, to sort systems into classes, and to detect changes in the dynamic behaviour of the system.

Statistical measures and topological methods are the two major groups of approaches used when studying chaos in dynamical systems. These groups often give us different information about the same dynamical system (Tufillaro, 1994) and together may provide sufficient discrimination power for the classification of observed physical systems. Statistical measures rest on the determination of metric properties of orbit points on an attractor. Topological invariants are purely geometric properties of the vector field describing the dynamics.

Statistical measures include, for example, recurrence period density entropy (RPDE) (Little, McSharry, Roberts, Costello, & Moroz, 2007), approximate entropy (Pincus, 1995) and multiscale entropy (Costa, Goldberger, & Peng, 2005). Topological invariants include simple non-linear prediction error (Hegger, Kantz, & Schreiber, 1999), Lyapunov exponents (Wolf, Swift, Swinney, & Vastano, 1985; Sano & Sawada, 1985; Rosenstein, Collins, & De Luca, 1993; Goldsmith, 2009), various fractal dimensions (Moon, 2008; Lopes & Betrouni, 2009), correlation dimension (Ghez & Vaienti, 1992; Casaleggio, Corana, & Ridella, 1995), and multi-fractal scaling functions such as detrended fluctuation analysis (DFA) (Peng, Havlin, Stanley, & Goldberger, 1995).

The most commonly used dynamic invariants are the largest Lyapunov exponent, which contains information on how far in the future predictions are possible, the correlation dimension, which indicates the complexity of the dynamical system, and RPDE, which characterizes the repetitiveness of a time series in state space. They are used in this dissertation and explained in the following.

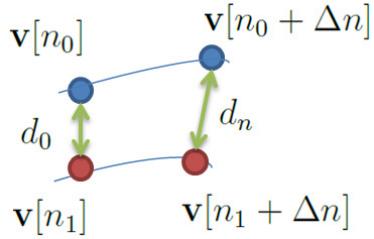


Figure 3.5: Defining $d_0 = \|\mathbf{v}[n_0] - \mathbf{v}[n_1]\|$ and $d_n = \|\mathbf{v}[n_0 + \Delta n] - \mathbf{v}[n_1 + \Delta n]\|$, the maximal Lyapunov exponent is calculated from averaging the logarithmic distance ratio $\frac{1}{\Delta n} \ln \frac{d_n}{d_0}$ over the available data.

3.3.1 Maximal Lyapunov exponent

Lyapunov exponents or Lyapunov spectra $\lambda_1, \lambda_2, \dots, \lambda_i$ describe the action of the dynamics defining the evolution of the trajectories, i.e. the average expansion, respectively, contraction, of the evolution of the trajectories in different local directions on an attractor of a dynamical system. The number of Lyapunov exponents i is equal to the dimension of the phase space and ordered in decreasing value. The most robust and important one is the maximal Lyapunov exponent λ_1 . The non-leading exponents are notoriously difficult to estimate from time series data and the exponents beyond the maximal exponent have not so far been demonstrated to be a useful concept (Schreiber, 1999).

When the system is chaotic, nearby trajectories on the attracting manifold diverge, on average, at an exponential positive rate characterized by λ_1 . Consider the representation of the time series data as trajectories in the embedding space and let $\mathbf{v}[m]$ and $\mathbf{v}[l]$ be two close points of trajectories in state space with distance $\|\mathbf{v}[m] - \mathbf{v}[l]\| = \delta_0$. Then one can consider the distance δ_0 as a small perturbation, which should grow exponentially in time. Denote by $\delta_{\Delta n}$ the distance after a time Δn between the two trajectories emerging from these points, $\delta_{\Delta n} = \|\mathbf{v}[m + \Delta n] - \mathbf{v}[l + \Delta n]\|$. Then the Lyapunov exponent λ_1 is determined by

$$\delta_{\Delta n} \approx \delta_0 e^{\lambda_1 \Delta n}, \quad 1 \leq \Delta n < N - 1. \quad (3.7)$$

The maximal Lyapunov exponent quantifies the degree of local instability in a state space, but stochastic noise in the recordings of the variables of dynamical systems (e.g. observational noise, dynamical noise, or both) can trigger findings of positive Lyapunov exponents even for systems that are not ‘chaotic’ (Dingwell, 2006). For slight noise levels, there may still be a range of length scales where the separation proceeds exponentially until a finite size saturation is reached (Schreiber, 1999). Several methods have been proposed for the calculation of the maximal Lyapunov exponent from experimental data. For examples, see (Wolf et al., 1985; Sano & Sawada, 1985; Rosenstein et al., 1993; Goldsmith, 2009).

Due to its robustness, the algorithm of Kantz (1994) is often used. The maximal Lyapunov exponent λ_1 is calculated from a discrete time series using embedding points defined by Eq. 3.2. See Fig. 3.5. A reference point $\mathbf{v}[l]$ is chosen and its neighbours are selected from the neighbourhood of reference point $U_\varepsilon(\mathbf{v}[l]) := \left\{ \mathbf{v} \mid \|\mathbf{v}[l] - \mathbf{v}[n]\| < \varepsilon \right\}$ with distance to the reference point less than a certain distance, ε . Then, one computes the distances of all selected neighbours to the reference point following the trajectories as a function of the relative time Δn for $0 \leq \Delta n \leq \Delta n_{max}$. Repeating the latter for the L available reference points $\mathbf{v}[l]$ and calculating the averages of these results $S(\Delta n)$,

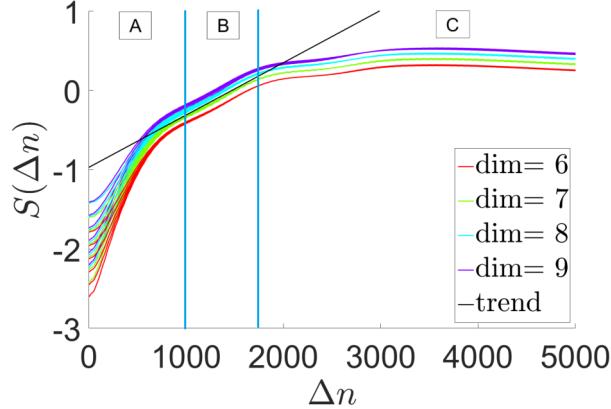


Figure 3.6: Example of curves for the calculation of maximal Lyapunov exponent from the noisy positional data described in Section 4.2. Three different regions denoted as A, B, and C can be appreciated: Noise ($\Delta n < 1000$), scaling ($1000 < \Delta n < 1800$), and saturation ($1800 < \Delta n$). In order to avoid the subjective selection of the scaling region, it is recommended to calculate $S(\Delta n)$ for different embedding dimensions and neighbourhood sizes. If linear increases are found with similar slopes for different parameters and for a reasonable range of Δn , then the average slope can be taken as an estimate of the maximal Lyapunov exponent λ .

$$S(\Delta n, \varepsilon) = \frac{1}{L} \sum_{l=1}^L \ln \left(\frac{1}{|U_\varepsilon(\mathbf{v}[l])|} \sum_{\mathbf{v}[n] \in U_\varepsilon(\mathbf{v}[l])} \|\mathbf{v}[l + \Delta n] - \mathbf{v}[n + \Delta n]\| \right), \quad (3.8)$$

with $|U_\varepsilon(\mathbf{v}[l])|$ being the number of nearest neighbours, noise in the data and fluctuations of the effective divergence will average out. If for some range of Δn the function $S(\Delta n)$ exhibits a robust linear trend, its slope is a stable and robust estimate of the maximal Lyapunov exponent. If λ is positive, there is a strong signature of chaos in the time series data. The higher the instability and the divergence of neighbouring trajectories in embedding space, the larger is the value of λ_1 .

In order to avoid the subjective selection of the scaling region, it is recommended to calculate $S(\Delta n)$ for different embedding dimensions and neighbourhood sizes. If linear increases are found with similar slopes for different parameters and for a reasonable range of Δn , then the average slope can be taken as an estimate of the maximal Lyapunov exponent λ . See Fig. 3.6 for an example.

3.3.2 Fractal dimension and correlation dimension

The fractal dimension is a topological property of an invariant set (attractor) of a dynamical system. It supplies a parametrization of the conceptual complexity of any geometrical object and of the space it is contained in. Further, it gives a lower bound for the number of equations or variables needed for modelling the underlying dynamical process.

The fractal dimension has a number of related definitions. Some definitions are equivalent, some are not. For a table of correspondences among fractal measures, see (Balocchi, 2011). Some well known definitions are the correlation dimension, the Hausdorff dimension, the box-counting dimension, the information dimension, and the Kaplan–Yorke dimension. For details, see (Tykierko, 2007; Moon, 2008; Lopes & Betrouni, 2009).

One of the most popular definitions is the correlation dimension, which has the advantage of

being simply and quickly calculated from a time series. One considers the probability that two embedding points ($\mathbf{v}[i]$, $\mathbf{v}[j]$) defined by Eq. 3.2 chosen randomly from the reconstructed invariant set, and evolving according to the dynamics of the system, will be within distance ϵ of each other. Then one examines how this probability changes as the distance ϵ varies. The correlation sum $C(N, \epsilon)$,

$$C(N, \epsilon) = \frac{2}{(N - n_{min})(N - n_{min} - 1)} \sum_{i=1}^N \sum_{j=i+n_{min}}^N \Theta(\epsilon - \|\mathbf{v}[i] - \mathbf{v}[j]\|), \quad (3.9)$$

approximates the desired probability, where Θ is the Heaviside step function, i.e. $\Theta(x) = 0$ for $x \leq 0$ and $\Theta(x) = 1$ for $x > 0$. The sum in Eq. 3.9 counts the number of pairs of embedding points ($\mathbf{v}[i]$, $\mathbf{v}[j]$) with distance less than ϵ from each other. In order to avoid the influence of temporal data correlation in the correlation sum C , one excludes those pairs of embedding points which are close not because of the geometry of the attractor but just because they are closer than n_{min} in (discrete) time. The time window n_{min} is called the Theiler window. For more details, see (Kantz & Schreiber, 2004).

As the data size grows, $N \rightarrow \infty$, and for small $\epsilon \rightarrow 0$, one expects that the correlation sum scales with a power law $C(N, \epsilon) \propto \epsilon^{d_c}$. The exponent d_c is called the correlation dimension and defined as,

$$d_c = \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \frac{\partial \ln(C(N, \epsilon))}{\partial \ln \epsilon}. \quad (3.10)$$

In practice, the correlation dimension d_c may be estimated from a time series by calculating the slope of the curve $C(N, \epsilon)$ vs ϵ in log – log scale as follows. First, an m -dimensional trajectory is constructed using embedding points defined by Eq. 3.2. Then, the correlation sum $C(N, \epsilon)$ is computed according to Eq. 3.9 and the function $\log C(N, \epsilon)$ versus $\log \epsilon$ (i.e. a typical log-log plot) is inspected for a linear region (referred to as the scaling region) for small ϵ . The slope of this linear region, if there is one, is taken as an estimate of the correlation dimension of the projection of the attractor to the m -dimensional embedding space. See (Casaleggio et al., 1995; Grassberger & Procaccia, 1983b, 1983a) for details.

Fig 3.7 shows an example of curves with three typical regions (denoted by A, B, and C) obtained from a noisy time series. For series characterized by chaotic behaviour, the correlation sum shows a noise region (A), followed by a linear region up to a curve peak (B) and finally by drop region (C). The drop with a large $\log \epsilon$ is produced because the average size of the attractor is reached and only a few distances of pairs of embeddings points are larger than ϵ in Eq. 3.9. The presence of noise and the embedding parameters may influence the extension of the regions for the computation of d_c . For example, the noise region depends on the amount of noise and it may be reduced using non-linear filtering. See (Casaleggio et al., 1995) and (Kantz & Schreiber, 2004) for examples before and after filtering.

Since the attractor dimension, in the case of experimental signals, is unknown, one computes $C(N, \epsilon)$ for several embedding dimensions d and determines their slopes within the scaling region. The average of these slopes is an approximation of the correlation dimension d_c . The scaling region should appear with $d \geq \lceil d_c \rceil$, where $\lceil d_c \rceil$ standing for the ceiling of the correlation dimension d_c . For details, see (Ding, Grebogi, Ott, Sauer, & Yorke, 1993).

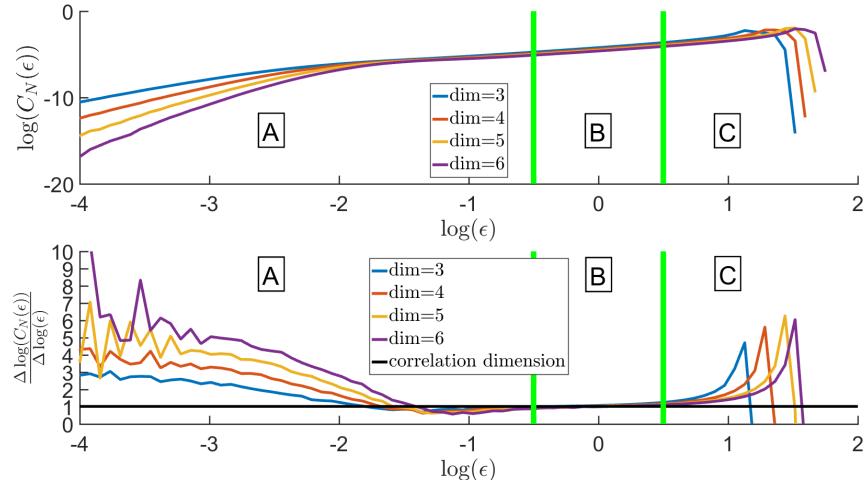


Figure 3.7: Top: Example of $\log C(N, \epsilon)$ versus $\log \epsilon$ curves for the calculation of correlation dimension from the noisy positional data described in Section 4.2. Three different regions can be appreciated (denoted by A, B, and C) obtained from a noisy time series. For series characterized by chaotic behaviour, the correlation sum shows a noise region (A), followed by a linear region up to a curve peak (B) and finally by drop region (C). The drop with a large $\log \epsilon$ is produced because the average size of the attractor is reached and only a few distances of pairs of embeddings points are larger than ϵ in Eq. 3.9. Bottom: Local slopes of the top figure. In this representation the scaling deviation and also the deviations from it are more visible.

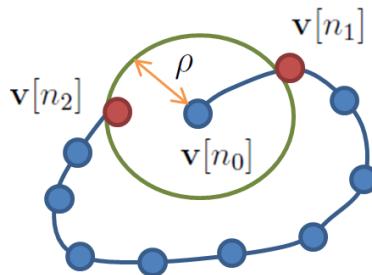


Figure 3.8: Graphical description of the procedure for estimating RPDE. Then, the trajectory of the embedded point $\mathbf{v}[n_0]$ is followed forward in time until it has left this ball, (i.e. the time n_1 where $\|\mathbf{v}[n_0] - \mathbf{v}[n_1]\| > \rho$, where $\|\cdot\|$ is the Euclidean distance), and continued further until the trajectory returns again to the ball (i.e. the time n_2 where $\|\mathbf{v}[n_0] - \mathbf{v}[n_2]\| \leq \rho$ and $n_2 > n_1$). The difference between the departure time n_1 and the return time n_2 is called the recurrence time.

3.3.3 Recurrence period density entropy (RPDE)

The information about how often a dynamical system returns to some place in the original or in the reconstructed phase space allows estimates of relevant quantities, such as the entropy, in terms of the probability distribution of the return times.

The recurrence period density entropy (RPDE) is a simple method which characterizes the repetitiveness of a time series in an embedding state space. The RPDE is based on the statistics of the first Poincaré return map (Lathrop & Kostelich, 1989), which measures how much time a trajectory of a dynamical system takes to make two consecutive returns to a given transversal section to the trajectory, which is known as a Poincaré section.

The Poincaré return map was originally designed for studying recurrence in deterministic chaotic dynamics without noise. In the presence of noise, the position of the Poincaré section (i.e. transversal section) may be not tangential to the trajectory and the statistics created using these bad oriented sections may be wrong. To solve this problem, the RPDE places for an embedding reconstruction with dimension d_e instead of the transversal section a d_e -dimensional ball with radius ρ around the embedding data point $\mathbf{v}[n_0]$. Then, the trajectory of the embedded point $\mathbf{v}[n_0]$ is followed forward in time until it has left this ball, (i.e. the time n_1 where $\|\mathbf{v}[n_0] - \mathbf{v}[n_1]\| > \rho$, where $\|\cdot\|$ is the Euclidean distance), and continued further until the trajectory returns again to the ball (i.e. the time n_2 where $\|\mathbf{v}[n_0] - \mathbf{v}[n_2]\| \leq \rho$ and $n_2 > n_1$). See Fig. 3.8 for an illustration.

The difference between these two discrete times (n_2 and n_1) is used to define the discrete recurrence time $\Delta n = n_2 - n_1$. This procedure is repeated for all L embedding reference points $\mathbf{v}[l]$ with $l = 1, \dots, L$, forming a histogram $R[\Delta n]$ of recurrence times Δn . This histogram is normalized to sum to unity so as to give the recurrence time probability distribution

$$P(k) = \frac{R[k]}{\sum_{i=1}^{\Delta n_{\max}} R[i]}, \quad (3.11)$$

where Δn_{\max} is the maximum recurrence time found in the embedded space. The uncertainty of the recurrence time Δn is defined as the normalized entropy with units of nats (by convention, $0 \ln 0$ is taken to be zero)

$$\text{RPDE} = -\frac{\sum_{i=1}^{\Delta n_{\max}} P(i) \ln P(i)}{\ln \Delta n_{\max}}. \quad (3.12)$$

The choice of the radius ρ is critical and depends on the embedding parameters. If the trajectory is nearly periodic, we require that ρ be large enough to capture all the recurrences, but not so large as to find recurrences that are due to spurious intersections of the d_e -dimensional ball with other parts of the trajectory, violating the conditions for proper recurrence (Little et al., 2007). Thus, the choice of ρ must be based on the maximal distance between nearest neighbour points in the embedding space. For a size ρ small enough, a value of 0 means that the data is periodic and a value of 1 means that it is white noise. For details, see (Little et al., 2007).

3.4 Definition of chaos

Non-linear deterministic systems often exhibit complex and chaotic behaviour. ‘Chaos’ is a tricky thing to define. Even textbooks devoted to chaos do not really define the term. The term ‘chaos’ in colloquial language is often related to a state of disorder, lacking any order, and wild in appearance. Concerning chaos theory, deterministic chaos (or simply chaos) refers to a non-periodic long-term behaviour with apparent lack of order in a system that nevertheless obeys particular laws or rules.

Definitions of chaos given in (Devaney, 1989; Gleick, 1987; Broer & Takens, 2010; Strogatz, 2000; Boccaletti et al., 2000) suggest that chaotic dynamical systems are characterized by:

- A deterministic system with at least one non-linearity.
- Aperiodic long-term behaviour.
- Sensitivity to the initial conditions: ‘Similar causes have similar effects’ is invalid in chaotic systems except for short periods. However, sensitive dependence on initial conditions alone does not give chaos.
- Topological transitivity or ‘mixing’: Neighborhoods of points in the state space will evolve over time so that they do not necessarily stick together in one localized clump.
- A dense collection of points with periodic orbits: Every point in the state space is approached arbitrarily closely by periodic orbits.

3.4.1 Which dynamical systems do exhibit chaos?

Sensitivity to the initial conditions is one feature of deterministic chaotic systems. This property implies that two trajectories emerging from two different close by initial conditions separate exponentially in the course of time (Boccaletti et al., 2000), i.e. a positive maximal Lyapunov exponent can be measured from the time series produced by such system.

Since, by definition, the phase trajectory cannot intersect with itself, e.g. by the uniqueness of the solutions, continuous systems in a 2-dimensional phase space cannot have such exponential divergence. Chaotic behaviour can only be observed in deterministic continuous systems with a phase space of dimension larger than 2. Such a phase space is large enough to have bounded and locally separating trajectories that do not cross.

Further, non-autonomous 2-dimensional systems are equivalent to autonomous 3-dimensional systems (see Section 8.1 for an example). Therefore, they can exhibit chaotic behaviour. Furthermore, it is well known that dissipative systems with a non-linear time-delayed feedback or ‘memory’ can produce chaotic dynamics, and the dimension of their chaotic attractors can be made arbitrarily large by increasing their delay time sufficiently (Voss, 2000). Thus, non-linear time-delayed systems are infinite dimensional systems. Also discrete systems can exhibit chaotic behaviour even if they are 1-dimensional, e.g. the logistic map. In addition, systems described by partial differential equations are infinite dimensional systems by definition even if they are described by only one equation because there is space to have tangled trajectories without violating the uniqueness of the solutions. All these systems exhibit a similar variety of dynamical behaviour to systems described by ordinary differential equations, e.g. chaotic solutions.

Chapter 4

Analysis of evidence of chaos in cycling ergometry

Time series measurements in the real world are sequential records of some physical variables from a (dynamical) process of interest. In particular, time series with erratic periodic behaviour (i.e. a pattern of recurrence with a component of unpredictability) are common in physiological signals related to the human body and in related kinematic and kinetic variables of cyclical human motion. In most cases, such time series can be classified either as pseudo-periodic or chaotic. A pseudo-periodic time series is defined as a representative of a periodic orbit (in a state space) perturbed by dynamical noise, observational noise, or with the combination of both types of noise. See Section 3.2.2 for a description of the different types of noise.

The mathematical tools to analyse a given time series must be selected according to the type of data. However, chaotic and pseudo-periodic time series often look similar. Thus, we might not be able to distinguish them from each other only through visual inspections and quantitative techniques are needed instead (Luo, Nakamura, & Small, 2005). It follows that the first step for the analysis of the irregular behaviour of the time series is to investigate whether the irregularity is brought by stochasticity or by non-linear deterministic chaos.

A positive maximal Lyapunov exponent is one of the main characteristics of chaotic systems and their observed variables. For this reason, the maximal Lyapunov exponent is included as a part of the evidence of chaos. However, this exponent alone cannot be used for proving strictly deterministic chaos because random data may also produce a positive maximal Lyapunov exponent. Therefore, the evidence of chaos is complemented with surrogate data hypothesis testing. With surrogate data hypothesis testing, one can never conclude that the data is anything, but only that it is not, i.e. hypothesis testing tells us which property a time series lacks by rejecting the corresponding null hypothesis (Luo, 2005).

Evidence of deterministic non-linearity and chaos in cyclical motion has been reported based on surrogate data hypothesis testing in the literature on gait (Theiler, 1995; Small, Judd, & Mees, 2001; Miller, Stergiou, & Kurz, 2006) but not in pedalling motion. Knowing this, we proposed in (Quintana-Duque & Saupe, 2012) to analyse whether the intercycle dynamics of motion in cycling ergometry come from a deterministic chaotic system using the pseudo periodic surrogate (PPS) described in (Small, Judd, & Mees, 2001). The PPS creates surrogate data from the original time series, preserving the original cyclic pattern while changing the original intercycle dynamics. In this chapter, additional tests for the evidence of chaos are applied, based on (Small, Nakamura, & Luo, 2007; Nakamura & Small, 2005; Luo, Nakamura, & Small, 2005; Dolan & Spano, 2001).

In addition to the evidence of chaos in cycling ergometry, we found a lack of studies on the

dynamic structure of pedaling motion and its potential benefits for distinguishing between subtle differences in pedaling motion patterns. In (Quintana-Duque, 2012) two different methods for embedding (uniform and non-uniform) were evaluated, exploring their influence on two selected dynamic invariants based on embedding space, i.e. the maximal Lyapunov exponent (Wolf et al., 1985) and the recurrence period density entropy (RPDE) (Little et al., 2007).

The aim of this chapter is to explore the potential use of non-linear tools for the analysis of the evidence of chaos in the motion data of cycling ergometry. Section 4.1 describes surrogate data hypothesis testing for the evidence of chaos. Finally, Section 4.2 presents the results of the analysis of evidence of chaos in cycling motion data.

4.1 Surrogate data hypothesis testing

Surrogate data hypothesis testing is currently an important empirical technique for testing a certain null hypothesis based on a discriminating statistic. Once the null hypothesis to be rejected is defined, an ensemble of artificial surrogate time series, called surrogate data, is generated. The surrogate time series should preserve some characteristics of the original data, which are consistent with the null hypothesis, and destroy others. Using the ensemble of surrogates, one applies some discriminating statistic to both the surrogates and the original data, e.g. the dynamic invariants explained in Chapter 3. If the discriminating statistic value for the data is different from the ensemble of values estimated for the surrogates, one may reject the given null hypotheses as being a likely origin of the data. If the discriminating statistic value for the data is not distinct from that for the surrogates, one may not reject the null hypotheses. Obviously, the surrogate data method is not to define what the observed data is but it provides a rigorous way to exclude a significant hypothesis with which the data is inconsistent (Zhao, Sun, & Small, 2008).

The statistical power of hypothesis testing for the evidence of chaos depends on the choice of the discriminating statistic. If we fail to reject the null hypothesis, our null hypothesis might be true; however, we shall also note the possibility that our discriminating statistic is not powerful enough to detect the difference between the surrogates and the original time series (Schreiber & Schmitz, 1997). Statistical methods, e.g. mean, range, and standard deviation, are not suitable for testing for evidence of chaos because they cannot distinguish between chaotic and random time series with similar statistical features. However, the dynamic invariants can distinguish between chaotic and random time series easily. For example, we could look for self-similar geometry over a reasonable range of length scales and demonstrate that the uncertainties indeed grow exponentially over a certain period of time (Schreiber, 1999). See (Maiwald, Mammen, Nandi, & Timmer, 2008) for a comparison of possible discriminating statistics.

Human cyclic motion is strongly but not exactly periodic. That is, the kinematics recorded from motion using sensors exhibit both temporal and spatial deviations from precise periodicity within and across cycle repetitions. However, the exact nature of these deviations is unknown and possible explanations are that they are chaotic or the product of noise combined with linear dynamics.

Surrogate data hypothesis testing can be used to examine the nature of the fluctuations overlying the periodic movement patterns in kinematic motion data. The standard surrogate data repertoire provides algorithms for testing the hypotheses that the data has been created by, e.g. (i) independent and identically distributed noise, (ii) linearly filtered noise, or (iii) a monotonic non-linear transformation of linearly filtered noise. Algorithms for each of these three hypotheses generate an ensemble of fake (artificial) time series data assuming data without any trend or pseudo-periodicity. However, the mentioned hypotheses are neither sufficient nor even particularly useful for time se-

ries exhibiting irregular oscillations. The mentioned hypotheses will be always rejected because the original trend or pseudo-periodicity is destroyed in the surrogate data. See (Small, Yu, & Harrison, 2001) for a discussion. Thus, for time series exhibiting irregular oscillations, e.g. cyclic motion, other hypotheses are needed.

Let $x[n]$ denote the original time series data consisting of $n = 1, \dots, N$ measurements. If the next null hypotheses described by Small et al. (2007) are rejected, then we may claim the time series exhibiting irregular oscillations is very likely to be chaotic.

- Hypothesis 1: Data $x[n]$ has been created by a observational white noise process plus long-term trends or periodicities.
- Hypothesis 2: Data $x[n]$ has been created by a linear filtered noisy process, i.e. observational colored noise, plus long-term trends or periodicities.
- Hypothesis 3: Data $x[n]$ has been created by a stationary linear system plus observational noise.
- Hypothesis 4: Data $x[n]$ has been created from a noisy driven periodic orbit, i.e. no determinism other than the periodic behaviour plus dynamical noise.

The discriminating statistics, the methods for generation of surrogate data, and the criterion for rejection of the null hypothesis are main aspects of the surrogate data methodology. Some discriminating statistics based on the dynamic invariants were explained in Chapter 3. In the following, Section 4.1.1 describes methods for surrogate data and Section 4.1.2 reviews the rejection of the null hypothesis with a proposed method.

4.1.1 Methods for surrogate data

The surrogate data methodology is applied in non-linear time series analysis to formally reject the hypotheses that the observed data were produced by different noise processes. In order to test such hypotheses in data with irregular oscillations, surrogate data must be constrained to look like the data and to have the obvious periodic components. Recently, a series of algorithms have been proposed that mimic the underlying dynamics with the addition of different types of noise. In the following subsections, some surrogate data techniques are explained to test the hypotheses mentioned above. For more details, see (Small et al., 2007; Small & Tse, 2002; J. Zhang, 2007; Luo, 2005; Nakamura, Small, & Hirata, 2006).

Small-shuffle surrogates for data exhibiting cyclic components

The small-shuffle surrogate (SSS) algorithm, proposed by Nakamura and Small (2005), is used to investigate hypothesis 1. The assumption is that the frequencies of irregular fluctuations are higher than those of long-term trends or periodicities which are contaminating the data.

Let $x[n]$ denote the original time series data consisting of $n = 1, \dots, N$ measurements. The algorithm is essentially just a shuffling of the data index on a small scale, i.e. reordering the series $x[n]$ where each observation is shifted a little in time. The algorithm is the following:

1. Add Gaussian random numbers to the index of the original data, i.e. let the perturbed index be a real number $i[n] = n + Ag[n]$, for some constant parameter A .

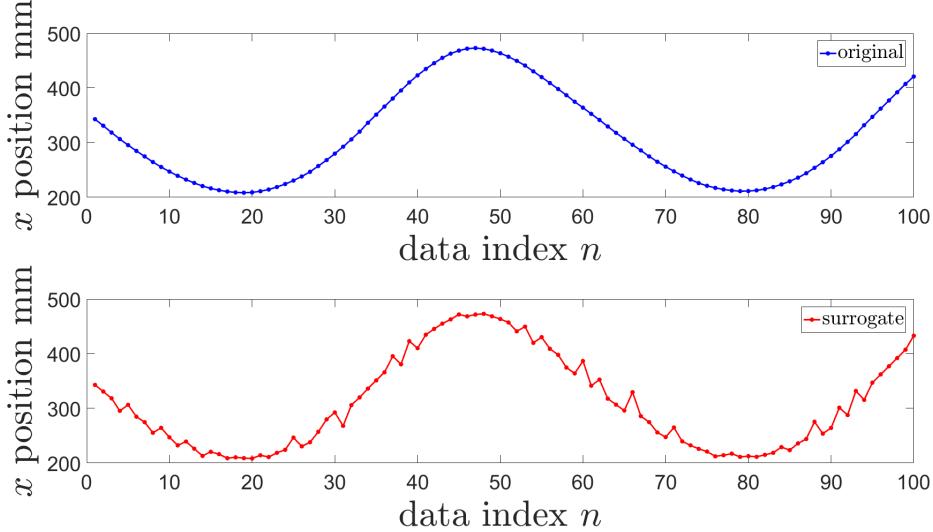


Figure 4.1: Top: Example of one original time series of the noisy positional data described in Section 4.2. Bottom: Surrogate data created by the small-shuffle surrogate and $A = 2$. The amplitude distribution is the same but any stationary dynamic is destroyed. The slow trend in the data is preserved.

2. Sort the elements of $i[n]$ in ascending order. Let $r[n] \in \mathbb{Z}$ be the index list of the sorted elements of $i[n]$.
3. The surrogate $s[n]$ is obtained from $s[n] = x[r[n]]$, i.e. reorder the original data with the indexes of the sorted perturbed indices.

Note that if A is too small, the data are shuffled only a tiny amount. Then, if $A \rightarrow 0$, $r[n] = i[n] = n$ and $s[n] = x[n]$. Conversely, if $A \gg N$, then $s[n]$ is almost equivalent to a random shuffling of the original data (Small et al., 2007). When the amplitude A is selected appropriately, the index is shuffled only on a small scale, where the generated surrogate data loses local structures or correlations, but preserves the long-term trends and periodicities. See Fig. 4.1 for an example with $A = 1$ where the original periodicity was preserved and the local correlation is destroyed. With $A = 1$, about 50% of the data points in the SSS data have the same index as the original data.

Truncated Fourier transform surrogate

The truncated Fourier transform surrogate (TFTS) algorithm, proposed by Nakamura et al. (2006), is used to test hypothesis 2. As mentioned before, surrogate data should be similar to the original data. That is, when the original data have long-term trends, it is preferable to preserve those trends. Normally, data with irregular fluctuations and long-term trends exhibit large power peaks in the low frequencies and almost white noise power in the high frequencies. The truncated Fourier transform surrogate (TFTS) algorithm works by preserving the power spectrum and low-frequency phases in the Fourier transform (i.e. preserving the global ‘slow’ behaviours), but randomising the high-frequency phases only (i.e. destroying local structures). Thus, it is possible to discriminate between linear and nonlinear data because the superposition principle is valid only for linear data (Nakamura, Luo, & Small, 2005). When data are nonlinear, even if the power spectrum is

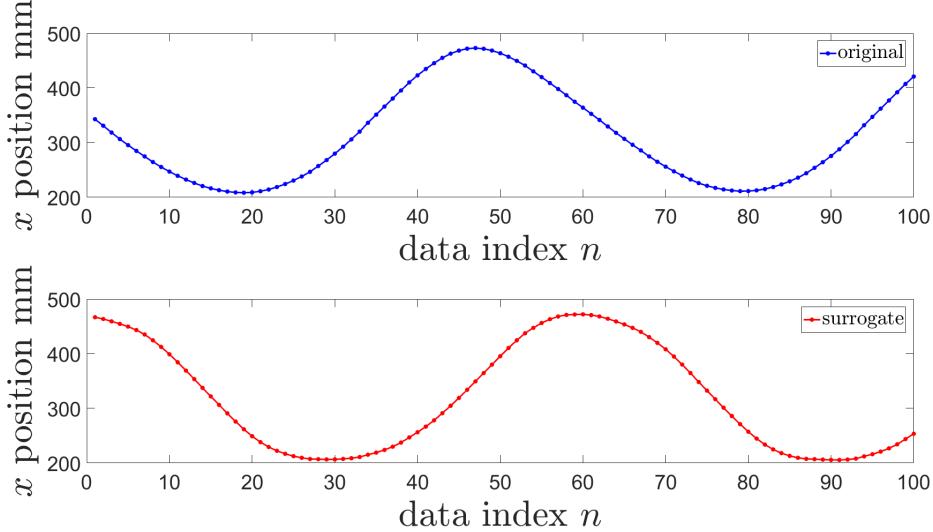


Figure 4.2: Top: Example of one original time series of the noisy positional data described in Section 4.2. Bottom: Surrogate data created by the truncated Fourier transform surrogate with f_c selected as the frequency at where the accumulated power starting from 0 Hz was 90% of the total power. Only the frequencies larger than f_c in the Fourier transform are randomised in phase, i.e. the surrogate appears like linear noise on short time scales. The power spectrum and slow trends in the data (or non-stationarity) are preserved.

preserved completely, the inversed Fourier transform data using randomized phases will exhibit a different dynamical behavior, i.e. the dynamic invariants calculated from this surrogate data will have different values than the dynamic invariants calculated from the original data. Hence, the null hypothesis addressed by the truncated Fourier transform surrogate (TFTS) algorithm is that irregular fluctuations are generated by a linear system. The algorithm is the following:

1. Let $x[n]$ denote the original time series data consisting of $n = 1, \dots, N$ measurements with a sampling frequency f_s , and let $X[k]$ with $|k| \leq N/2$ be the complex discrete Fourier transform of $x[n]$. Let $|X[k]|$ and $\angle X[k]$ be the magnitude and angle representation respectively.
2. Keep the original phases $|(f_s k)/N| \leq f_c$. Generate random phases for $\angle X[k]$ if $(f_s k)/N > f_c$. Finally, make $\angle X[-k] = -\angle X[k]$ for $(f_s k)/N > f_c$ to obtain real number when calculating the inverse Fourier transform.
3. Compute the inverse Fourier transform of the complex series $\{X\}$. This is the surrogate data $s[n]$ consisting of $n = 1, \dots, N$.

In the time domain, the surrogate data appears like linear noise at short time scales, and the structure of the original data (low frequencies) is preserved in the surrogate data at long time scales. For details, see (Small et al., 2007).

The f_c can be selected roughly from the power spectrum of the original data where the power spectrum is almost white for preserving the long-term behaviour. For details, see (Nakamura et al., 2006). If f_c is too large, the randomization of phases is very little or none at all, and then the TFTS data are almost identical to the original data. If the f_c is too small, the number of randomized phases is increased, and the TFTS data do not preserve long-term trends.

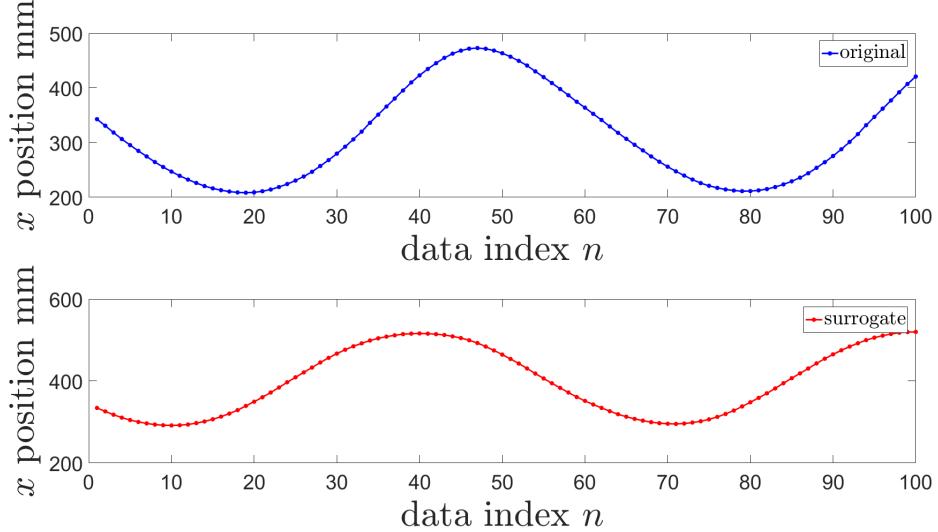


Figure 4.3: Top: Example of one original time series of the noisy positional data described in Section 4.2. Bottom: Surrogate data created by the temporal shift surrogate. The dynamic invariants as the ubiquitous correlation dimension of $x[n]$ and $s[n]$ are the same only if $x[n]$ is periodic.

There is a problem with linear surrogate algorithms using the Fourier transform. When there is a large difference between the first and last points of the time series, the Fourier transform will treat this as a sudden discontinuity in the time series. As a result, this will introduce significant spurious high-frequency power into the power spectrum (Nakamura et al., 2006). To ameliorate this artifact, the original data is symmetrized, i.e. the original time series and the reverse time version of the same are concatenated ($x[1], x[2], x[3], \dots, x[N], x[N], x[N-1], \dots, x[2], x[1]$). By this procedure, there is no end-point mismatch in the data (Small et al., 2007).

Temporal shift surrogate

The temporal shift surrogate (TSS) algorithm, proposed by Luo, Nakamura, and Small (2005); Nakamura et al. (2005), is used to test hypothesis 3. As before, let $x[n]$ denote the observed scalar time series. For sufficiently large Δn ,

$$s[n] = \alpha x[n] + \beta x[n + \Delta n] \quad (4.1)$$

defines the surrogate time series. If the observed system is periodic (or linear), two independent realizations of that system a and b are periodic (resp., linear). Similarly, the superposition $a + b$ of those two realizations maintains that periodicity (or linearity), i.e. the data $x[n]$ and the surrogate $s[n]$ have the same deterministic dynamics. For pseudo-periodic or non-linear time series, the latter is not true, i.e. the data and surrogate are effectively independent with different dynamic properties.

To generate a suitable ensemble of surrogates, one chooses a sufficiently large value of Δn larger than the first zero of the autocorrelation function. For each surrogate realization, the parameters are selected randomly subject to the constraint that $\alpha^2 + \beta^2 = 1$. This constraint ensures that the power spectra of the surrogate and data are identical.

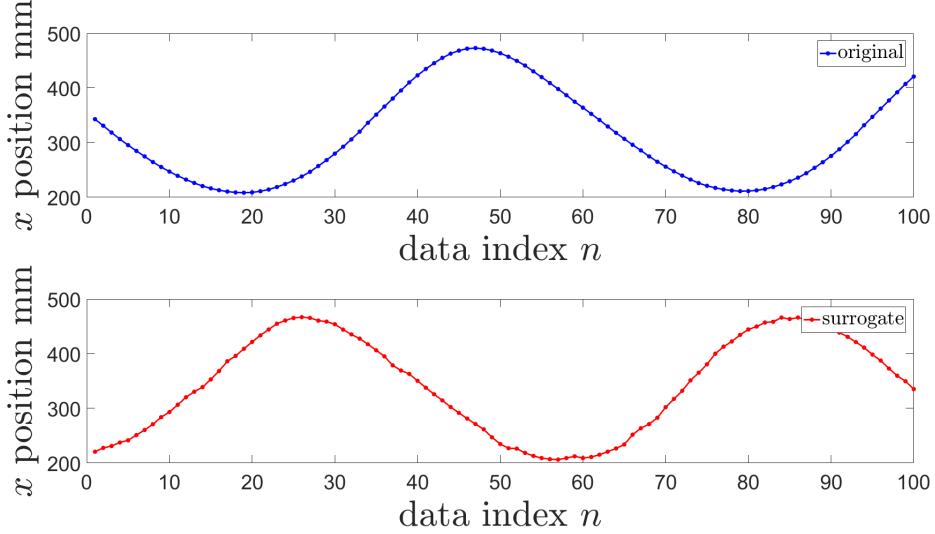


Figure 4.4: Top: Example of one original time series of the noisy positional data described in Section 4.2. Bottom: Surrogate data created by the pseudo-periodic surrogate. The macroscopic dynamics of the original system are preserved but microscopic dynamics are dominated by the random perturbations.

Pseudo-periodic surrogate

The pseudo periodic surrogate (PPS) algorithm, proposed by Small, Yu, and Harrison (2001), is used to test hypothesis 4. The PPS generates surrogate data that preserve the largescale deterministic features (such as periodic trends) but destroys the fine structures (such as linear or non-linear determinism) present in the observed data (Small & Tse, 2002). The PPS algorithm is inspired by the application of local modelling techniques to generate simulated time series, and by contaminating a trajectory on the attractor with dynamic noise. For this algorithm, the attractor of the observed dynamical system is first reconstructed from the time series data $x[n]$ using an embedding space defined by Eq. 3.2, i.e. $\mathbf{v}[n] = (x[n], x[n - \tau], x[n - 2\tau], \dots, x[n - (d_e - 1)\tau])$ with $n = 1, \dots, N$ where the embedding dimension d_e and the embedded time delay τ can be calculated using the techniques described in the last chapter. If the time series is nearly cyclic, the trajectory in the embedding space will describe cycles. Then, the embedding points of neighbouring cycles in the reconstructed attractor are used to create a new attractor with a noisy trajectory. The PPS algorithm, as described in (Small et al., 2007), is the following:

1. One selects an initial random embedding point $\mathbf{s}[1] = \mathbf{v}[m]$ with $1 \leq m \leq (n - 2 * (d_e - 1)\tau)$ from the reconstructed attractor as start point of the surrogate trajectory. Let $i = 1$.
2. Compute the distance $d[n] = \|\mathbf{s}[i] - \mathbf{v}[n]\|$ between the current state $\mathbf{s}[i]$ and every point on the attractor \mathbf{v} .
3. Assign the weight $w[n] = \exp\left(\frac{-d[n]}{\rho}\right)$ and the probability $p[n] = \frac{w[n]}{\sum_k w[k]}$ to each point $\mathbf{v}[n]$. The parameter ρ is the noise radius that indicates the amount of randomization and will be discussed latter.
4. Choose a neighbour of $\mathbf{s}[i]$ according to the implied probability distribution. That is, the probability of choosing $\mathbf{v}[n]$ is $p[n]$.

5. Suppose that $\mathbf{v}[j]$ with $1 \leq j \leq (n - 2 * (d_e - 1)\tau)$ is the chosen neighbour in step 4. Then let the successor of $\mathbf{s}[i]$, $\mathbf{s}[i + 1] = \mathbf{v}[j + 1]$.
6. Increment i .
7. Repeat from step 2 until $i > N$.
8. At this point, stop. The surrogate data is formed from the scalar first components of the embedding points of the noisy attractor $\mathbf{s}[i]$ with $i = 1, \dots, N$.

The resulting one-dimensional surrogate data contain a similar periodic pattern as the original data, but any subtle dynamic deterministic scale is destroyed with an appropriate choice of the embedding parameters (τ and d_e) and the noise parameter ρ . If ρ is large (i.e. very little randomization), then the surrogate data and the original signal will be nearly identical. If ρ is very small (i.e. too much randomization) then this algorithm is equivalent to random sampling. The selection of the parameter ρ must be somewhere between these two extremes. One effective way to do this is to choose the value of ρ in order to have the probability that we don't just follow the current state is around 0.5. See for details, the implementation of this idea in the ANTA software described in (Small, 2005).

probability (i.e. the probability that we don't just follow the current state), if the transition probability p is 0.5 the probability of a sequence of length 2 (the shortest allowable) $p(1-p)$ is maximal. Lower the probability for longer sequences.

4.1.2 Rejection of the null hypothesis

The discriminating statistic and the criterion for rejection of the null hypothesis are critical aspects of the surrogate data methodology. If multiple realizations are available for the observational data, then it may be possible to compare the two empirical probability distribution functions (PDFs) of the discriminating statistic calculated on both observed experimental data and surrogate data. For example, the distributions can be compared with the Kolmogorov–Smirnov test and their means can be compared with Student's t -test. However, most of the time, only one experimental time series and its single estimated discriminant statistic are available. Then, the distribution of the estimated discriminant statistic cannot be modelled from a single value and methods considering the distances between two Gaussian distributions as the Kullback-Leibler divergence are not considered in the literature.

It is always possible to create several realizations out of that single set by chopping up the data. However, the numerical algorithms of the dynamic invariants are compromised by shortened data sets, reducing their power to reject a null hypothesis (Theiler et al., 1992). Thus, one needs to check whether the single value of the estimated discriminant statistic of the experimental time series falls within or outside the value distribution of the discriminant statistic of the surrogate data. Two ways for the rejection of the null hypothesis are available:

- Parametric: This assumes that the discriminating statistic values follow a Gaussian distribution and the distribution parameters (mean and variance) can be estimated. If the distribution of the discriminant statistic of the surrogate data is represented by an error bar, the significance, i.e. the confidence level, is calculated from the number of ‘sigmas’ that the discriminant statistics of the original data lies outside these bounds.

- Non-parametric or ranking criterion: Theiler and Prichard (1997) proposed a modified version of Student's t -test. Suppose the statistic of the original time series is ϕ_0 and the surrogate values are $\{\phi_i\}_{i=1}^N$ given N surrogate realizations. Then if the statistic of both the original time series and the surrogates follows the same distribution, the probability is $1/(N+1)$ for ϕ_0 to be the smallest or largest among all of the values $\{\phi_0, \phi_1, \dots, \phi_N\}$. After ranking $\{\phi_i\}_{i=0}^N$, if one finds that ϕ_0 is the smallest or the largest value among all of the values in $\{\phi_i\}_{i=1}^N$, i.e. if ϕ_0 obviously falls outside the distribution of $\{\phi_i\}_{i=1}^N$, it is quite possible that ϕ_0 instead follows a different distribution from that of $\{\phi_i\}_{i=1}^N$. The false rejection rate is considered as $1/(N+1)$ for one-sided tests and $2/(N+1)$ for two-sided ones (Luo, Zhang, & Small, 2005).

For the study of evidence of chaos using the surrogate data methodology, the non-parametric criterion and the dynamic invariants as discriminating statistics are commonly used. However, the standard deviation of the discriminant statistic values is ignored and some hypotheses could be falsely rejected. For instance, the dynamic invariants calculated from a time series are ideally invariant for an increasing number of embedding dimensions. In the real world, they can fluctuate due to noise for increasing number of embedding dimensions, different neighbour sizes, or different sets of randomly selected points for the calculation. Thus, the mean and the standard deviation of the calculation of the dynamic invariants should be included for hypothesis testing.

In Fig. 4.5 is shown an example of a possible wrong hypothesis rejection that I found in (Luo, 2005). For this example, the discriminant statistic was the correlation dimension and a small set of data points are used as the reference. Thus, some statistical fluctuations are common even for the same data set when calculating its correlation dimension. When the ranking criterion was applied without checking the standard deviation of the discriminant statistic, i.e. if only the mean correlation dimension is used with the ranking criterion, then the hypothesis is rejected because the mean correlation dimension of the original time series is the lowest value. However, Luo (2005) ignores the analysis including the standard deviation of the correlation dimension values of both surrogate and original time series. If the standard deviation is included in the analysis, one should not reject the null hypothesis because the correlation dimension estimates of 7 surrogates fall within the distribution of the correlation dimension estimates of the original time series.

I propose to check the null hypothesis including the standard deviation from the calculation of the discriminating statistic and modifications of the parametric method. Assuming that the estimates of the dynamic invariant of each surrogate data are normally distributed, as the parametric rejection of the null hypothesis does, a probability density function can be calculated from the estimates of the discriminating statistic of the surrogates as the average of curves. Then, one observes the interval $\bar{x}_{\text{ori}} - \sigma_{\text{ori}} < x < \bar{x}_{\text{ori}} + \sigma_{\text{ori}}$, where the x is the value of the discriminating statistic, \bar{x}_{ori} and σ_{ori} are the mean and the standard deviation of the discriminating statistic of the original data. If the area under the empirical probability density curve in this interval is smaller than a certain confidence level, e.g. 0.05, then the hypothesis is rejected. This area can have a value between 0 and 0.68. On the one hand, if the mean of the empirical probability density curve calculated from surrogate data is far away from the mean of the discriminating statistic calculated from the original data \bar{x}_{ori} and the standard deviations of both density curves are small, then the area in mentioned interval is almost 0. On the other hand, if the mean and the standard deviation of the empirical probability density curve are equal to the \bar{x}_{ori} and σ_{ori} , respectively, the area in mentioned is about 0.68 assuming a normal distribution.

Fig. 4.6a and Fig. 4.6b compare the results between the ranking criterion and the proposed method calculated from one of the time series of the noisy positional data described in Section 4.2. In these figures, the black line and the dashed black lines represent the mean and the standard

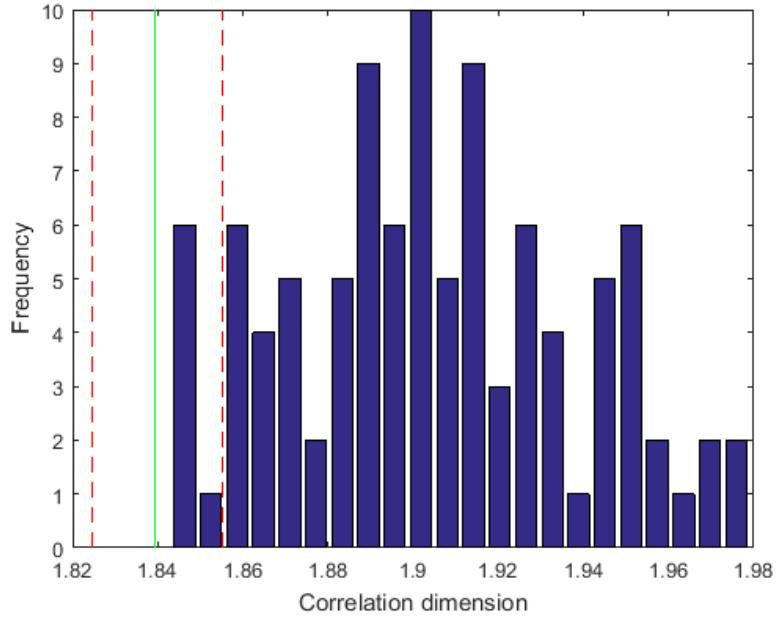
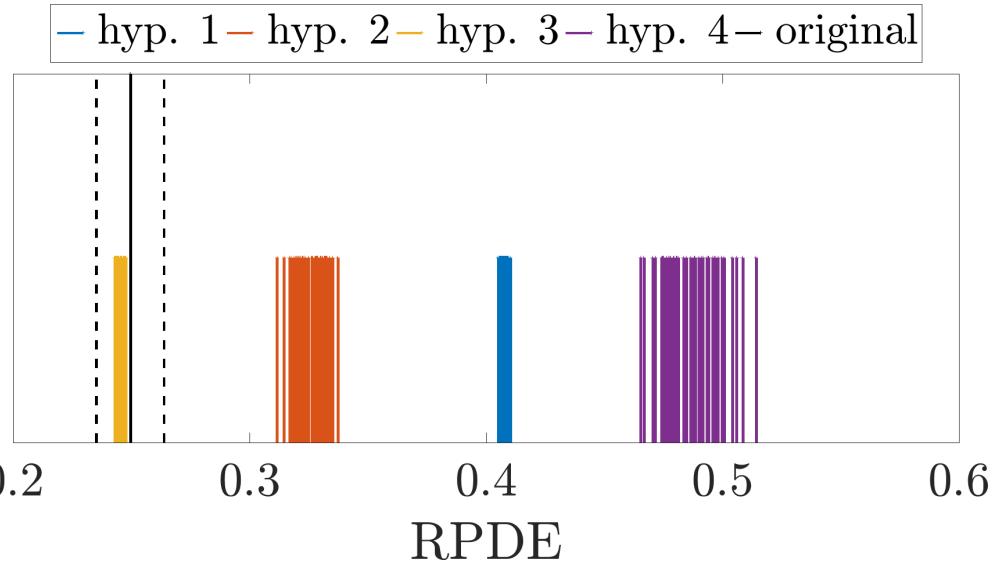
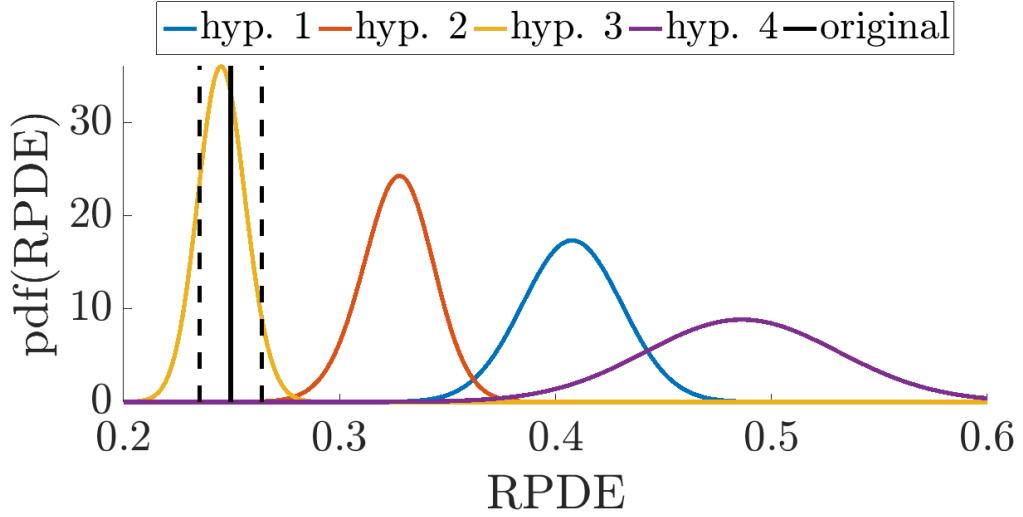


Figure 4.5: An example of a possible wrong hypothesis rejection published on page 69 in (Luo, 2005). The histogram was created using the reported data of 100 surrogates. The abscissa gives the correlation dimension values and the ordinate gives the corresponding frequency for each histogram. Further, the green line represents the mean correlation dimension of the original time series calculated 100 times with different small sets of reference points, and the dashed red lines denote the correlation dimension mean of the original time series \pm the standard deviation. If only the mean correlation dimension is used with the ranking criterion, then the hypothesis is rejected because the mean correlation dimension of the original time series is the lowest value. However, Luo (2005) ignores the analysis including the standard deviation of the correlation dimension values of both surrogate and original time series. If the standard deviation is included in the analysis, one should not reject the null hypothesis because the correlation dimension estimates of 7 surrogates fall within the distribution of the correlation dimension estimates of the original time series.



(a) Example of hypothesis rejection using the method described in (Theiler & Prichard, 1997).



(b) Example of hypothesis rejection using the proposed method.

Figure 4.6: Example of hypothesis rejection using (a) the method proposed in (Theiler & Prichard, 1997) and (b) the proposed method calculated from the time series of 130N 85C P3 described in Section 4.2. The black line and the dashed black lines represent the mean and the standard deviation of the discriminant statistic values of the original data, respectively. The RPDE is the value of the discriminating statistic. Notice that the Theiler's method rejects the null hypothesis 3 but the discriminating statistic plus the standard deviation is in the distribution of the discriminating statistic of the surrogates. Further, with the method of the distribution it is shown clearly that the hypothesis cannot be rejected.

deviation of the discriminant statistic values of the original data, respectively. Further, the RPDE is the value of the discriminating statistic calculated with 5 different dimensions. Following the ranking criterion shown in Fig. 4.6a, the null hypothesis 3 can be rejected because the mean of the discriminant statistic of the original data is beyond the extreme of the discriminant statistic of the surrogates. However, the values between the mean and the standard deviation of the original data overlap the values of the discriminating statistic of the surrogates. See Fig. 4.6b. The area under the empirical probability density curve in the interval $\bar{x}_{\text{ori}} - \sigma_{\text{ori}} < x < \bar{x}_{\text{ori}} + \sigma_{\text{ori}}$ is clearly larger than 0.05. Therefore, the hypothesis cannot be rejected.

4.2 Analysis of evidence of chaos in motion data of cycling ergometry

Evidence of chaos in signals recorded from human bodies has been published, e.g. for gait data (Nessler et al., 2009; Dingwell & Cusumano, 2000; Stergiou, 2004), electrocardiogram (ECG) data (Zhao et al., 2008) and electroencephalogram (EEG) data (Theiler, 1995). The analysed data contained (strong) cyclic components and had positive Lyapunov exponents. Further, surrogate data was used to reject the hypotheses that the data has been created by: (i) independent and identically distributed noise; (ii) linearly filtered noise; or (iii) a monotonic non-linear transformation of linearly filtered noise. However, as was mentioned in Section 4.1.1, these hypotheses are inappropriate for time series exhibiting (strong) cyclic components because any of them destroys the essential periodic features of the original data. Thus, the conclusions reported in those papers was only that the gait motion data did not originate from a random process, which is obvious.

For data exhibiting strong cyclic components, surrogate data must be constrained to look like the data and to have the obvious periodic components. See (Miller et al., 2006) for a discussion. For a complete evidence of chaos, it is more natural to ask if at least the 4 null hypotheses mentioned in Section 4.1 can be rejected.

As far as we know, no previous publications about the study of evidence of chaos in pedaling motion have been published before our publications (Quintana-Duque & Sauer, 2012; Quintana-Duque, 2012). In this section, an extension of the results of the time series of the x -coordinate position of the knee joint recorded during pedaling is presented. In order to check the evidence of chaos within pedaling motion data, the maximal Lyapunov exponent and the correlation dimension were calculated (see Section 3.3 for a description of the methods). Further, the 4 null hypotheses mentioned in Section 4.1 were tested using surrogate data created with the corresponding methods described in Section 4.1.1. The recurrence period density entropy (RPDE) (see Section 3.3.3 for a description of the algorithm) was used as discriminating statistic to compare the original with the surrogate data.

Only if (1) the maximal Lyapunov exponent is positive, (2) the correlation dimension can be calculated from a scaling region and (3) all 4 null hypotheses mentioned in Section 4.1 are rejected, can one be assured (or at least persuaded) that there is evidence of chaos in a pedaling motion data. If one fails to reject one of the null hypotheses, the null hypothesis might be true. However, one should also note the possibility that the discriminating statistic is not powerful enough to detect the difference between the surrogates and the original time series.



Figure 4.7: Marker positions for motion capture during cycling in red circles. The x -coordinates of the knee joint recorded with the marker in the blue circle were used as data.

4.2.1 Methods

The knee motion during pedaling of one healthy rider on a configurable static bike (Cyclus2) was recorded using a motion capture system (Lukotronic, LUKOtronic-Steinbichler Optotechnik GmbH, Neubeuern, Germany). Using the motion capture system with a sampling rate of 200 Hz we recorded the position of the markers. See the circles in Fig. 4.7 for the placement of the markers on the knee (blue) and on the bike frame (red). The x -coordinates of the knee joint were used as data. During the recording, we used three reference markers attached to the bicycle frame. The markers defined the coordinate system, which was re-calculated for each measurement. In this way, the oscillations and deflections of the bicycle frame due to the pedalling motion and bike sway are taken into account during the recording of the cycling motion. In order to keep the markers always visible with distances between them as large as possible, the markers were placed on the head tube of the bicycle, on the seatstay, and on the top tube.

The three reference markers for the definition of the coordinate system formed a triangular constellation defining the xy plane. The marker placed on the head tube defined the origin of the relative coordinate system. The x axis was defined by the vector formed between the markers on the head tube and on the seatstay. The y axis was calculated as the orthogonal vector to the x -coordinate passing through the origin with the direction of the same side of the marker on the top tube. The direction of the z axis was calculated using the cross product between the vector defining the x axis and the vector defined by the origin and the marker on the top tube. See the vectors in Fig. 4.7.

One healthy rider participated in our tests. The average cadence was selected by the rider between 80 and 90 revolutions per minute (rpm). The free selection of the cadence is important because some studies have suggested that the amount of chaos in motion decreases when a pace maker is given for synchronization or a too narrow fixed cadence is prescribed, see (Nessler et al., 2009) for details. The pedal brake force (in Newton) was fixed with the bike simulator for each test according to the desired work load intensity (low, medium and high). To control the pedal brake force, we used a Cyclus2 ergometer (RBM Elektronik-Automation GmbH, Leipzig, Germany). During the ergometer's operation, the flywheel on the rear part is supplied with kinetic energy which maintains the angular momentum. The angular speed of the flywheel is decelerated by means of an eddy current brake. The eddy current brake guarantees a non-slipping transmission



Figure 4.8: Test protocol: 6-minute tests with low (185 W), medium (210 W) and high intensity power (240 W) were done. Each test was divided into three intervals (P1, P2, and P3) and the invariants were calculated for each one.

of the braking resistance. Operating the Cyclus2 in pedal force mode, a constant pedal brake force ($\pm 5\%$ error) is imposed.

Three tests, each one of 6 minutes, were done with low (around 185 W with mean cadence of 80 rpm and pedal brake force of 120 N), medium (around 210 W with mean cadence of 85 rpm and 130 N), and high intensity power (around 240 W with mean cadence of 90 rpm and 140 N), respectively, with a short pause of two minutes between them. Each test was divided into three intervals (denoted in the following as P1, P2, and P3) of two minutes, and one minute was used from each interval as data for the calculation of the invariants. See Fig. 4.8 for the time schedule of the 9 parts.

Some dynamic invariants and the parameters for the uniform embedding (τ and d_e) were calculated for each of the 9 time series. The selected dynamic invariants were the maximal Lyapunov exponent and correlation dimension which provided us information about the statistical and geometric properties of the attractor respectively. The embedding delay τ was selected from the first minimum of the mutual information curve of the time series $x[n]$, see an example in Fig. 3.4. Then, the embedding dimension d_e was calculated with the embedding delay τ as parameter using the false nearest neighbor method FNN. See Section 3.2 for a description of the methods. The mutual information curve was calculated using openTStools described in (Merkwirth, Parlitz, Wedekind, Engster, & Lauterborn, 2009) and the FNN was calculated using ANTA software described in (Small, 2005).

For the calculation of the maximal Lyapunov exponent from $S(\Delta n)$ from Eq. 3.8, we used five different neighbourhood sizes distributed in the interval $(\sigma(x[n])d_e)/20 < \epsilon < 50$, with $\sigma(x[n])$ as the standard deviation of the discrete time series $x[n]$ to confirm that linear increases in $S(\Delta n)$ exist for different neighbourhood sizes. Further, in order to get reliable results of this invariant, as suggested in (Kantz, 1994; Kantz & Schreiber, 2004), we calculated for each time series the average of the maximal Lyapunov exponent values obtained from embedding spaces created using embedding dimensions $d_e, d_e + 1, d_e + 2, d_e + 3$, and $d_e + 4$, 400 reference points with at least 50 neighbours (this avoids oscillations in the linear increase, if any exist). If the scaling region exists for d_e , it exists for larger dimensions as well as for intermediate values of ϵ .

In order to avoid a subjective selection of the linear increase (i.e. scaling region) in $S(\Delta n)$, an semi-automatic algorithm was implemented to fit suitable lines with increasing size. The idea behind the algorithm is simple. On the one hand, if a linear increase exists, then the fitted lines with increasing size should be parallel to the linear increase in $S(\Delta n)$. Furthermore, if the scaling region is consistent for different neighbourhood sizes and embedding dimensions, the average of the resulting fitted lines should be parallel to all linear increases. On the other hand, if the linear increase in $S(\Delta n)$ is too short or does not exist, then it must be obvious after comparing the $S(\Delta n)$ curve and the average of the resulting fitted lines. In order to restrict the search of the linear increase, the algorithm finds firstly the maximal value of $S(\Delta n)$ denoted by Δn_{peak} . For

$1 \leq \Delta n \leq \Delta n_{\text{peak}}$, the best fit (in the least-squares sense) of the lines is found for different sizes. The lines were increased in steps of $\lfloor \Delta n_{\text{peak}}/6 \rfloor$ for $\Delta n/7 \leq l \leq \Delta n/2$. Finally, one checks which is the most common interval in the abscissa covered by all lines and the lines containing this interval are used to calculate the average line.

The maximal Lyapunov exponent was calculated both from the raw data and from non-linearly filtered data in order to check whether the scaling region is more visible with filtered data. The non-linear filter described in Section 8.2 was used instead of the traditional FIR and IIR low-pass filters because traditional filters are based on the assumption that the signal and noise components can be distinguished in the spectrum. For coarsely sampled signals from non-linear systems this poses a problem since the signal itself can have a broadband spectrum and possible high-frequency components that must be preserved after filtering. See for details, Section 3.2.2. The TISEAN software, described in (Hegger et al., 1999), was used to calculate the maximal Lyapunov exponent based on the method of Kantz (1994) and to perform non-linear filtering based on the method of Schreiber (1993).

The correlation dimension was calculated using Eq. 3.9 with a modified version of the ANTA software and uniform embedding with dimensions $d_e - 1, d_e, d_e + 1, d_e + 2$. For the calculation of this invariant, the embedding dimension can be smaller than d_e . Instead of $d_e > 2d_c$ as in the reconstruction theorem, $d_e > d_c$ is sufficient for the correlation dimension estimates since self-intersections of the reconstructed set with itself have measure zero and hence do not distort this particular statistic (Sauer et al., 1991). The calculation of the correlation sum was done in three steps. First, a coarse calculation was performed for all desired embedding dimensions for $-4 < \log \epsilon < 2$. The standard deviation for each ϵ was calculated and the interval of ϵ with a standard deviations less than 10% of the maximum standard deviation was selected. Then, a fine calculation was performed for all desired embedding dimensions in the selected interval. The intercept with the abscissa of the best fit (in the least-squares sense) of a line with a length of 50% of the length of the selected interval was used as an estimate of the correlation dimension for each embedding dimension. Finally, the average of the correlation dimension among all embedding dimensions was used as a robust estimate.

For the rest of the calculations, only raw data was used. In order to confirm that irregularities of pedaling motion data are most likely due to deterministic chaos rather than random inputs, surrogate data hypothesis testing was applied to check the 4 null hypotheses mentioned in Section 4.1. These hypotheses contain, among other things, typical characteristics of random, linear, or periodic data. The surrogate data methodology consisted of the comparison between the discriminating statistic values (i.e. RPDE values) calculated from the surrogate and original data sets. The RPDE was calculated for the embedding dimensions $d_e, d_e + 1, d_e + 2, d_e + 3$, and $d_e + 4$ and with a radius proportional to the number of embedding dimensions and a proportion to the standard deviation of each time series, i.e. $\sigma/20$. This proportion was used for all tests and was the lowest proportion which produced stable results for an increasing number of embedding dimensions.

Two methods for the rejection of the 4 null hypotheses, explained in Section 4.1.2, were used in order to illustrate the effects of ignoring the standard deviation of the measurements: The ranking criterion proposed in (Theiler & Prichard, 1997), and the empirical probability density function. For the Theiler's method, a two-sided test is used, i.e. 39 sets of surrogate data were compared for a 95% confidence level. For the proposed method, 50 sets of surrogate data were compared. The surrogate data were created using a modified version of the algorithms in the ANTA software.

Like other dynamic invariants, estimates of the Lyapunov exponent and correlation dimension from time series may be affected by correlations between the reference points and the neighbours. Therefore, a minimum time equal to $d_e \tau$ was used for the calculations of these invariants.

4.2.2 Results

Tables 4.1 and 4.2 contain the results of the calculation of the embedding parameters (i.e. embedding dimension and embedding delay) and the selected dynamic invariants (i.e. maximal Lyapunov exponent and the correlation dimension). Tables 4.3 and 4.4 contain the hypothesis test results with the ranking criterion proposed in (Theiler & Prichard, 1997) and the empirical probability density function. See Section 4.1.2 for a description of these methods. In what follows, the evidence of chaos is analysed by parts.

Correlation dimension

For all data sets, the correlation sum was calculated with embedding dimensions $d_e - 1, d_e, d_e + 1$ and $d_e + 2$. The smaller the embedding dimension, the larger the scaling region in the curves of the correlation sum. Further, the local slopes of these curves, shown in Fig. 4.10, contained regions typical of chaotic dynamics similar to Fig 3.7 with three typical regions. At smaller scales, i.e. $\log(\epsilon) \leq -0.5$, the noise affected the correlation sum, increasing the estimation of the correlation dimension because the noisy data points were not confined to the attractor but separated from the attractor's trajectory, causing a wrong statistic for small ϵ . On larger length scales, i.e. for $-0.5 \leq \log(\epsilon) \leq 0.5$, the local scaling exhibited a plateau and is almost the same for all embedding dimensions larger than the correlation dimension. For a large ϵ , i.e. $\log(\epsilon) > 0.5$, the correlation sum did not exhibit a plateau in the local slopes (i.e. a linear scaling) since the macroscopic structures of the attractor determined the value of the correlation sum on this region and its value depended on both embedding dimension d_e and ϵ . The resulting curves, shown in Fig. 4.9, always had a drop for large ϵ , indicating that a chaotic dynamical system may be involved. For examples of chaotic systems with similar curves, see (Kantz & Schreiber, 2004).

The correlation dimensions d_c in all tests were in the range of $0.8 < d_c < 2$. As the ceil of the correlation dimension describes the number of variables that contribute to the dynamical system, the results showed that pedaling motion can be described as a 1- or 2-dimensional dynamical system.

Maximal Lyapunov exponent

The curves for the estimation of the maximal Lyapunov exponent were calculated using Eq. 3.8 for different embedding dimensions and neighbourhood sizes. One expected that a linear increase could be identified on those curves only when the time series exhibited low dimensional chaos with a reasonable noise level. In order to explore whether noise was the explanation for the absence of a linear region for some data sets, non-linear filtering was applied to the time series data. See Fig. 4.11 for results using raw data and Fig. 4.12 for results using filtered data.

Following the methodology described in Section 4.2.1, it was possible to see a distorted linear increase for the data sets 120N 80C P1, 120N 80C P3, 140N 90C P3 and 130N 85C P1, for both raw and filtered data. See their corresponding curves for raw data in Figures 4.11a, 4.11c, 4.11d and 4.9i; and for filtered data in Figures 4.12a, 4.12c, 4.12d and 4.9i. For the other data sets, one cannot draw any reasonable conclusion even after filtering. Most of the curves on both figures showed three typical regions but with a lack of linear increase. For small Δn , an initial jump was visible when the neighbourhood size ϵ was smaller than the noise, i.e. due to noise one found some 'false neighbours' for which the distances between nearby trajectories grew faster than the exponential divergence. The second region appeared with increasing Δn . Only when the exponential divergence dominated

Table 4.1: Embedding parameters and dynamic invariant obtained after processing raw data. Columns correspond to the name of the data set, the embedding dimension d_e , embedding delay τ , the maximal Lyapunov exponent λ , the correlation dimension d_c , and the recurrence period density entropy, respectively.

| test | d_e | τ | λ | d_c | RPDE |
|-------------|-------|--------|-------------------|-------------------|-------------------|
| 120N 80C P1 | 4 | 42 | 0.136 ± 0.005 | 1.012 ± 0.004 | 0.283 ± 0.010 |
| 120N 80C P2 | 4 | 39 | 0.095 ± 0.015 | 0.979 ± 0.029 | 0.234 ± 0.011 |
| 120N 80C P3 | 5 | 39 | 0.117 ± 0.013 | 0.943 ± 0.083 | 0.301 ± 0.008 |
| 130N 85C P1 | 4 | 41 | 0.012 ± 0.001 | 0.837 ± 0.034 | 0.287 ± 0.005 |
| 130N 85C P2 | 4 | 41 | 0.093 ± 0.006 | 1.009 ± 0.002 | 0.250 ± 0.014 |
| 130N 85C P3 | 4 | 43 | 0.035 ± 0.002 | 0.965 ± 0.101 | 0.334 ± 0.008 |
| 140N 90C P1 | 4 | 40 | 0.018 ± 0.007 | 1.012 ± 0.005 | 0.267 ± 0.009 |
| 140N 90C P2 | 4 | 40 | 0.113 ± 0.004 | 0.966 ± 0.090 | 0.248 ± 0.011 |
| 140N 90C P3 | 4 | 40 | 0.140 ± 0.022 | 0.908 ± 0.134 | 0.369 ± 0.004 |

Table 4.2: Embedding parameters and dynamic invariant obtained after processing filtered data. Columns correspond to the name of the data set, the embedding dimension d_e , embedding delay τ , the maximal Lyapunov exponent λ , the correlation dimension d_c , and the recurrence period density entropy RPDE, respectively.

| test | d_e | τ | λ | d_c | RPDE |
|-------------|-------|--------|-------------------|-------------------|-------------------|
| 120N 80C P1 | 4 | 42 | 0.088 ± 0.003 | 1.019 ± 0.005 | 0.287 ± 0.010 |
| 120N 80C P2 | 4 | 39 | 0.097 ± 0.017 | 1.011 ± 0.007 | 0.231 ± 0.011 |
| 120N 80C P3 | 5 | 39 | 0.120 ± 0.014 | 0.951 ± 0.073 | 0.301 ± 0.008 |
| 130N 85C P1 | 4 | 41 | 0.110 ± 0.067 | 0.941 ± 0.065 | 0.293 ± 0.006 |
| 130N 85C P2 | 4 | 41 | 0.096 ± 0.010 | 1.011 ± 0.005 | 0.248 ± 0.015 |
| 130N 85C P3 | 4 | 43 | 0.035 ± 0.002 | 0.874 ± 0.023 | 0.333 ± 0.008 |
| 140N 90C P1 | 4 | 40 | 0.015 ± 0.012 | 1.004 ± 0.013 | 0.265 ± 0.010 |
| 140N 90C P2 | 4 | 40 | 0.077 ± 0.003 | 0.860 ± 0.072 | 0.247 ± 0.012 |
| 140N 90C P3 | 4 | 40 | 0.139 ± 0.025 | 0.898 ± 0.072 | 0.369 ± 0.004 |

was a possible linear increase visible. Finally, for increasing Δn , $S(\Delta n)$ grew until it reached a saturation with possible oscillations. Tables 4.1 and 4.2 contain the results of the calculations.

Surrogate data analysis

The 4 null hypotheses mentioned in Section 4.1 were tested with surrogate data that were constrained to have the same obvious periodic components as the original data of the knee during pedalling. The RPDE was used as the discriminating statistic. Table 4.3 and Fig. 4.13 contain the results of the comparison of 39 surrogate data with original data and Theiler's method. Table 4.4 and Fig. 4.14 contain the results of the comparison of 50 surrogate data with the original data and the method of the empirical distribution. For Theiler's method, 4 time series had a strong evidence of chaos whereas for the method of the distribution only one time series had a strong evidence of chaos. The difference was in the testing results of hypothesis 3 which was rejected wrongly by the Theiler's method.

Table 4.3: Results of the comparison of 39 surrogate data with original data using RPDE as discriminating statistic and Theiler's method. When the value of the RPDE of the original data was the smallest or the largest than the 39 RPDE values of the surrogate data, the hypothesis was considered rejected and marked as 1. Otherwise, as 0. Columns correspond to the name of the test and the results for each of the 4 hypotheses.

| data set | Hyp. 1 | Hyp. 2 | Hyp. 3 | Hyp. 4 |
|-------------|--------|--------|--------|--------|
| 120N 80C P1 | 1 | 0 | 0 | 1 |
| 120N 80C P2 | 1 | 0 | 0 | 1 |
| 120N 80C P3 | 1 | 0 | 1 | 1 |
| 130N 85C P1 | 1 | 1 | 0 | 1 |
| 130N 85C P2 | 1 | 0 | 0 | 1 |
| 130N 85C P3 | 1 | 1 | 1 | 1 |
| 140N 90C P1 | 1 | 1 | 1 | 1 |
| 140N 90C P2 | 1 | 1 | 1 | 1 |
| 140N 90C P3 | 1 | 1 | 1 | 1 |

Table 4.4: Results of the comparison of 50 surrogate data with original data using RPDE as discriminating statistic and the method of the empirical distribution. If the area under the empirical probability density curve in the interval $\bar{x}_{\text{ori}} - \sigma_{\text{ori}} < x < \bar{x}_{\text{ori}} + \sigma_{\text{ori}}$ is smaller than 0.05, then the hypothesis is rejected and marked as 1. Otherwise, as 0. Columns correspond to the name of the test and the results for each of the 4 hypotheses.

| data set | Hyp. 1 | Hyp. 2 | Hyp. 3 | Hyp. 4 |
|-------------|--------|--------|--------|--------|
| 120N 80C P1 | 1 | 0 | 0 | 1 |
| 120N 80C P2 | 1 | 0 | 0 | 1 |
| 120N 80C P3 | 1 | 0 | 0 | 1 |
| 130N 85C P1 | 1 | 1 | 0 | 1 |
| 130N 85C P2 | 1 | 0 | 0 | 1 |
| 130N 85C P3 | 1 | 1 | 0 | 1 |
| 140N 90C P1 | 1 | 1 | 1 | 1 |
| 140N 90C P2 | 1 | 1 | 0 | 1 |
| 140N 90C P3 | 1 | 1 | 0 | 1 |

4.3 Conclusions

The method of surrogate data has become a central tool for validating the evidence of chaos. A dynamic invariant as discriminating statistic is calculated for both the original time series and the surrogate data sets. The dynamic invariant is normally calculated for different sets of parameters and their average value is used as the measure to compare the original and the surrogate data. Finally, a statistical criterion is used to determine whether the dynamic invariants of the original and surrogate are sufficiently different. If they are, the null hypothesis that the original and the surrogate data come from the same process is rejected.

Commonly, the statistical criterion known as the ranking criterion is used in the literature, which is a modified version of Student's t -test. However, this statistical criterion is incomplete and wrongly interpreted in many cases, because it does not include the error of the dynamic invariant measure. To improve the hypothesis testing, a novel method was suggested in this Chapter including the standard deviation from the calculation of the discriminating statistic and modifications of the parametric method described in Section 4.1.2. Both methods of testing hypotheses, the ranking criterion and the proposed method, were compared and three false rejections of one null hypothesis were found with the ranking criterion.

The hypotheses mentioned in Section 4.1 were not rejected for all data sets. However, one still cannot conclude that there is no evidence of chaos based on surrogate data. As was mentioned in Section 4.1, if we fail to reject the null hypothesis, our null hypothesis might be true. However, we shall also note the possibility that our discriminating statistic is not powerful enough to detect the difference between the surrogates and the original time series.

Only one dataset, 140N 90C P1, had a strong evidence of deterministic chaos based on surrogate data with both statistical criteria. Although the correlation sum curves, shown in Fig. 4.10, had a scaling region for a medium ϵ for all data sets indicating that a chaotic dynamical system may be involved, this data set did not have a clear linear increase for the calculation of the Lyapunov exponent for both the raw data and the non-linear filtered data. Future work must seek more powerful and robust discriminating statistics in order to check the evidence of chaos.

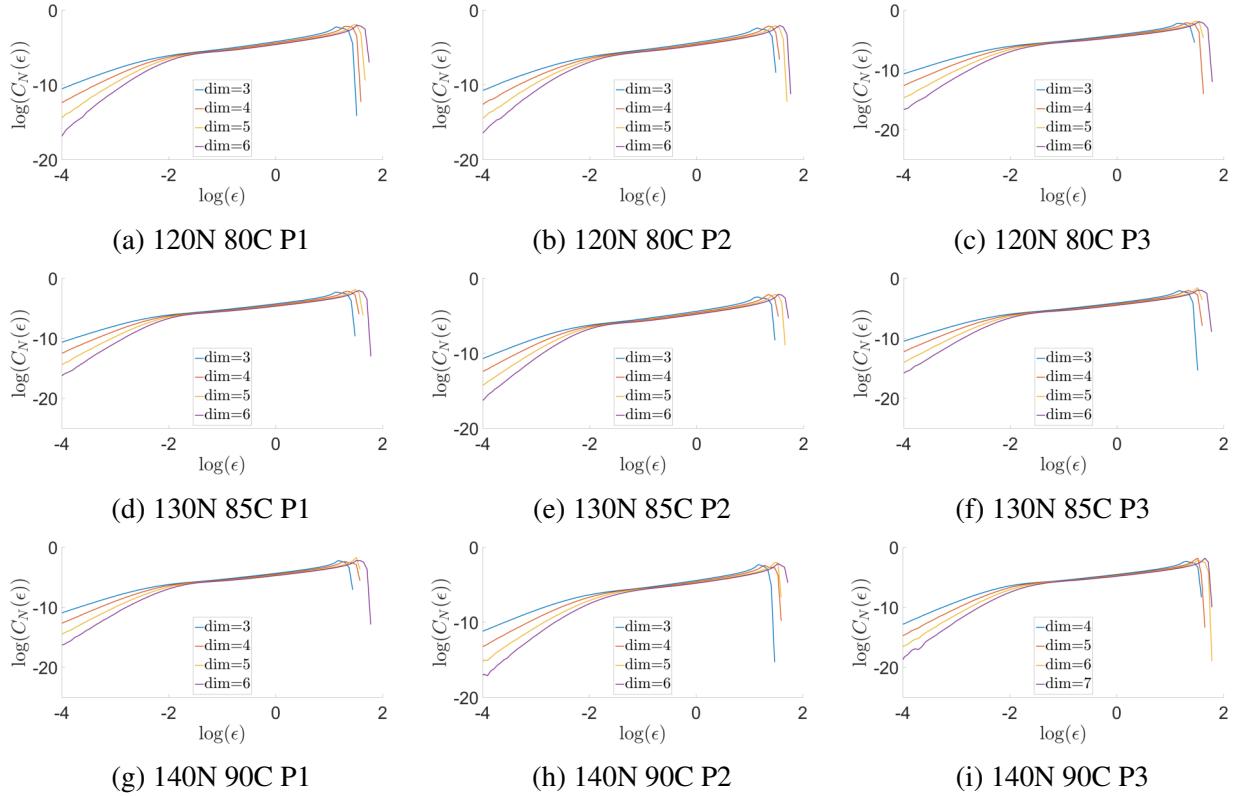


Figure 4.9: Curves for the calculation of the correlation dimension using raw data. The colored curves correspond to the results obtained with embedding dimension $d_e - 1$, d_e , $d_e + 1$ and $d_e + 2$. The resulting curves, shown in Fig. 4.9, had always a scaling region for a medium ϵ , indicating that a chaotic dynamical system may be involved.

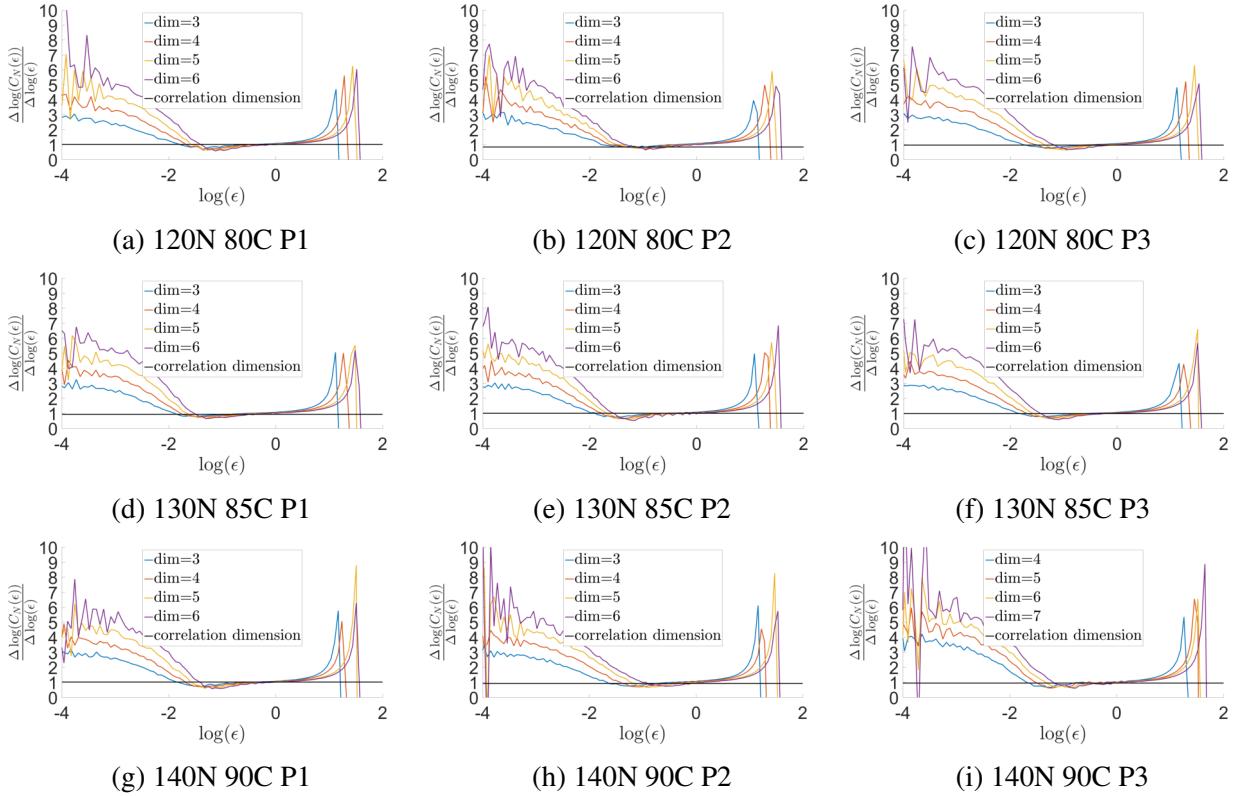


Figure 4.10: Local slope of the curves for the calculation of the correlation dimension using raw data. Each figure contains 5 curves. The colored curves correspond to the results obtained with embedding dimension $d_e - 1$, d_e , $d_e + 1$ and $d_e + 2$. The black line is the mean correlation dimension. The curves contain three regions which are typical for chaotic dynamics. See the text for details.

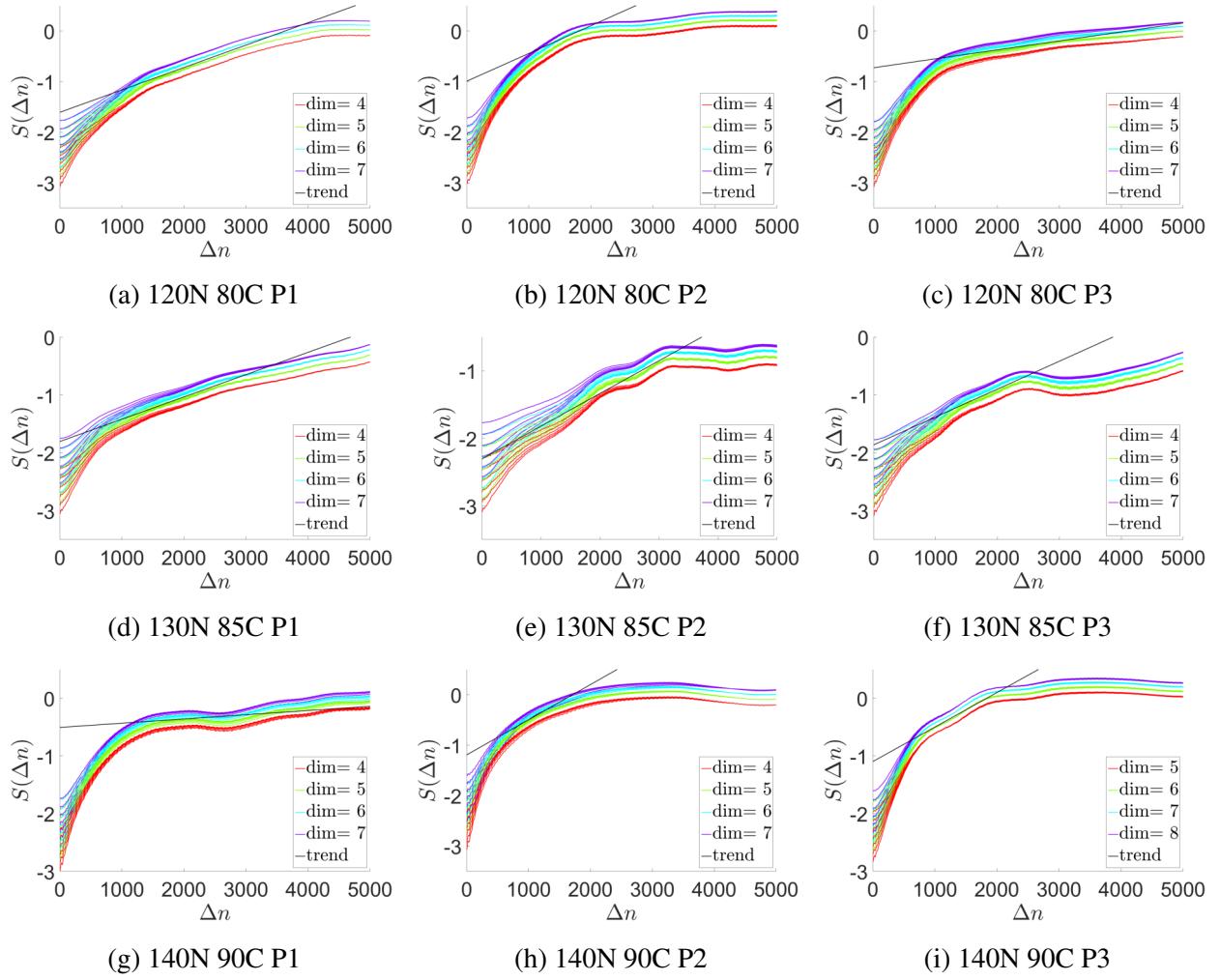


Figure 4.11: Curves for the calculation of the maximal Lyapunov exponent from raw data for each data of the 9 data sets.

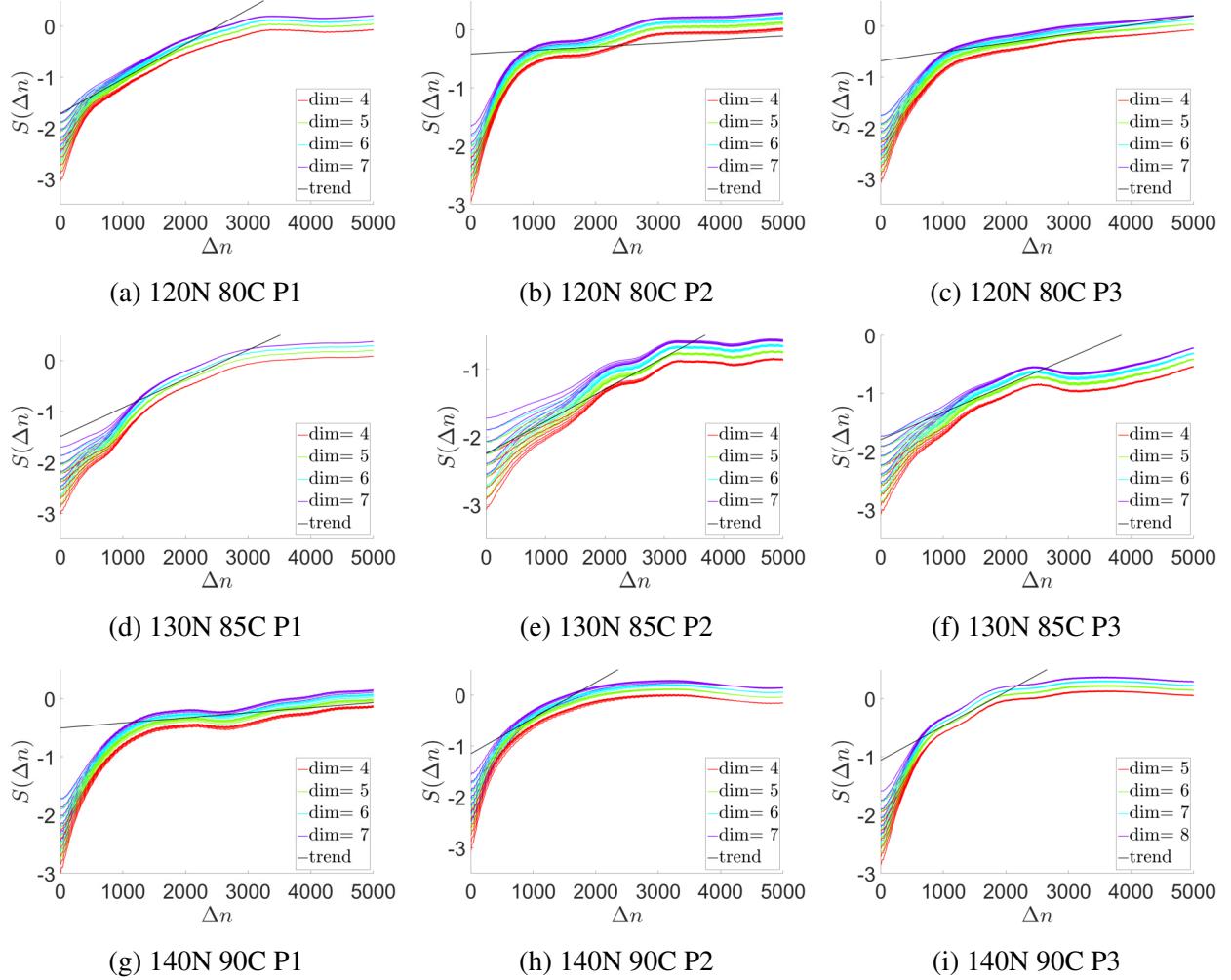


Figure 4.12: Curves for the calculation of the maximal Lyapunov exponent from non-linearly filtered data. For the non-linear filter, the embedding dimension was selected as twice the required by the embedding theorems and the embedding delay was equal to 1 and the ϵ was selected as three times the noise level.

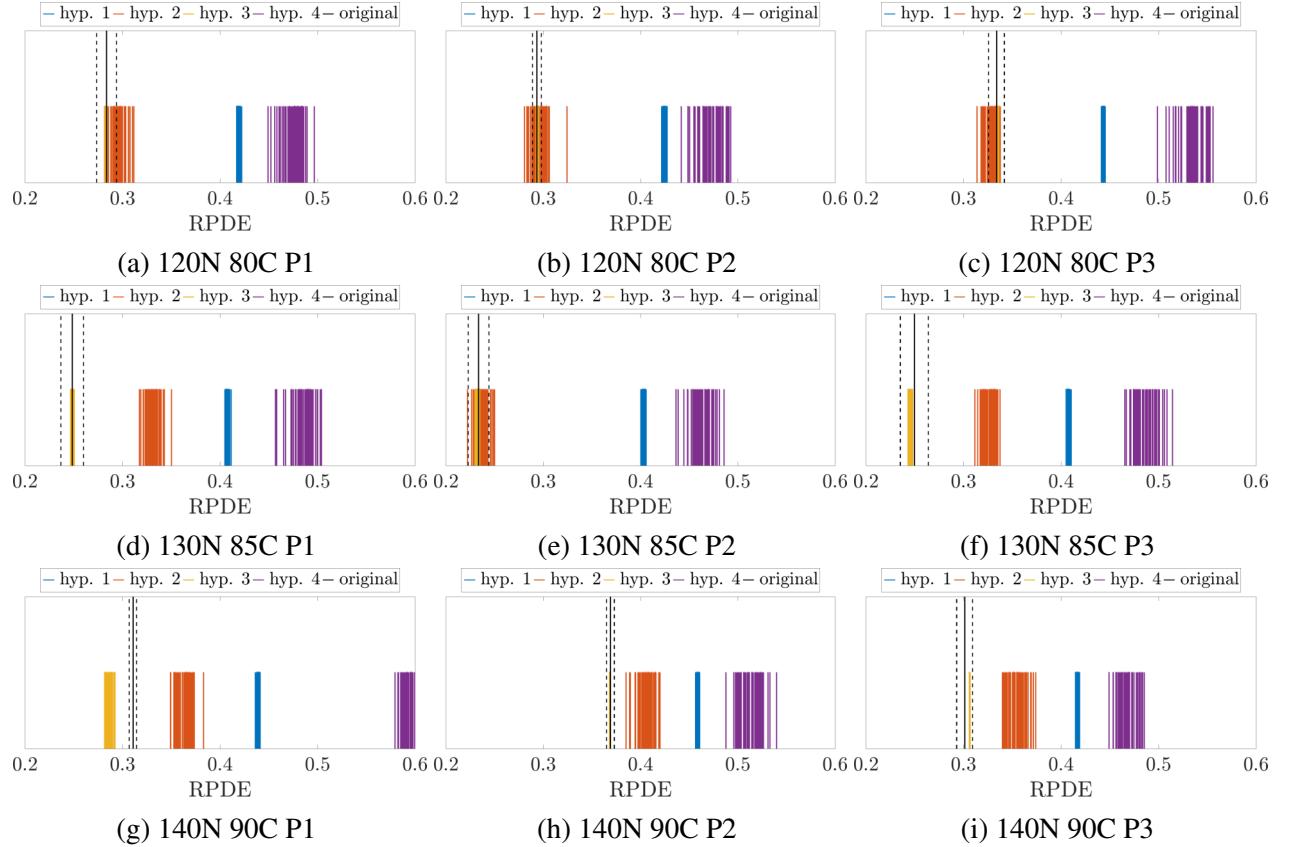


Figure 4.13: Results of hypothesis rejection using the method proposed in (Theiler & Prichard, 1997). The black line and the dashed black lines represent the mean and the standard deviation of the discriminant statistic values of the original data, respectively. The x is the value of the discriminating statistic. Notice that the Theiler's method rejects all null hypotheses for 130N 85C P3, 140N 90C P1, 140N 90C P2 and 140N 90C P3.

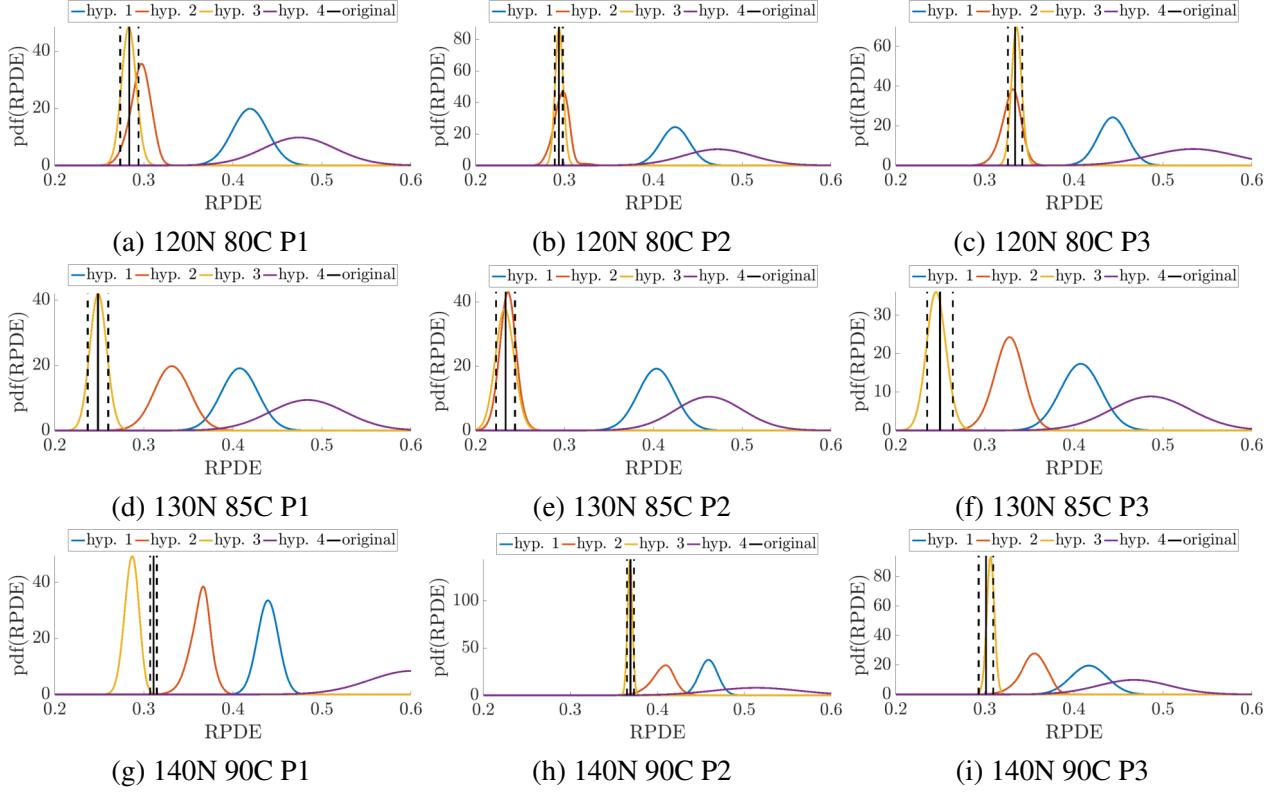


Figure 4.14: Results of hypothesis rejection using our proposed method calculated from one original time series of the noisy positional data described in Section 4.2. The black line and the dashed black lines represent the mean and the standard deviation of the discriminant statistic values of the original data, respectively. The x is the value of the discriminating statistic. Notice that our method rejects all null hypotheses only for 140N 90C P1.

Chapter 5

Parametrization of cyclic motion and self-organizing maps

Cyclic motion is at the core of many sports, such as running, swimming, or cycling. The study of the corresponding kinematic (e.g. joint angle), kinetic (e.g. ground reaction force), spatio-temporal (e.g. stride interval) and electromyographic variables is fundamental to the evaluation of training routines and the assessment of performance. These variables are often presented in the form of single-cycle curves, representing a time-varying value over one complete cycle.

Cyclic motion is the result of the continuous adjustment of the timing between muscular, neurological, and respiratory systems, according to environmental, bio-mechanical, and morphological constraints. Therefore, single-cycle curves are not exact copies of a single prototypical curve but a family of curves, each one slightly different from the others. That is, amplitude and phase variations can be found within samples of single-cycle curves of motion.

Conventional motion analysis derives features from a few single-cycle curves and from their corresponding prototypical average curve. The features are extracted from recognizable phases presented in normal motion. For instance, the maximum peaks can be calculated for each of the typical phases in gait motion shown in Fig. 5.1.

Computing cross-sectional averages over a family of misaligned curves without registration can lead to the cancellation of critical shape characteristics (Kneip & Gasser, 1992), modifications of landmarks¹ (Chau et al., 2005), and to inflated amplitude variability estimates (Sadeghi et al., 2003). Thus, the calculation of the prototypical average curve requires that the single-cycle curves must first be registered. See Fig. 5.2 for an example.

Curve registration is the alignment of curves by minimizing the discrepancies from an iteratively estimated sample mean or by alienation specific curve landmarks. When the phase shift is constant across the time scale, it is easy to correct it using a landmark registration procedure. However, many data sets require more complicated transformations of the time axis in order to align landmarks like the peaks and valleys. The transformations of the time axis, so called time-warping functions, may be constants, linear functions, or non-linear functions. See (Ramsay, 1997) for a description of standard registration techniques.

Frequently, the analysis of single-cycle curves and their corresponding prototypical average curve is performed manually by experts. However, their conclusions may not agree because of the difficulty in comprehending large amounts of information, different levels of expertise that might lead to different interpretations, etc. There is a clear need for a method which overcomes these

¹A landmark is a feature with a location that is clearly identifiable in all curves (Ramsay, Hooker, & Graves, 2009).

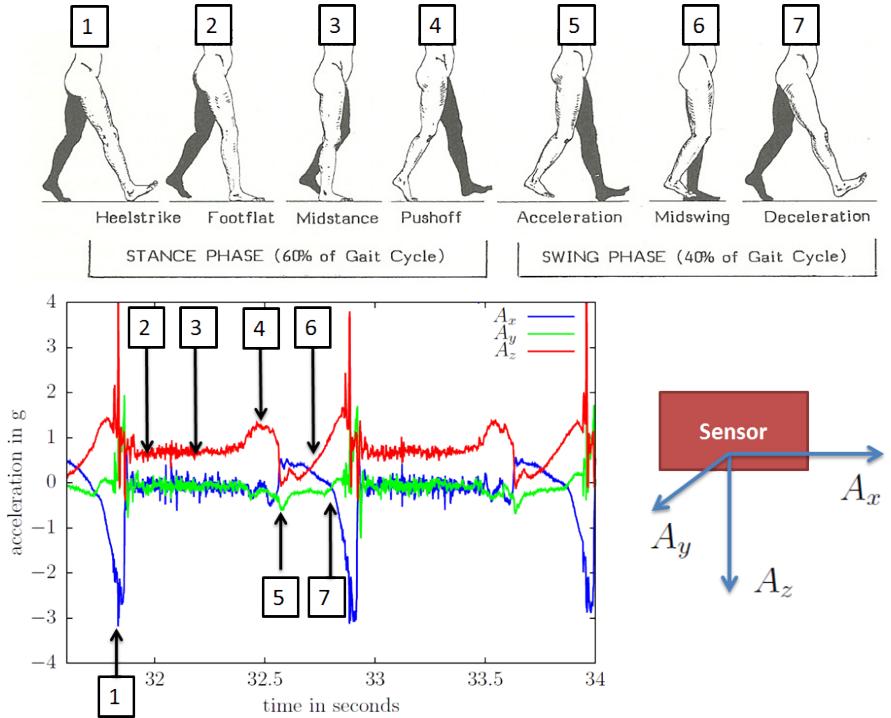


Figure 5.1: Top: Walking phases. Figure adapted from Schafer (1987). Bottom: Acceleration curves obtained from a device with three accelerometers placed on the right foot during walking. In a normal gait pattern the peaks can be used for cycle segmentation.

limitations and enables large quantities of data to be analysed and interpreted objectively (Barton, Lees, Lisboa, & Attfield, 2006). Curve registration can be used advantageously, in conjunction with other methods such as neural networks, to address such challenges as the clustering, classification and visualization of high dimensional data collected from motion.

The self-organizing map (SOM), described in (Kohonen, 1990), is a non-supervised algorithm that reduces the complexity of high dimensional data to usually two-dimensional regular spaced grids. Humans can easily interpret the two-dimensional representation of the data, e.g. visualization using SOM was applied to gait stem-patterns in (Barton et al., 2006) and to young–old gait patterns in (Begg & Kamruzzaman, 2003). Furthermore, one can build simple local or nearest-neighbour models on top of the SOM, e.g. recognition of gait patterns (Koktas, Yalabik, & Yavuzer, 2006) and of emotions in gait (Janssen et al., 2011).

This chapter is based on our publications (Quintana-Duque, Saupe, & Vieten, 2014; Quintana-Duque & Saupe, 2015), which included results from this thesis research, and presents our novel methods for curve registration and cycle averaging as well as the benefits when they are combined with SOM. Section 5.1 describes our motivation for the methods for curve registration and cycle averaging. Section 5.2 introduces our method for cycle averaging, called equalized DBA (eDBA). Section 5.3 describes the novel approach to phase registration using eDBA average cycle and the proposed quality assessment of the transversal sections. Section 5.4 describes concepts related with SOM and its topological quality. Finally, Section 5.5 shows the results of parametrizing cycles based on eDBA and a comparison of the quality and the classification performance of the trained SOM with and without data preprocessing using the phase registration of cycles.

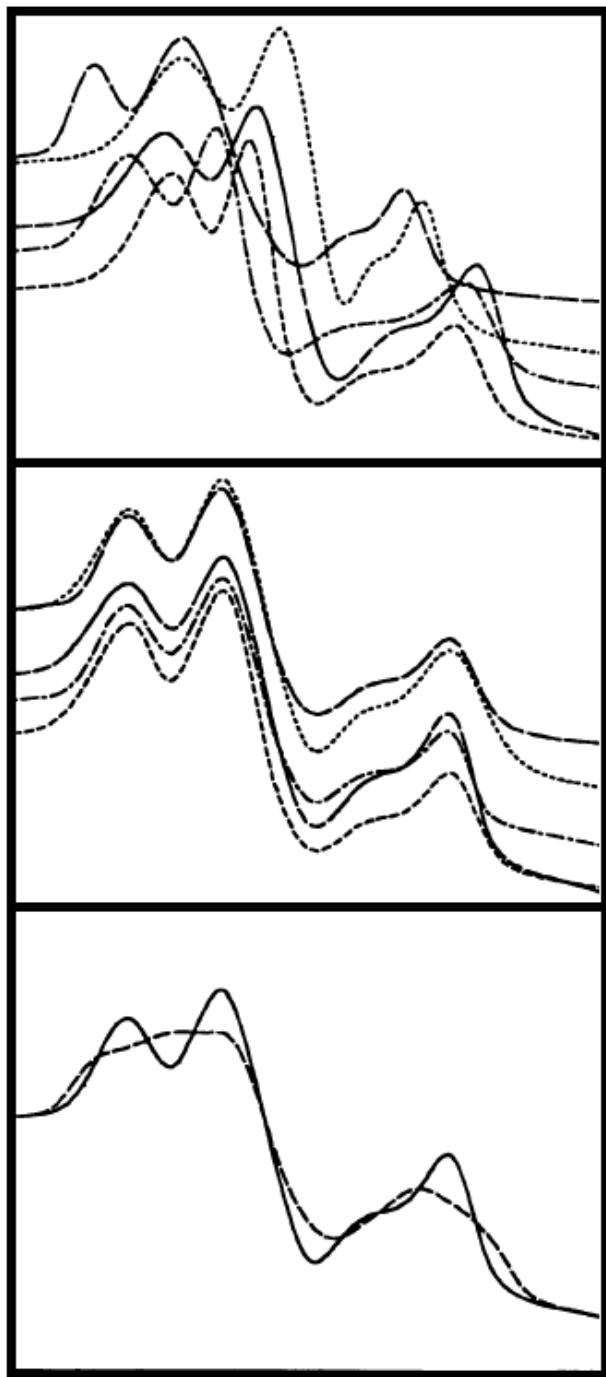


Figure 5.2: Examples of averaging with and without registration. Top: Five curves from a synthetic process are shown which differ in dynamics and amplitude. Middle: Transformed curves after curve registration. Bottom: Cross-sectional average (dashed curve) versus average after curve registration (solid curve). Without curve registration, the peaks and the curve structure of the average curve can be smeared or might even disappear. Figure modified from (Kneip & Gasser, 1992).

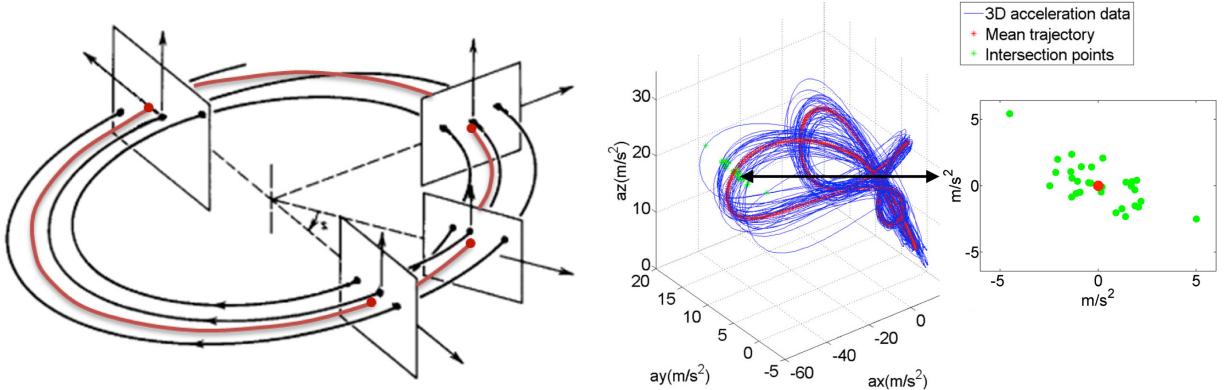


Figure 5.3: Transversal sections. Left: Schematic plot of parameterized transversal (Poincaré) sections belonging to a periodic trajectory (red). Figure modified from (Moon, 2008)). Middle: 3D plot of a cyclic trajectory of acceleration data (blue), the periodic mean trajectory (red), and a transversal section (green points). Right: 2D close-up view of the transversal section.

5.1 Analysis of multi-dimensional variables with transversal sections

Most current methods for the analysis of multi-dimensional variables of cyclical motion are based on the individual analysis of single variables, ignoring that the multidimensional space of all observed variables may reveal important additional insights into the patterns of motor control. However, an approach to the analysis of multi-dimensional variables of cyclical motion is to view them as a (reconstructed) state space and to intersect the cyclic orbit with a co-dimension 1 hyperplane, called the transversal section (Vieten, Sehle, & Jensen, 2013).

The transversal section and the analysis of intersections are similar to some tools for the analysis of dynamical systems near a periodic solution. The transversal section corresponds to the so-called Poincaré sections and the intersections with one transversal plane correspond to the so-called Poincaré map. Here we will use the names Poincaré sections and transversal sections interchangeably.

Poincaré maps are commonly employed to get insight into the phase space of a dynamical system. A Poincaré map is the intersection of an orbit in the state space of a continuous dynamical system with a certain lower-dimensional subspace, called the Poincaré section or Poincaré plane, transversal to the flow of the system (Grimaldi & Manto, 2012). See some examples in Fig. 5.3. With the Poincaré map, a periodic trajectory becomes a point, while a non-periodic trajectory is represented by a set of points (i.e. the intersections of the orbits and the Poincaré plane). The points on the Poincaré plane preserve many of the properties of the periodic and quasiperiodic orbits of the original system. In pure dynamical systems, the information in the Poincaré maps at different phases are equivalent despite the chaotic nature of the underlying problem. In embedding space reconstruction from a noisy time series, the information in the maps depends on the quality of the reconstruction of the (original) state space and the adequate placement of the Poincaré section.

Vieten et al. (2013) assume that the space constructed by the observed variables is enough to distinguish changes in long time series recorded from cyclic motion. The transversal section is placed according to an average curve instead of one orbit of the limit cycle attractor as in a dynamical system. Thus, the results obtained using this method strongly depend on the selection of the cycle average which defines the phase and the position of its corresponding transversal section.

Further, the phase variation between cycles can lead to spurious statistical results. In particular, individual realizations of movements with a typical shape show different dynamics and intensity.

In order to provide a tool for cycle registration that facilitates the analysis of cycles in cyclical motion data, we proposed in (Quintana-Duque et al., 2014) a novel method, called equalized DBA (eDBA), to calculate the average of a set of cycles based on dynamic time warping (DTW) and a modification of DTW barycentric averaging (DBA). The eDBA allows studying the features of the kinematic variables in cyclic motion depending on the phase using the eDBA average cycle as the reference for phase registration. Features can be extracted from the intersection points on each transversal section to characterize the overall motion. For example, the variance in one selected section may indicate the degree of regularity of the cyclic motion. The eDBA is explained in what follows.

5.2 The equalized DBA (eDBA) method

Cycles in motion data can have different durations because repetitions of a motion pattern are executed with different accelerations and decelerations. Thus, pointwise averaging of unaligned curves may not adequately represent the set of sequences under study.

The calculation of an average cycle is the first step for phase registration, data analysis by traditional statistics, and novel methods based on transversal sections. The first goal of this contribution is to align all K cycles so that they are parameterized by a phase from $[0, 2\pi]$. In this section, the eDBA algorithm is defined mathematically, its properties are discussed, and possible applications are proposed for feature analysis in cyclic motion depending on the phase.

5.2.1 Data and assumptions

The eDBA method is concerned with a set of smooth single-cycle curves of observed motion variables which can be considered as independent realizations. Such curves often possess a typical structural pattern common to all the curves of the sample but with different dynamics and intensity.

Let $x_M[n], n = 0, 1, \dots$ be a d -dimensional time series of measurements in \mathbb{R}^d , uniformly sampled from continuous motion in the state space \mathbb{R}^d and including measurement noise. For applications in motion science we assume that the motion is band limited at about 12 Hz, the sampling rate is above the Nyquist rate of 24 Hz, and the motion is supposed to be cyclic with a fundamental frequency of about 1 Hz, as, e.g. in typical human gait. To reduce the noise and to smooth the time series $x_M[n]$, we apply a Savitzky–Golay low-pass filter, explained in Section 8.3, with a cutoff frequency of 12 Hz.

The cycles should be readily segmented, e.g. by placing segment boundaries at pronounced extremal values of one of the d time series components. Then we segment the data obtaining K cycles $x_{M,k}[n], n = 0, \dots, N_k - 1, k = 0, \dots, K - 1$. The cycles may hold differing numbers of samples, N_k , and so we uniformly resample these sequences with the same number of samples, N . This yields a set of K cycles, each with N samples, $x_k[n] = 0, 1, \dots, N - 1, k = 0, \dots, K - 1$.

5.2.2 Calculation of the average cycle based on eDBA

The averaging is accomplished by a modification of the DTW barycentric averaging (DBA) algorithm (Petitjean, Ketterlin, & Gançarski, 2011), called equalized DBA (eDBA). In the following, these algorithms are explained.

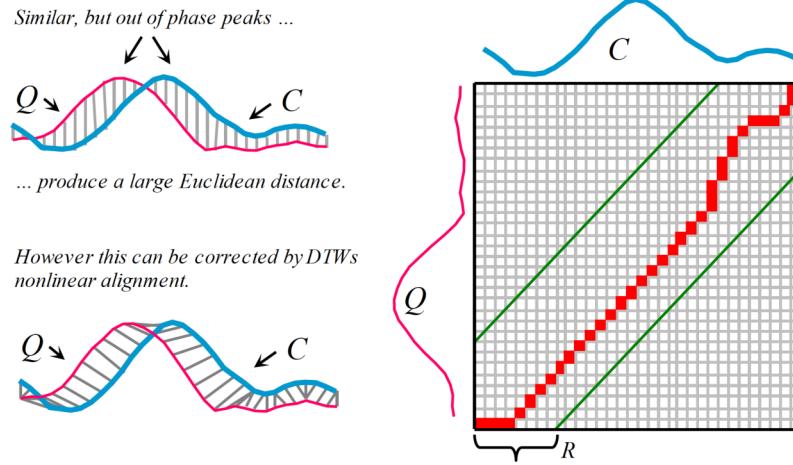


Figure 5.4: Left: Two time series which are similar but out of phase. Right: To align the sequences we construct a warping matrix, and search for the optimal warping path (red/solid squares). Note that a Sakoe–Chiba Band with width R is used to constrain the warping path (Rakthanmanon et al., 2012).

Dynamic time warping (DTW)

The dynamic time warping (DTW) is a well-known technique to quantify the dissimilarity and to find an optimal alignment between two given (time-dependent) d -dimensional sequences, $Q[n]$ and $C[m]$, $n = 0, \dots, N - 1$ and $m = 0, \dots, N - 1$.

DTW warps two sequences in a non-linear fashion to match each other. An alignment is given by two monotonically increasing index sequences s_0, \dots, s_M and t_0, \dots, t_M with $s_0 = t_0 = 0$, $s_M = t_M = N - 1$, and $0 \leq s_{k+1} - s_k \leq 1$, $0 \leq t_{k+1} - t_k \leq 1$ for $k = 0, \dots, M - 1$, implying that the samples $Q[s_k], C[t_k]$, $k = 0, \dots, M$ are pairwise aligned. A time warping incurs a cost $\sum_k \delta(Q[s_k], C[t_k])$, where the local cost is usually taken as the square difference of the time series values, i.e. $\delta(Q[i], C[j]) = (Q[i] - C[j])^2$ in the case of scalar time series. For d -dimensional time series, this generalizes to the square Euclidean norm of the difference $\delta(Q[i], C[j]) = \|Q[i] - C[j]\|^2$.

To align two sequences using DTW, an $N \times N$ warping matrix D is constructed, with the (i^{th} , j^{th}) element of the matrix being the cost $\delta(Q[i], C[j])$, where $D(0, 0) = 0$ and $D(N - 1, N - 1)$ is the overall similarity measure. The optimal warping path containing the associated elements between $Q[n]$ and $C[m]$ is calculated from the warping matrix D minimizing the sum of costs for the alignment. See an example in Fig. 5.4. The overall similarity measure (i.e. DTW distance) can be recursively computed by

$$D(m, t) = \delta(Q[m], C[t]) + \min \begin{cases} D(Q[m - 1], C[t - 1]), \\ D(Q[m - 1], C[t]), \\ D(Q[m], C[t - 1]). \end{cases} \quad (5.1)$$

Pseudocodes for the calculation of the DTW distance matrix and the alignment are given in Algorithms 1 and 2, respectively.

Algorithm 1 DTW_Distance_Matrix(Q, C)

Input: Q : array with size $N \times d$

Input: C : array with size $N \times d$

Output: D : the accumulated cost matrix of DTW

```
1:  $D = \text{empty array with size } N \times N$ 
2: for  $i = 1$  to  $N$  do  $D[i, 0] = \infty$ 
3: for  $i = 1$  to  $N$  do  $D[0, i] = \infty$ 
    $D[0, 0] = 0$ 
4: for  $i = 1$  to  $N$  do
5:   for  $j = 1$  to  $N$  do
6:      $D[i, j] = \delta(Q[i], C[j]) + \min(D[i - 1, j], D[i, j - 1], D[i - 1, j - 1])$ 
return  $D$ 
```

Algorithm 2 DTW_multiple_alignment(x_a, x)

Input: x_a : the sequence for which the alignment is computed

Input: x : the sequence to align to using DTW

Output: alignment: the vector with values of x after alignment

Step 1: compute the accumulated cost matrix of DTW

```
1: cost = DTW_Distance_Matrix( $x_a, x$ )                                 $\triangleright$  This part changes with eDBA
   Step 2: store the elements associated to
2:  $L = \text{length}(x_a)$ 
3:  $\text{alignment} = [\emptyset, \dots, \emptyset]$                                       $\triangleright$  array of  $L$  empty sets
4:  $i = \text{rows}(\text{cost})$                                                $\triangleright$   $i$  iterates over the elements of  $x_a$ 
5:  $j = \text{columns}(\text{cost})$                                               $\triangleright$   $j$  iterates over the elements of  $x$ 
6: while ( $i > 1$ )&&( $j > 1$ ) do
7:    $\text{alignment}[i] = \text{alignment}[i] \cup x[j]$ 
8:   if  $i == 1$  then
9:      $j = j - 1$ .
10:  else if  $j == 1$  then
11:     $i = i - 1$ 
12:  else
13:     $\text{score} = \min(\text{cost}[i - 1][j - 1], \text{cost}[i][j - 1], \text{cost}[i - 1][j])$ 
14:    if  $\text{score} == \text{cost}[i - 1][j - 1]$  then
15:       $i = i - 1$ 
16:       $j = j - 1$ 
17:    else if  $\text{score} == \text{cost}[i - 1][j]$  then
18:       $i = i - 1$ 
19:    else
20:       $j = j - 1$ 
return  $\text{alignment}$ 
```

DTW barycentric averaging (DBA)

A characteristic average time series of a set of time series is an essential part for analysis and for many learning algorithms such as k -means clustering. As mentioned before, computing cross-sectional averages over a family of misaligned curves can lead to the cancellation of critical shape characteristics and landmarks (Kneip & Gasser, 1992). Thus, the sequences must be aligned in order to calculate an average resembling the original structure. However, finding the multiple alignments of a set of sequences, or its average sequence, is a typical chicken-and-egg problem: knowing the average sequence provides multiple alignments and vice versa (Petitjean et al., 2014).

The DTW similarity measure is probably the most used and useful tool to analyse sets of sequences. Unfortunately, computing the average of a collection of time series in a way that is consistent with DTW is nontrivial, because as the average is unknown, so too is the temporal alignment of each time series to the average. Many attempts at finding an averaging method for DTW have been proposed but their results depend on the order of the cycles. See for examples (Niennattrakul & Ratanamahatana, 2007; Hautamäki, Nykänen, & Fränti, 2008).

The DTW barycentric averaging (DBA) presented in (Petitjean et al., 2011) is a heuristic strategy, designed as a global averaging method based on a local optimization algorithm similar to clustering methods like k -means. DBA iteratively refines an average sequence $x_a[i]$ with $i = 0, \dots, N - 1$ from the set of sequences to average $S = \{x_0, \dots, x_k, \dots, x_{K-1}\}$ in order to minimize the total cost given by the sum of all squared costs (i.e. DTW distances) for the DTW alignments of S with the average series x_a . For each iteration, DBA works in two steps:

1. Consider x_a fixed and compute DTW between each sequence in S and the temporary average sequence x_a to be refined, in order to find associations between sample points of the average sequence and sample points of the set of sequences.
2. Updated all sample points i of the average series x_a by the arithmetic mean, i.e. barycenter with all weights equal to one, of all corresponding samples of the given series S found in the previous step with DTW. Note that a sample point of one of the sequences may contribute to the new position of several sample points of the average. Conversely, any sample point of the average is updated with contributions from one or more sample points of each sequence.

Pseudocodes for the calculation of the DBA average are given in Algorithms 3 and 4.

Modifications of DBA: The eDBA method

DBA is considered currently as the state-of-the-art method to average a set of sequences consistently with DTW. However, the DBA algorithm yields sharp peaks of the average cycle that are not characteristic of the underlying motion but can be regarded as DBA artifacts. See Fig. 5.5 for an example. The DBA artifacts are caused by two problems in the averaging step of DBA:

1. Many samples of a given cycle may be aligned with the same sample of the average cycle. The latter strongly pulls the sample of the average cycle towards the given cycle, yielding the spurious peaks. This first problem was identified in the red line in Algorithm 4.
2. It is in the nature of DTW that several consecutive samples may be aligned with the same sample of the average cycle leading to the same phase value in the final registration process. However, such a piecewise constant phase is not appropriate for continuous motion from physical processes such as in sports. This second problem was identified in the red line in Algorithm 2.

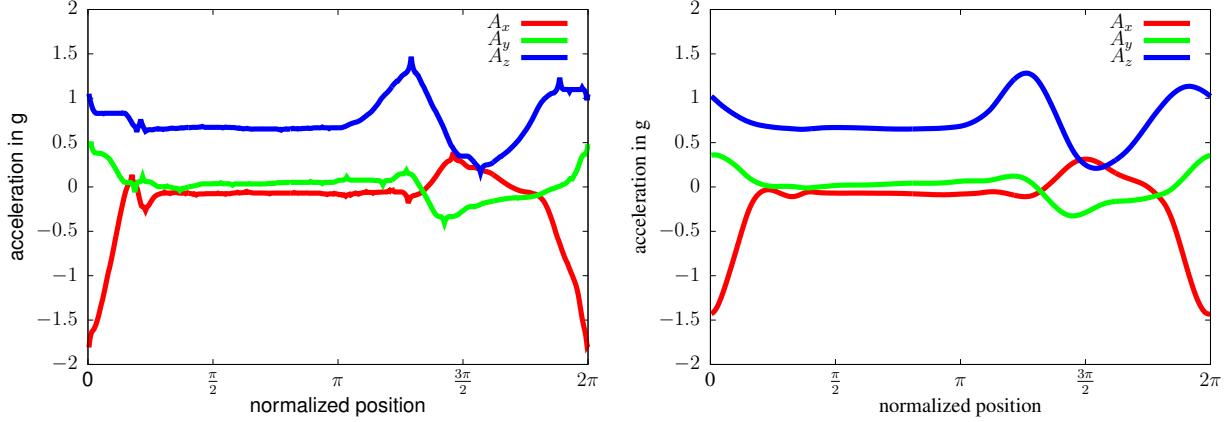


Figure 5.5: Average results of 3D walking acceleration data, i.e. (A_x, A_y, A_z) of a sensor placed on the foot, using DTW barycentric averaging (DBA) shown on the left figure and equalized DBA (eDBA) shown on the right figure.

Algorithm 3 DBA(S, I, x_{ini})

Input: S : the set of sequences to average
Input: I : the number of iterations
Input: x_{ini} : the initial average sequence
Output: x_a : the refined average sequence

- 1: $i = 0$
- 2: $x_a = x_{\text{ini}}$
- 3: **while** $i < I$ **do**
- 4: $x_a = \text{DBA_update}(x_a, S)$
- 5: $i = i + 1$

return x_a

Algorithm 4 DBA_update(x_a, S)

Input: x_a : the average sequence to refine (of length L)
Input: S : the set of sequences to average
Output: x_a : the refined average sequence (of length L)

Step 1: compute the multiple alignment for x_a

- 1: $\text{alignment} = [\emptyset, \dots, \emptyset]$ ▷ array of L empty sets
- 2: **for all** x in S **do**
- 3: $\text{alignment_for_x} = \text{DTW_multiple_alignment}(x_a, x)$
- 4: **for** $i = 0$ to L **do**
- 5: $\text{alignment}[i] = \text{alignment}[i] \cup \text{alignment_for_x}[i]$

Step 2: compute the multiple alignment for the alignment

- 6: **for** $i = 0$ to L **do**
- 7: $x_a = \text{mean}(\text{alignment}[i])$ ▷ arithmetic mean on the set. This part changes with eDBA

return x_a

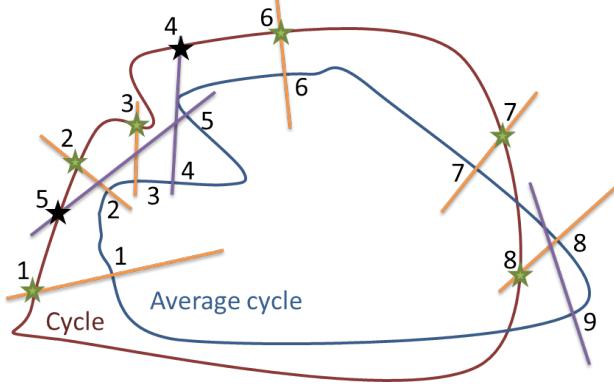


Figure 5.6: For noisy real-world data Poincaré sections may not be well-defined: Nine points, labeled 1 to 9, on the average cycle are taken as origins of planar Poincaré sections. The corresponding intersections may not be in sequence (point 5 is before point 2) and the section for point 9 does not intersect with the given cycle at all.

Our solution to the first problem is to introduce weights for the averaging such that each cycle is taken into account with the same weight. Our solution to the second problem is to include a regularization term in the cost function of DTW to penalize piecewise constant phase given by

$$\delta(i, j) = (1 - \lambda) \frac{(\|x_a[i] - x_k[j]\|^2)}{E} + \frac{\lambda(i - j)^2}{T}. \quad (5.2)$$

where $x_a[i]$ is the average sequence, $x_k[j]$ is the cycle to align, E and T are normalization constants. We used $E = 1/N \sum_{n=0}^{N-1} \|x_a[n]\|^2$ and $T = (N/4)^2$. The parameter $\lambda \in [0, 1]$ determines the trade-off between the classical DBA ($\lambda = 0$) and plain arithmetic averaging ($\lambda = 1$). For $\lambda = 0.5$, on average the two terms contribute about the same costs in the DTW algorithm.

5.3 Approach to phase registration using eDBA average cycle and transversal sections

The second goal of this contribution is to align all K cycles so that they are parameterized by a phase from $[0, 2\pi]$. An average periodic cycle, $x_a[n], n = 0, \dots, N - 1$, is constructed from all K cycles. This periodic signal defines the phase as $2\pi n/N$ for the n -th sample $x_a[n]$ of the average cycle. Each individual cycle is aligned to the average cycle by dynamic time warping (DTW) so that if $x_k[m]$ registers with $x_a[n]$, then the phase for $x_k[m]$ is defined as $2\pi n/N$.

The time series x_k can be registered by DTW with the average eDBA cycle x_a so that to each sample $x_k[j]$ there is attributed the phase of the corresponding point $x_a[j]$ on the mean cycle. For each phase value $\phi = 2\pi i/N, i = 0, 1, \dots, N - 1$ let the set A_ϕ denote the set of all points from the entire motion trajectory that are assigned to the phase ϕ . By design of the DTW algorithm, the sets A_ϕ need not be disjoint and may contain several points of a single cycle $x_k[n]$. Such cases are artifacts resulting in physically unrealistic phase discontinuities and time intervals of constant phase, respectively.

Consider the transversal section at phase $2\pi i/N$ given by intersections of the motion trajectory with the normal plane at the sample $x_a[i]$ of the average cycle, i.e. the linear subspace that is anchored at $x_a[i]$ and orthogonal to the tangent of the average cycle at $x_a[i]$. We imagine that

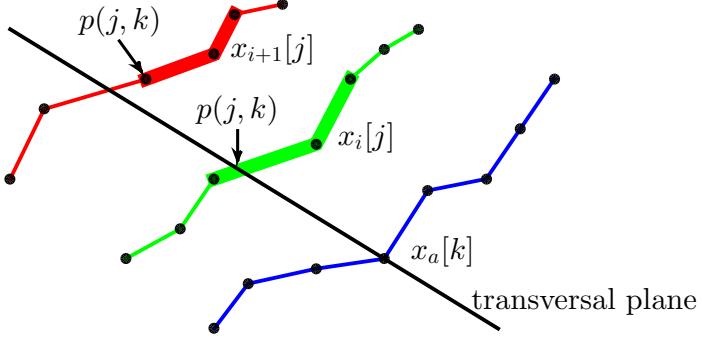


Figure 5.7: Examples of intersection (green line) and no intersection (red line) with the transversal section (black line) for the point in the average curve $x_a[k]$. An intersection $p(j, k)$ between the plane and the curve exists in $x[j - 1] \leq p(j, k) \leq x[j + 1]$. If there is no intersection, $p(i, j)$ is taken to be the point on the segment closest to the plane, i.e. one of the three segment end points.

the Poincaré section is then pushed forward to the next sample $x_a[i + 1]$ with updates of the corresponding intersections. However, it is not necessarily the case that (a) the intersections for the next Poincaré section correspond to points of cycles that stem from the future instead of the past, and that (b) such intersections always exist, see Fig. 5.6. To address these problems we propose using the DTW algorithm to compute the Poincaré sections. To this end, we apply the cost function

$$\delta(i, j) = (1 - \lambda) \frac{\|x_a[i] - p(i, j)\|^2}{E} + \lambda \frac{(i - j)^2}{T}, \quad (5.3)$$

where $p(i, j)$ denotes the intersection of the polygonal line from $x_k[j - 1]$ to $x_k[j]$ and on to $x_k[j + 1]$ with the normal plane at $x_a[i]$, i.e. transversal section, if an intersection exists. In principle, the DTW algorithm enforces the desired monotonicity of alignment, and here, it limits the first problem (a) to small deviations. If there is no intersection, $p(i, j)$ is taken to be the point on the segment closest to the plane, i.e. one of the three segment end points. If two intersections exist, then the $p(i, j)$ is taken to be closest intersection to $x_a[k]$. Thereby, although an intersection could not be found, at least a nearby sample point is taken, ameliorating the second problem (b). See Fig. 5.7 for examples of intersection and no intersection with the transversal section.

5.3.1 Quality of transversal sections

The Poincaré map is the intersection of flow data in the state space with a certain lower dimensional subspace, called the Poincaré section, which is transversal to the flow of the system. For deterministic systems, i.e. systems governed by differential equations without noise, all sections are equivalent, conveying the same information. However, in practice, measurements are noisy, and outlier cycles may lack intersections with some of the transversal sections as was pointed out in Fig. 5.6. Therefore, the placement of the transversal section is of high relevance for the usefulness of the results.

When analysing the dynamics of movement by means of transversal sections in observed space or Poincaré sections in embedding space, the question arises, which section to take, as any one of the sections belonging to phase parameters in $[0, 2\pi]$ may be used. Hegger et al. (1999) suggested selecting the section that maximizes the number of intersections. The sections are preselected by

the zero crossing of the temporal derivative of one of the variables, i.e. collecting all maxima or all minima, respectively. However, this method suffers from noise since for small time derivatives additional extrema can be produced by perturbations.

An optimal section of the transversal section should maximize the signal-to-noise ratio (SNR) of the resulting intersection points since the noise level is dependent on the section. We suggest using the SNR to quantify the quality of a section given by

$$\text{SNR}(\phi) = \log_{10} \frac{\text{variance}(\phi)}{\text{noise}^2(\phi)}. \quad (5.4)$$

To calculate the SNR, let A_ϕ denote the set of points in the transversal section at phase ϕ . Some of these points may be outside the planar section. The covariance matrix captures the variance of the signal at phase ϕ . The signal energy in the transversal section is given by the sum of the eigenvalues of the matrix. We divide this energy by the average power of the noise at the points, estimated by the squared Euclidean distances between the measurements and the smoothed data. An $\text{SNR} = 0 \text{ dB}$ indicates that the noise has the same power as the signal. This implies that motion analysis based on transversal sections should be such that the corresponding SNR is maximal, and at least greater than zero.

Another requirement for valid analyses of transversal sections should be that the section set A_ϕ contains only intersections of the motion trajectory with the corresponding normal plane to the average trajectory. Therefore, we suggest monitoring the percentage of points in A_ϕ that are true intersections and selecting only sections with 100% intersections. An example based on acceleration data is shown in Fig. 5.14 with our proposed quality method.

5.4 Self-organizing map (SOM)

The self-organizing map (SOM) is an unsupervised neural network algorithm which implements a non-linear projection of multi-dimensional data onto a low dimensional array of vectors, i.e. a finite set of neurons. The low-dimensional discretized representation of the input space of the training samples is called a map, and it should represent meaningful entities in the data. The mapping process retains the topology of the input data after training, thereby revealing potential unknown correlations between input parameters, which can provide important insights into the data.

Typically, an SOM has two layers, known as the input and output layers. The neurons in the output layer can be mapped to the input layer as model vectors. See Fig. 5.8 for an example with 2D data. Each model vector of the SOM is determined as a weighted average of all of those training vectors that are mapped into the topological neighbourhood around the corresponding model (Kohonen, Nieminen, & Honkela, 2009). Thus, self-organizing maps learn both the distribution and topology of the input vectors they are trained on.

For SOM, it is assumed that a map with a sufficient number of neurons conveys information about the input data sets in a clear manner. However, the appropriate choice of the number of neurons is a problem in SOM. The right selection of the parameters of the map (e.g. the number of neurons, dimensions of the map grid, map lattice and shape) is a key issue due to the trade-off between the representation accuracy and the avoidance of over-fitting. Normally, different sets of parameters are tested and the quality of the resulting maps after training is compared to select the best one. Some quality measures of SOM are explained in the next section.

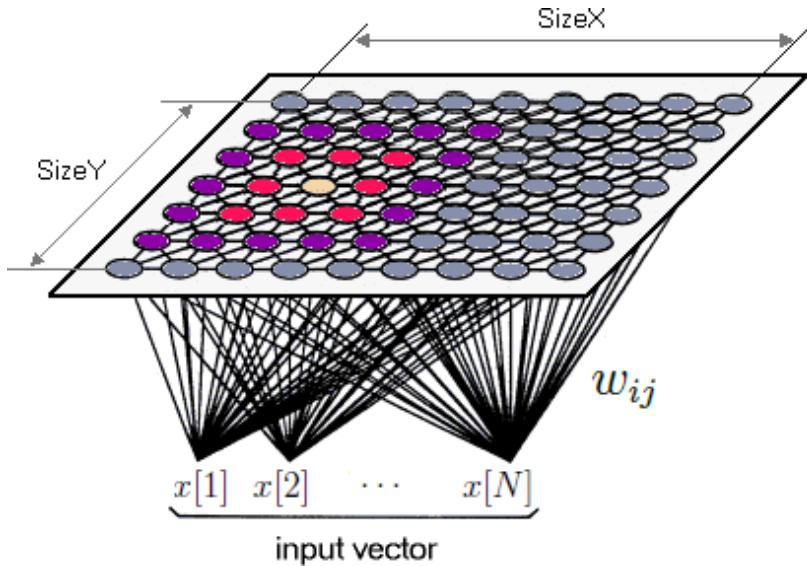


Figure 5.8: Example of a self-organizing map (SOM). The SOM is a neural network that maps a large training set via a process of competitive learning from a high dimensional input space spanned by $x[1], \dots, x[n]$ to a two-dimensional array of nodes (i.e. neurons) in the output space. The learning process adjusts the weights w_{ij} . Figure modified from (Joglekar, 2015).

5.4.1 Quality of self-organizing maps

After the SOM has been trained, it is important to know whether it has properly adapted itself to the training data. Map quality measures have been proposed for evaluating the precision of the mapping, the preservation of the topology, or a combination of both. Most of the quality measures are based on the closest model vector in the map to an input vector, e.g. the best matching unit (BMU). For surveys of quality measures for SOM, see (Pöhlbauer, 2004; Mayer, Neumayer, Baum, & Rauber, 2009). For this dissertation, the following three error measures provided in (Vatanen et al., 2015) were applied to describe the quality, i.e. the smaller the error, the better the quality.

1. Mapping precision: The neurons in the output space can be mapped in the input space as reference vectors. The mapping precision error is defined as the average quantization error (i.e. Euclidean distance) between data vectors and their first BMUs on the map in the input layer. See Fig. 5.9a.
2. Topological representation accuracy: The mapping process retains the topology of the input data after training, thereby revealing potential unknown correlations between input parameters, which can provide important insights into the data. The topological representation accuracy error is the simplest way to assess the preservation of the topology. It is measured as the proportion of data vectors for which the first- and second-BMUs are not adjacent units in the array topology. See Fig. 5.9b.
3. Topographic preservation average: In this chapter, the combined error measure proposed in (Kaski & Lagus, 1996) is used for the topographic preservation quality. The combined error is the average error over all input vectors of the sum of the accuracy of the map in representing the input data set (i.e. quantization error) and the distance from the best-matching

unit to the second-best-matching unit of each data vector along the shortest path following the neighbourhood relations (i.e. topographic error). If the map is defined by means of a graph, where the edge lengths are the distances between the nodes and the graph nodes are the model and input vectors, the topographic error is computed using Dijkstra's algorithm (Dijkstra, 1959) from the input vector to the second-BMU passing by the first-BMU. See Fig. 5.9c.

5.5 Results

5.5.1 Experimental Setup

Eight healthy adults (7 male, 1 female, age 26 ± 2.6 years old, weight 62 ± 7.2 kg) took 12 tests of 3 minutes each on a treadmill (walking, jogging, running). The protocol of the tests is given in Table 5.1. We varied the treadmill's slope, speed, and whether or not 1.5 kg weights were attached to the ankles. Before the recording session, the subjects practiced for a few minutes on the treadmill at different speeds in order to get familiar with the device. The test sequence was identical for all subjects. One minute of rest was given after 9 minutes of tests.

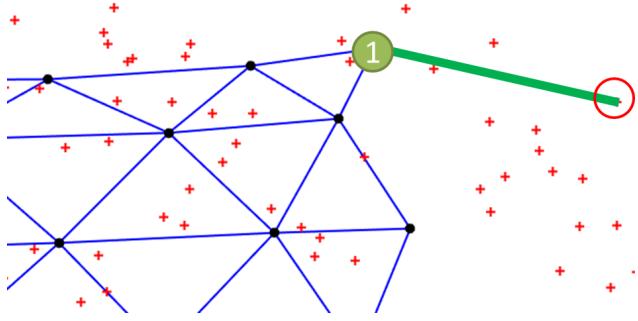
The motion data was acquired by a 3D accelerometer (RehaWatch, Hasomed) attached to the right shoe. The fastening of the sensor to the shoe was done by a special foot mounting with velcro. The sensor yielded time series A_x , A_y , A_z of acceleration (in units of g, acceleration by gravity) at a sampling rate of 600 Hz. See Fig. 5.10 for the direction of the acceleration vectors (A_x , A_y , A_z).

5.5.2 Data processing

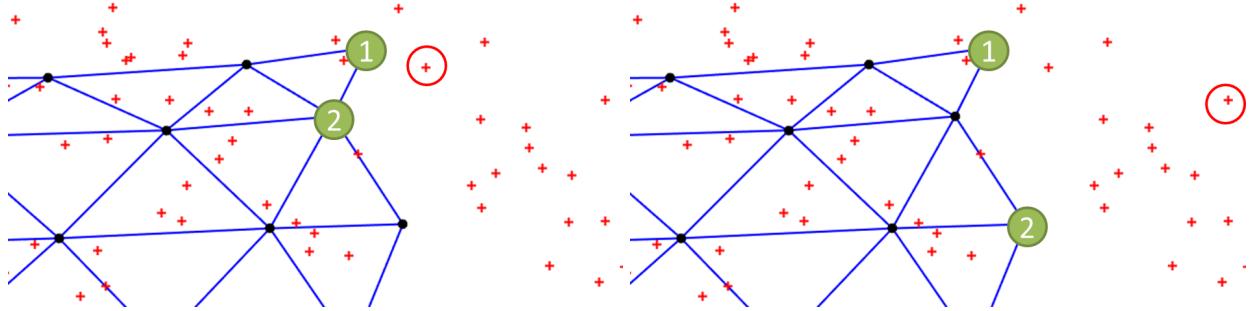
The noise in the 3D acceleration data was reduced by a Savitzky–Golay filter (Savitzky & Golay, 1964), i.e. by least squares regression on a sliding window with quadratic polynomials and a

| no. | activity | time | speed | slope | weights |
|-----|----------|-------|----------|-------|---------|
| 1 | walking | 3 min | 2.5 km/h | 0% | — |
| 2 | walking | 3 min | 2.5 km/h | 5% | — |
| 3 | walking | 3 min | 2.5 km/h | 10% | — |
| | rest | 1 min | | | |
| 4 | jogging | 3 min | 5.0 km/h | 0% | — |
| 5 | jogging | 3 min | 5.0 km/h | 5% | — |
| 6 | jogging | 3 min | 5.0 km/h | 10% | — |
| | rest | 1 min | | | |
| 7 | running | 3 min | 7.5 km/h | 0% | — |
| 8 | running | 3 min | 7.5 km/h | 5% | — |
| 9 | running | 3 min | 7.5 km/h | 10% | — |
| | rest | 1 min | | | |
| 10 | walking | 3 min | 2.5 km/h | 0% | 3 kg |
| 11 | walking | 3 min | 2.5 km/h | 5% | 3 kg |
| 12 | walking | 3 min | 2.5 km/h | 10% | 3 kg |

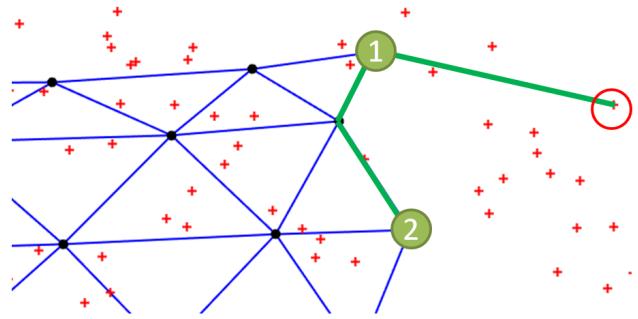
Table 5.1: Test protocol for each subject



(a) Mapping precision accuracy is defined as the quantization error between data vectors and their BMUs. In the example, 2D input vectors are represented by red crosses, reference vectors are represented by black points and neighbourhood connections given the topology in the output space are represented by the blue lines. The distance between the input vector in the red circle and its BMU is represented as the green line.



(b) Topological representation accuracy is defined as the percentage of the input data vectors for which the first- and second-BMUs are not adjacent units. On the left picture, the first- and second-BMUs are adjacent units. For the right picture, the first- and second-BMUs are not adjacent units. For both pictures, the 2D input vectors are represented by red crosses, the reference vectors are represented by black points and the neighbourhood connections given the topology in the output space are represented by the blue lines. Further, the first- and second-BMUs are represented by the numbers 1 and 2, respectively.



(c) The combined error considers the quantization and the topographic error. The 2D input vectors are represented by red crosses, the reference vectors are represented by black points and the neighbourhood connections given the topology in the output space are represented by the blue lines. Further, the first- and second-BMUs are represented by the numbers 1 and 2, respectively. The distance from the input vector in the red circle to the second-BMU passing by the first-BMU is represented as the green line.

Figure 5.9: Examples with 2D input data of some quality definitions for self-organizing map (SOM).

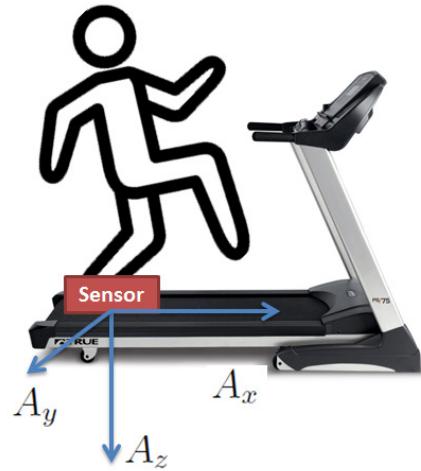


Figure 5.10: Data are the measurements of the 3D acceleration vectors of one sensor (A_x , A_y , A_z).

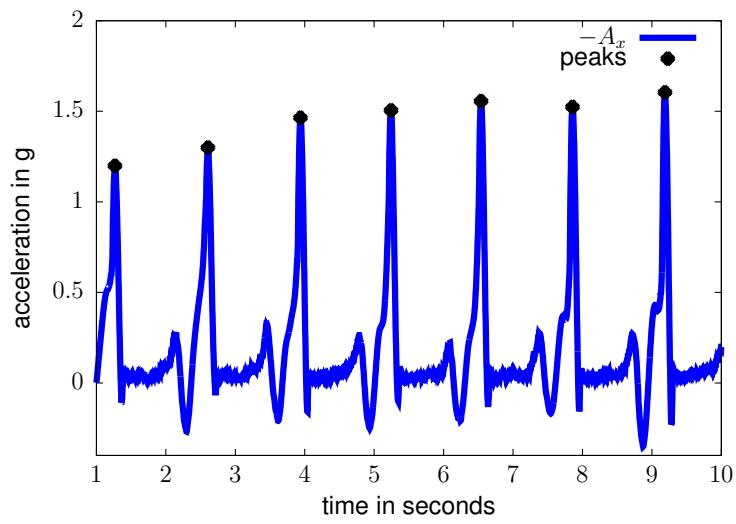


Figure 5.11: The cycle segmentation was performed using extremal values in A_x .

window of 91 samples, yielding a filter with a cutoff frequency of 12 Hz. For the SNR calculation, the difference between the original and the filtered signals was taken as noise.

The gait cycles could clearly be segmented by extremal values in A_x , designating the instant of heel strike in each cycle. Fig. 5.11 shows an example of the segmentation points for each cycle. In this figure and in all collected data, the cycles have different amplitudes and durations. Then, each cycle was Fourier transformed and upsampled by trigonometric interpolation to 512 samples per cycle. The first 20 cycles of each test were discarded to remove the transition from stance to swing motion.

For each single gait activity, i.e. walking without weights, jogging, running, and walking with weights, the average eDBA was calculated with $\lambda = 0.5$. The cycles of each activity were registered using their corresponding average eDBA. Furthermore, in order to test the effects of a unique average for all activities, we calculated the average eDBA of all activities together and the cycles were registered with this average. For an example of the results, see Fig. 5.12.

5.5.3 Parametrization and quality of transversal sections

The eDBA allows computing a representative average cycle and the study of kinematic variables depending on the phase using the eDBA average cycle as the reference for phase registration. It may be of interest to study how the kinematic variables change with the phase angle of the cyclic motion. Features characterizing the overall motion can be extracted either from the representative average directly or from the intersection points on each transversal section. For instance, Fig. 5.5 shows the average cycle of gait acceleration data and Fig. 5.13 shows the variance as a function of phase for two activities and two persons.

The intersection points on the plane preserve many properties of periodic and quasiperiodic orbits of the (original or reconstructed) state space of the system. In pure dynamical systems, the information in the Poincaré maps at different phases are equivalent despite the chaotic nature of the underlying problem. In the embedding space reconstruction from a noisy time series, the information in the maps depends on the quality of the reconstruction of the (original) state space and the adequate placement of the Poincaré section. However, a selected phase may be enough to characterize the overall motion. See Fig. 5.14 for an example of our quality method to select the best transversal section. The green curve corresponds to the percentage of intersections and the violet curve corresponds to the SNR curve against the phase. An analysis of the quality of the transversal sections computed using our eDBA and registration algorithm shows that for large phase intervals there are cycles that do not intersect properly with the section as required. For the remaining intervals, the figure shows the SNR. For many phases the SNR is negative, indicating that the measurement noise dominates the signal, rendering such transversal sections useless for motion analysis. The figure also shows the preferred transversal section with maximal SNR where the circles indicate the noise of the samples.

5.5.4 Parametrization and self-organizing maps

Methods

The experimental setup and the data processing are described in Sections 5.5.1 and 5.5.2. For each single gait activity (i.e. walking without weights, jogging, running, walking with weights) the average eDBA was calculated with $\lambda = 0.5$. The cycles of each activity were registered using their corresponding average eDBA. Furthermore, in order to test the effects of a unique average for all

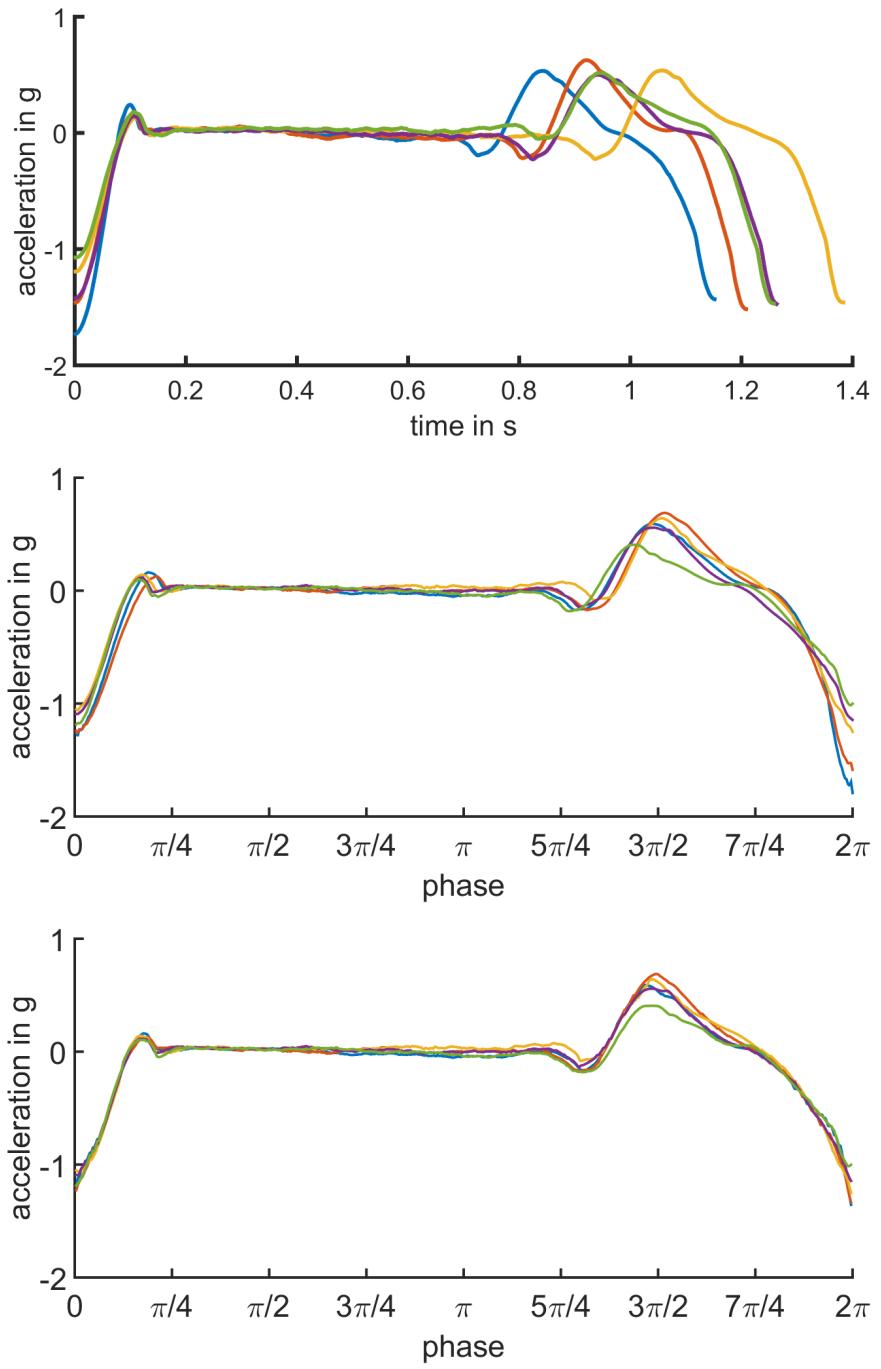


Figure 5.12: Example of registration using our eDBA approach with acceleration data from our tests. Top: Cycles after segmentation in time. Middle: Resampled cycles without registration. Bottom: Resampled cycles after registration using equalized DBA (eDBA) approach.

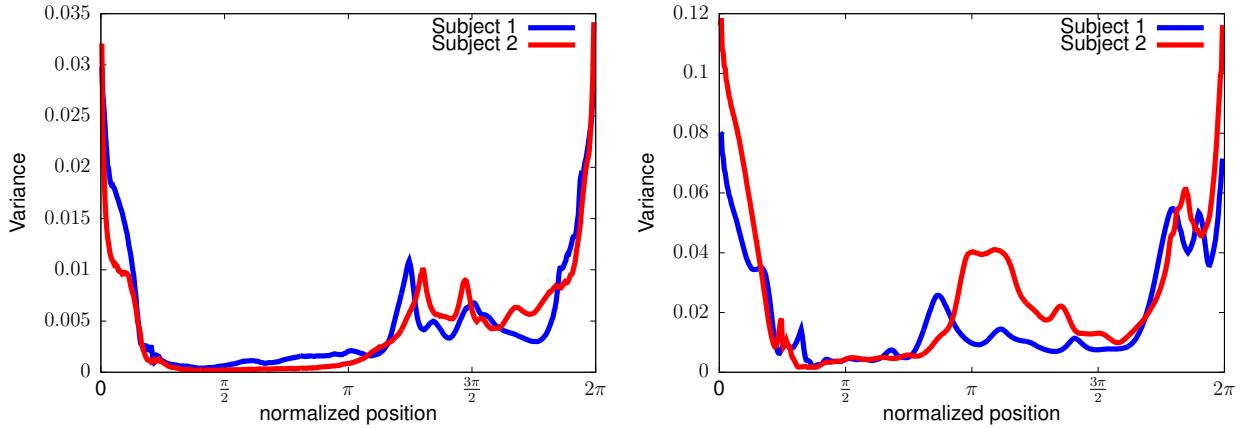


Figure 5.13: Time warped and aligned cycles may be used for within-stride analysis. Here the multidimensional variance is shown as a function of phase for two participants for walking (left) and running (right). The graph shows that it is possible to discriminate different persons as well as different activities from the average cycles.

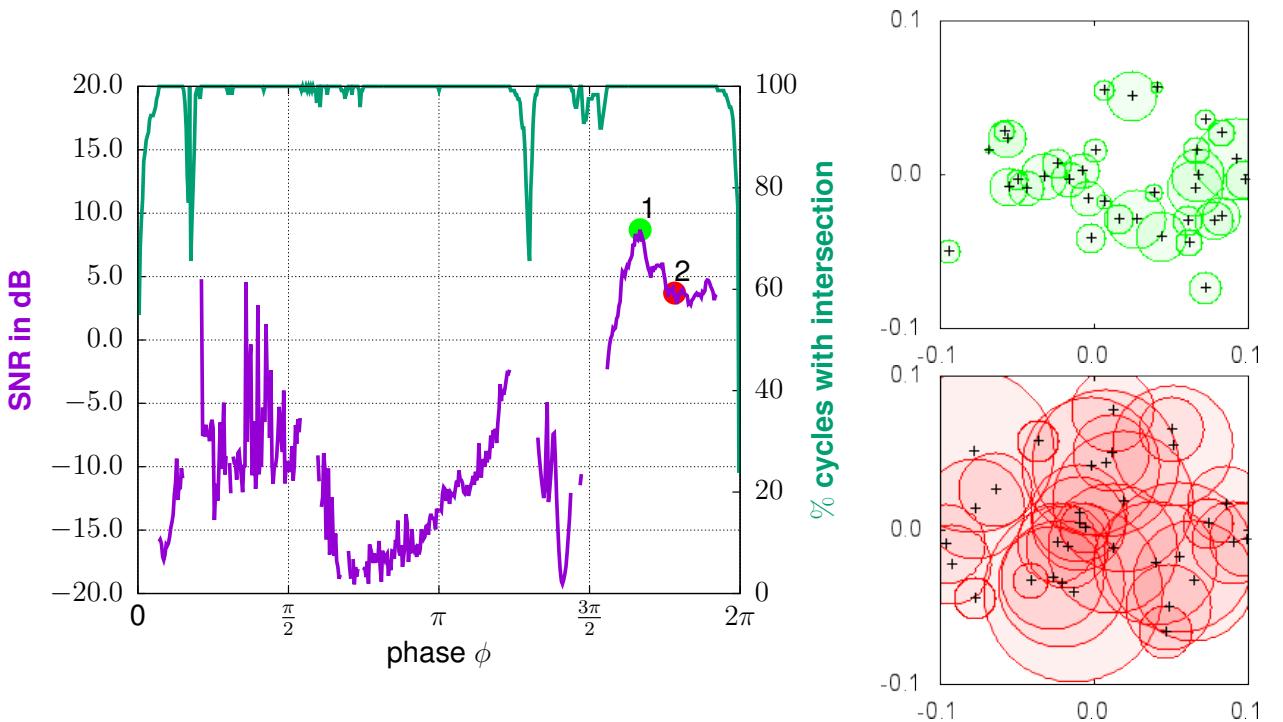


Figure 5.14: Left: Quality of Poincaré sections given by the SNR (violet curve) and the percentage of intersections (green curve). When the percentage of intersections was lower than 100% the SNR curve was not shown. Right: noise of the intersections between the cycles and two selected Poincaré sections.

activities we calculated the average eDBA of all activities together and the cycles were registered with this average.

Cycle classification was performed using SOM with sizes $n \times n$ and $n = 2, 4, \dots, 40$. The number of neurons and their topological relations were fixed from the beginning of each training: 2-dimensional map with hexagonal lattice and a sheet form was used for all tests. A sheet form means that the map sides were not connected to each other. Before the training, initial values were given to the model vectors by a linear initialization. The training was done with a batch algorithm. The input data was composed of the concatenated acceleration vectors (A_x, A_y, A_z) of each cycle. The maps and three quality measures were calculated using the MATLAB toolbox provided in (Vatanen et al., 2015).

In the following, we compare the classification performance, the representation accuracy and the topological quality of each SOM as a function of the number of neurons in the SOM. We trained the SOMs using the acceleration data from each single activity separately (i.e. walking without weights, jogging, running, walking with weights), and together.

Classification performance

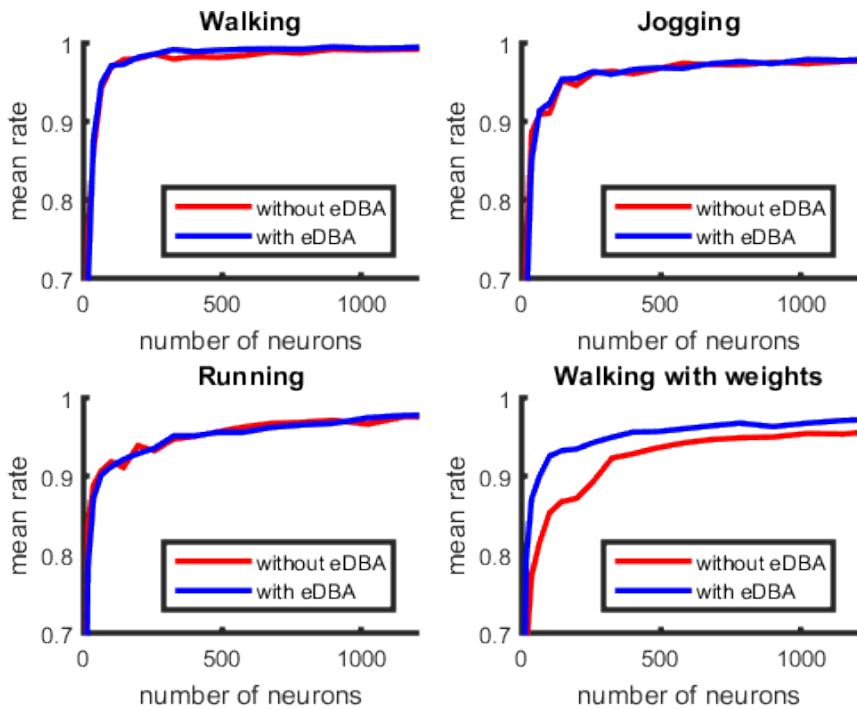
Walking, jogging and running are considered as distinct gait modes with different mechanics and energetics. Although they share some basic kinetics and kinematics, the gaits are also distinctly different. Walking gait may be defined by the existence of a double support phase during stance, whereas running has a ‘flight’ phase during which neither limb is in contact with the ground (Cappellini, Ivanenko, Poppele, & Lacquaniti, 2006).

Walking patterns were more irregular than jogging and running patterns, i.e. in our experiments the coefficient of variation of the step duration of walking was larger than that for running or jogging for all subjects. Thus, our improvement of classification performance by cycle alignment was better for activities containing a larger degree of variability, which was to be expected. Further, the test protocol described in Table 5.1 was not randomized for each user. However, with the randomization we would expect better results because the motion pattern recorded in each test would appear more irregular due to changes in gait speed between consecutive tests.

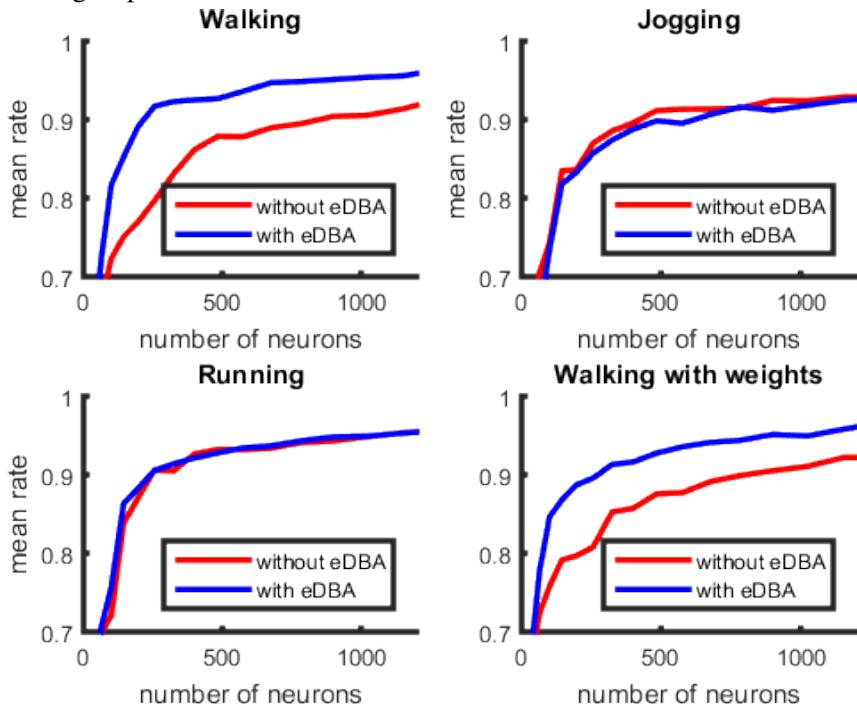
The mean success rate of subject classification and slope classification are shown in Fig. 5.15a and Fig. 5.15b, respectively. On the one hand, the results of the subject classification ranged between 0.9 and 1 when the number of neurons in the map exceeded 400. The results improved for walking data with weights when the phase registration was used. On the other hand, the results of slope classification ranged between 0.8 and 1 when the number of neurons in the map exceeded 400. The results improved in particular for walking data (i.e. with and without weights) when the phase registration was used.

Fig. 5.16a and Fig. 5.16b show the hit histogram for each neuron in a 20×20 SOM trained with acceleration data obtained from walking data with weights. Hit histograms indicate how the best matching units of some data sets are distributed on a map. In each figure, regions with the same labels are distinguishable. In Fig. 5.16a regions corresponding to subjects and in Fig. 5.16b small regions corresponding to the slopes performed by subjects are discernible in the map. The map clustered the slopes together for each subject. The latter indicates that the variability between subjects was larger than the variability of the tests of each subject.

Finally, we observed the effects of phase registration calculating the eDBA average from the cycles of all tests together, i.e. the 12 tests given in Table 5.1. The SOMs were trained for the same size as before with and without phase registration in the preprocessing step. The subject classification ranged from 0.9 to 0.95 did not improve, but the slope classification and activity

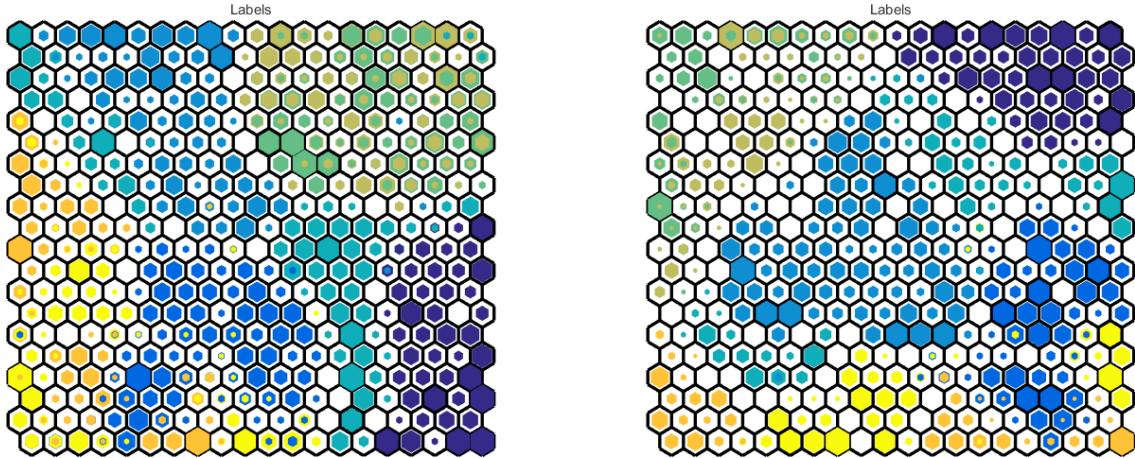


(a) Mean rate of subject classification success for each activity after training with and without phase registration in preprocessing step.

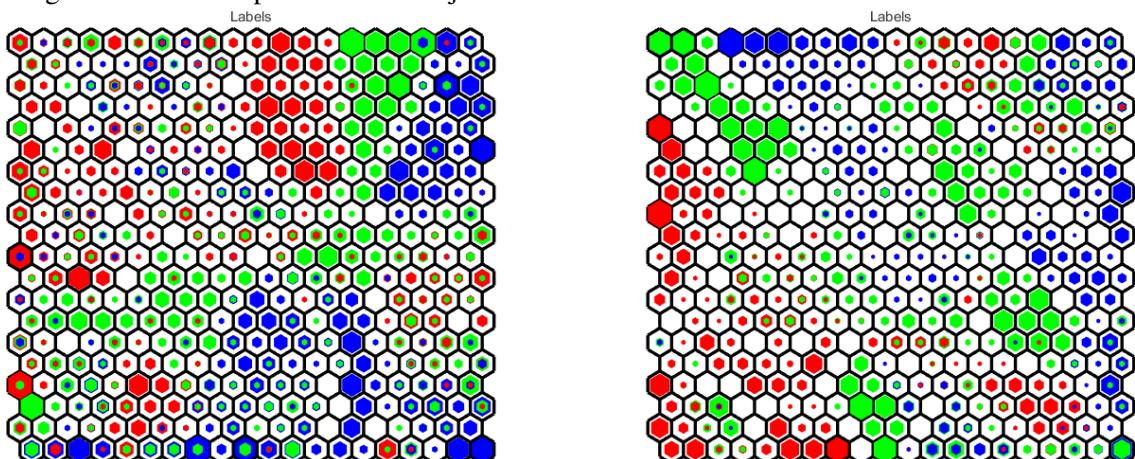


(b) Mean rate of slope classification success for each activity after training with and without phase registration in preprocessing step.

Figure 5.15: Mean rate of subject and slope classification success.



(a) Examples of subject classification in a 20×20 map of cycles obtained from the activity walking with weights after training with (right side) and without (left side) phase registration in preprocessing step. Each of the eight colours corresponds to one subject.



(b) Examples of slope classification in a 20×20 map of cycles obtained from the activity walking with weights after training with (right side) and without (left side) phase registration in preprocessing step. Each of the three colours corresponds to one slope.

Figure 5.16: Examples of subject and slope classification in a 20×20 map.

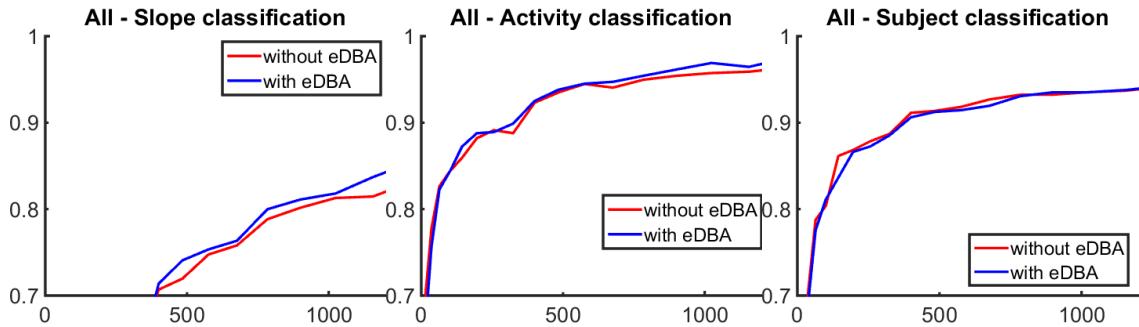


Figure 5.17: Results with all tests together. Left: Results of slope classification. Center: Results of activity classification. Right: Results of subject classification.

classification (i.e. walking, jogging, running, and walking with weights) improved slightly. The rate of slope classification ranged from 0.7 to 0.86 improved 0.005–0.01, and the rate of activity classification ranged from 0.9 to 0.95 improved 0.005–0.02. See Fig. 5.17 for the results.

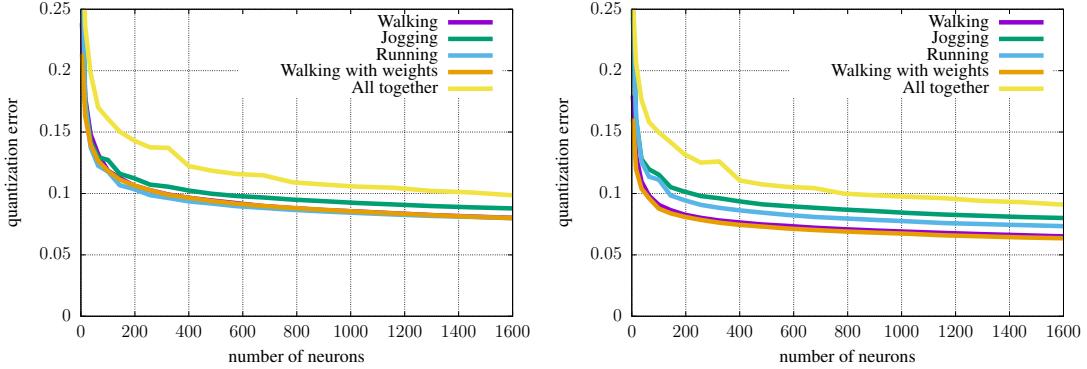
SOM quality

Fig. 5.18a shows the mean quantization error with and without registration based on eDBA averaging in the preprocessing step against the number of neurons in the map. Fig. 5.18b shows the combined error against the number of neurons in the map. The quantization and the combined errors were normalized by the norm of each input vector. For both curves, starting from the lowest number of neurons, when the number of neurons was incremented the error curves fell sharply until a bend was achieved. Then, the error curves decreased gently. Furthermore, the normalized error values with eDBA in the preprocessing step were smaller than without registration for walking with and without weights.

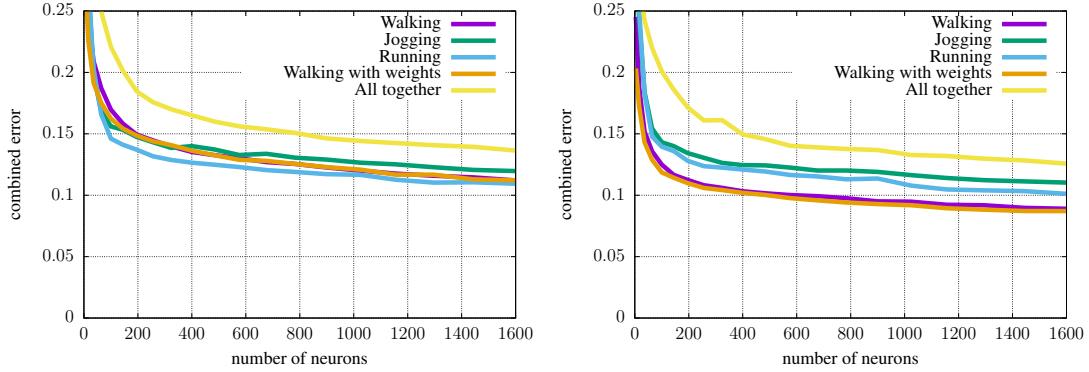
The error values and the position of the bend at the error curves are relevant to selecting the appropriate number of neurons of the map, i.e. avoiding overfitting while representing the input data distribution. For example, if the bend in the curve is defined as the bisection point for which the sum of the least squared errors of the fitting of two lines to all the points on the left and right sides is minimal, then the bend points in Fig. 5.18a and Fig. 5.18b are located at 36–196 neurons. The bend location did not reduce with eDBA averaging in the preprocessing step.

For data mining and visualization applications, it is important that two neighbouring neurons represent the same type of data and two distant neurons (according to the grid) represent different data. However, at the end of the learning process, some neighbouring neurons may not represent the same data. In Fig. 5.18c we plotted the percentage of data vectors for which the first and second BMUs are non-adjacent units against the number of neurons in the map. This percentage was smaller for maps trained with phase registration based on average equalized DBA (eDBA) in the preprocessing step except for walking. For walking and maps with neurons less than 256, the percentage was higher than without phase registration, although the quantization error and the combined error were the lowest among the activities.

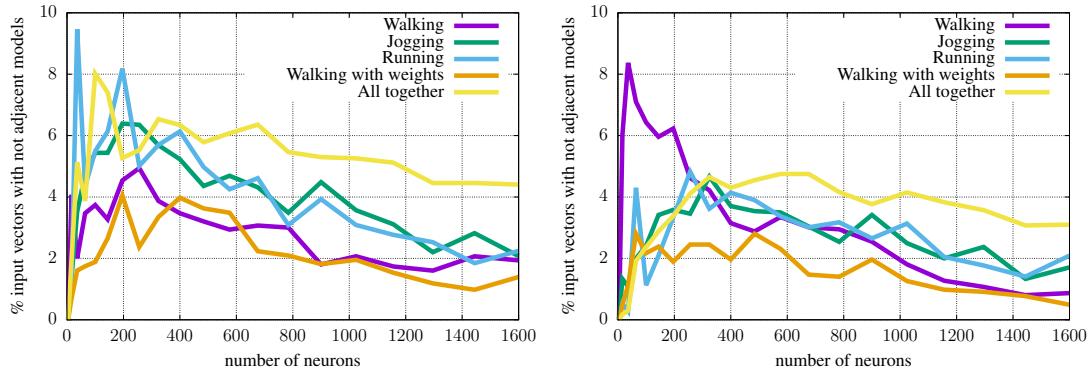
For all curves in Fig. 5.18c, when the number of neurons were increased the percentage grew until a peak was reached. Then, the percentage fell smoothly. The shape of the curves shows the trade-off between the size of a map and its fit to the given data. When the number of neurons was too small, topological errors were low but the quantization error was large as shown in Fig. 5.18a. When the number of neurons increased, the topological errors decreased. The curves do not indicate overfitting of the SOMs on the data because the maximal number of neurons (i.e. 1600 neurons)



(a) Mean quantization error without (left) and with (right) equalized DBA (eDBA) in preprocessing step.



(b) Mean combined error without (left) and with (right) equalized DBA (eDBA) in preprocessing step.



(c) Percentage of input vectors for which first- and second-BMUs are not adjacent units in the resulting map without (left) and with (right) equalized DBA (eDBA) in preprocessing step

Figure 5.18: SOM quality curves without (left) and with (right) equalized DBA (eDBA) in preprocessing step.

was much smaller than the number of training samples.

5.6 Conclusions

We proposed a new method to calculate the average of a set of cycles based on dynamic time warping (DTW) and a modification of DTW barycentric averaging (DBA). Our method allows the study of features of kinematic variables in cyclic motion depending on the phase using an average cycle as the reference. In addition, a new definition of the quality of intersections of cycles with transversal sections is given based on signal-to-noise ratio (SNR).

We tested self-organizing map (SOM) of different sizes with and without phase registration in the preprocessing step, and compared the resulting maps using three quality measures based on mapping precision, topology preservation and a combination of both. The phase registration as a part of the preprocessing step increased the mapping precision and the topology preservation quality. Furthermore, the phase registration improved the success rate of slope classification especially for walking gait data which were more irregular than the other activities (e.g. with 100 neurons or more the improvement was 5%–10%).

From our results, we suggest applying phase registration of the cycles with eDBA for each activity, in particular, when cycle classification is required (e.g. see the results with different slopes). For activity classification (i.e. walking, jogging, running) the registration using eDBA of all activities is recommended as well.

For future research, a framework to validate the clustering results with experts can be done. In addition, the effects on the SOMs with different parameters for calculating the eDBA average and the phase registration must be tested. Further, the cost function in the alignment method can be augmented by the motion velocity in state space and apply the method to time series obtained from a reconstruction of the embedding space using Takens' theorem. Furthermore, the extension of the method to continuous time must be explored (see Section 7.2.5 for a description).

While curve registration is useful for mitigating unwanted phase variation in cycle curves, there are instances where phase variability is itself of interest. Because curve registration changes the temporal location of data, it should not be applied in studies concerned with dynamic features explained in Chapter 3.

Chapter 6

Estimation of torque variation from pedal motion in cycling ergometry

Cycling is the result of the interaction between the cyclist, the bicycle, and the environmental constraints. The right selection of the variables involved in adjusting a bicycle (e.g. handle bar position, seat height, crank length, etc.), the correct body position on the bicycle, and a good pedalling technique are necessary to prevent cycling injuries and to optimize the force distribution during the pedal stroke.

Applying correctly oriented forces to the pedal (i.e. pedalling technique) is a major component of skilled performance on the bicycle. Although there is no agreement on the characteristics of the optimal pedalling technique, any desired pedalling technique can be learned and trained by cycling ergometry. A training session in the lab is based on the analysis and feedback of the distribution of the pedal forces applied during each revolution of the crank. In many cases, torque variation is enough for training pedalling techniques and the correction of force asymmetry between the legs. For some examples, see (Böhm et al., 2008; Faria, 2009).

The determination of the pedal forces is fundamental to analysing cycling performance from a biomechanical point of view. Pedal sensors implemented in the pedal have been validated in the literature for measuring force in one up to three dimensions, based on strain gauges or piezoelectric elements. For a list of sensors, see (Mornieux, Zameziati, Mutter, Bonnefoy, & Belli, 2006; Stapelfeldt, Mornieux, Oberheim, Belli, & Gollhofer, 2007). A direct measure of the torque (or, as well, the tangential force applied to the pedal that rotates the crank) requires expensive sensors, but indirect ways of calculating the torque variations are possible and presented in this contribution.

The study of pedal forces in cycling is of interest for several applications. For example, some related topics are: pedalling biomechanics (Kautz & Hull, 1993; Christensen et al., 2000; Hug, Turpin, Guével, & Dorel, 2010), limb coordination (Jirsa, Fink, Foo, & Kelso, 2000; Bini et al., 2010), human motion modelling (Hull, Kautz, & Beard, 1991; Höchtl, Böhm, & Senner, 2010), detection and correction of asymmetry (Sanderson, 1990; Smak, Neptune, & Hull, 1999; F. Carpes, Rossato, Faria, & Bolli Mota, 2007), evaluation of body performance given a pedalling technique (Cannon, Kolkhorst, & Cipriani, 2007; Ettema & Lorås, 2009; Theurel, Crepin, Foissac, & Temprado, 2011), cadence and workload effects on pedalling technique (Black, 1994; Stapelfeldt, Mornieux, & Gollhofer, 2006; Rossato, Bini, Carpes, Diefenthäler, & Moro, 2008) and the influence of pedalling technique on muscular efficiency (Theurel et al., 2011).

We propose to derive the variation in the net torque from measurements of pedal motion that can be made in the laboratory using motion-capturing or optical motion-tracking devices (e.g. by plain commercial cameras). For an ergometer with almost constant brake torque, we may assume

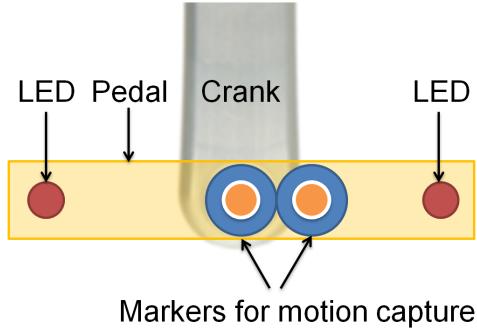


Figure 6.1: Pedal motion was recorded using video (through two LEDs) and motion capture (through two active infrared markers) in order to reconstruct the position and the orientation of the pedal. For motion capture, one marker was placed on the pivot point of the pedal rotation and the other marker was placed to the right side of the first marker. For video, the LEDs were placed on both sides of the pedal so that the midpoint between LEDs was aligned with the pivot point.

that variations in the net torque can be extracted from the pedal motion alone. The key problem is to reliably estimate the angular pedal acceleration from measurements of pedal motion.

The reconstruction of the position and the orientation of the pedal was done in two ways: (1) recording a video of two light-emitting diodes (LEDs), and (2) by motion capture using two active infrared markers. See the experimental setup in Fig. 6.1. In order to calculate the angular acceleration from the pedalling motion, the angular position of a point representing the angular position of the crank is differentiated twice. This is not a trivial task because the differentiation process amplifies noise by its very nature, (Ovaska & Valiviita, 1998). For this reason, a bandwidth differentiator with special characteristics is required for this calculation. We tested the Savitsky–Golay filter (Savitzky & Golay, 1964) from which we obtained the filtered components required for the calculation of the second derivative of the angular position of the crank. In addition, we observed the differences in our results when a novel method for the correction of the marker position was applied to the motion capturing data (MoCap) and video data. Without marker correction, the pedal torque results from MoCap were better than from video but the results from video were close to MoCap results when a correction of the marker position was applied. The physical relation between the angular acceleration of the crank and the net torque applied to the pedals is derived, allowing us to compute the net torque from our motion data. In our experiments, the pedal brake forces ranged between 100 and 250 N, and cadences of 60, 80, and 100 rpm were used. The comparison of the results with directly measured crank torque (using an SRM Torque Box¹) provides a validation of our method.

This chapter is based on our publication (Quintana-Duque, Dahmen, & Saupe, 2015) and divided into the following sections: Section 6.1 introduces the calculation of the second derivative from the positional data. Section 6.2 presents the novel method for the correction of the marker position. Section 6.3 reviews the concepts and the features of force and torque in pedalling. Section 6.4 describes the torque validation. Section 6.5 describes the experimental setup. Section 6.6 presents the results. Finally, Section 6.7 describes the conclusions.

¹<http://www.srm.de/products/torque-analysis-system/>

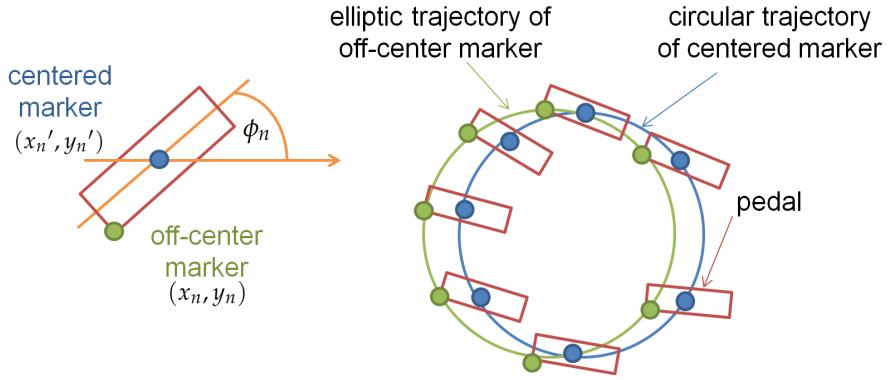


Figure 6.2: An off-centre marker on the pedal gives a trajectory that is not a circle when the pedal orientation is not constant.

6.1 Calculation of the second derivative from positional data

We assumed that the angular position of the pivot point of rotation of the pedal and the crank angle θ with respect to the centre of rotation of the crank are the same. If $(x(t), y(t))$ denote the marker coordinates of a point representing the angular position of the crank, with the origin of the coordinate system placed at the centre of the crank rotation, then the crank angle is given by $\theta(t) = \arctan(y(t)/x(t))$. The second derivative of θ (i.e. angular acceleration) is calculated using the chain rule for derivatives

$$\ddot{\theta} = \frac{d^2}{dt^2} \arctan \frac{y}{x} = \frac{(\ddot{y}x - y\ddot{x})(x^2 + y^2)}{(x^2 + y^2)^2} - \frac{(\dot{y}x - \dot{x}y)(2x\dot{x} + 2y\dot{y})}{(x^2 + y^2)^2}. \quad (6.1)$$

In our application, the Savitsky–Golay filter is applied separately to the x - and y -coordinates of the pedal motion to obtain smoothed data and its first and second derivatives. This filter is briefly reviewed in Section 8.3. Alternative ways of calculating the second derivative of θ , which depend directly on the angle, are:

- Finite central difference approximation to the second derivative

$$\frac{d^2\theta(t)}{dt^2} \approx \frac{\theta(t + \Delta t) - 2\theta(t) + \theta(t - \Delta t)}{\Delta t^2}. \quad (6.2)$$

- Second derivative of a fitted polynomial

$$\frac{d^2\theta(t)}{dt^2} \approx p''(t). \quad (6.3)$$

6.2 Correction of the marker position

The crank rotation during pedaling motion ideally takes place in a two-dimensional plane. We expect that any marker on the crank, except the pivot of crank rotation, describes a perfect (sampled) circular trajectory. The angular position of the crank can be obtained from the trajectory of a point representing the angular position of the crank. In the same way, the angular position can be calculated from the trajectory described by a marker on the pedal, provided that the marker is on

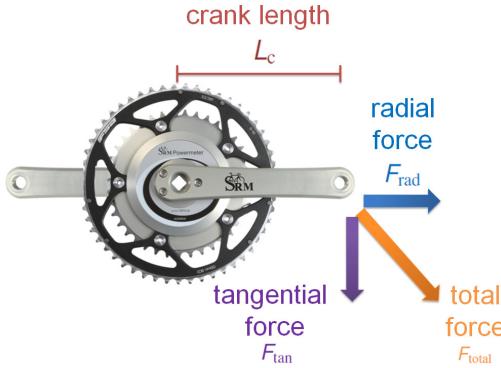


Figure 6.3: The total force applied to the pedal, F_{total} , is the sum of two perpendicular forces: the tangential force F_{\tan} and the radial force F_{rad} .

the pivot of pedal rotation. Otherwise, the marker trajectory turns into an approximate ellipse due to simultaneous rotations of the pedal and the crank. See Fig. 6.2.

However, although the marker position is on the pivot of the pedal rotation, crank and pedal rotations in the real world describe elliptical trajectories because the bearings at the pedals and in the bottom bracket may have a small amount of play and the crank sets are slightly elastic.

The eccentricity of an ellipse fitted to the marker trajectory indicates by how much the marker trajectory deviates from a circle. The eccentricity e is defined as $e = \sqrt{1 - b^2/a^2}$ with a and b denoting the lengths of the major and minor axes, respectively. The correction $(\Delta x, \Delta y)$ of the 2D position of a marker on the pedal can be calculated if the pedal orientation ϕ_n with respect to the horizontal axis is known. The correction of the marker position is given by

$$\begin{pmatrix} x_n' \\ y_n' \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \begin{pmatrix} \cos \phi_n & -\sin \phi_n \\ \sin \phi_n & \cos \phi_n \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix},$$

where x_n and y_n are the old coordinates, and x_n' and y_n' are the new coordinates after correction. The correction vector $(\Delta x, \Delta y)$ is defined such that: either 1) a prescribed target eccentricity of an ellipse fitted to the corrected positional data is achieved (in the following, called the “prescribed eccentricity”) or 2) the root mean square (RMS) distance between the fitted ellipse and the corrected data is minimal among all possible corrections and their respective fitted ellipses (in the following, called the “minimum error eccentricity (M.E.”)).

6.3 Related concepts: Force and torque

The total force F_{total} applied to the pedal is the sum of all vector forces, see Fig. 6.3, produced by the contractions and extensions of the leg and hip muscles. These can be decomposed into the tangential force, F_{\tan} , and the radial force, F_{rad} (Eq. 6.4). The force F_{\tan} is tangential to the crank rotation and the radial force F_{rad} is parallel to the crank. Only F_{\tan} contributes to the rotation of the crank.

$$F_{\text{total}} = F_{\tan} + F_{\text{rad}}. \quad (6.4)$$

The torque describes the effect of a force on the rotational motion of the pedal pivot point about the axis on the bearing. Mathematically, the torque is the cross product of the lever-arm length vector, L_c , and the force F_{\tan} acting on the end of the lever-arm. Here, the lever-arm length is equal

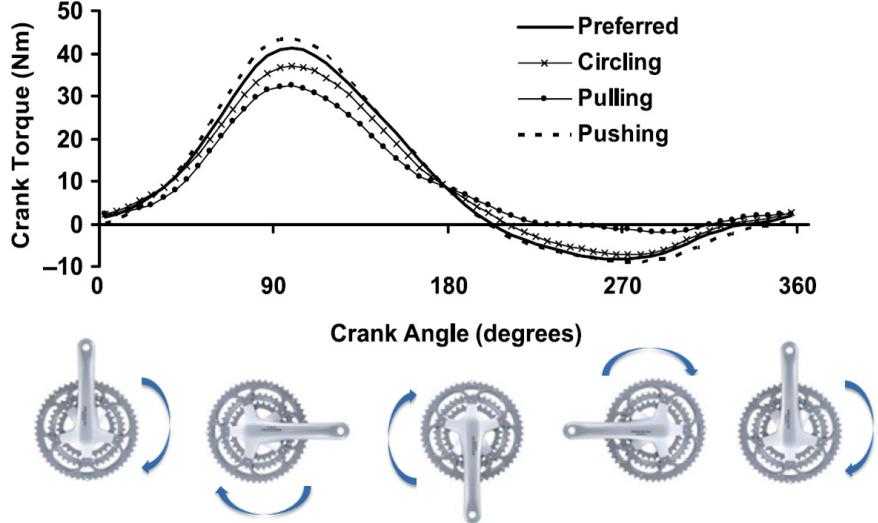


Figure 6.4: Right leg torque profiles using different pedalling styles at 90 rpm and 200 W. Figure adapted from Korff et al., 2007. Four pedalling styles are shown: individually preferred pedalling technique (preferred), pedalling emphasizing the transition phases through top dead centre at 0° and bottom dead centre of the crank cycle at 180° (circling), emphasizing an active pull during the upstroke of the crank cycle (pulling), and emphasizing the pushing action during the downstroke of the crank cycle (pushing).

to the crank length L_c . Thus, the magnitude of the torque is given by

$$\|\tau\| = \|\mathbf{L}_c \times \mathbf{F}_{\tan}\| = L_c F_{\tan}. \quad (6.5)$$

Fig. 6.4 shows the torque patterns of one pedal for different pedalling styles, where a positive peak around 90° (downstroke) and a negative peak around 270° (upstroke) can be seen for all styles. The measurements were made by Korff et al. (2007) using a custom made force pedal with two triaxial piezoelectric force sensors, where 0° is the highest position of the crank rotation.

During the pedalling motion, two main torques act on the crank at the same time: the brake torque τ_{brake} and the applied net torque τ_{net} . The brake pedal torque is the sum of all torques produced by the different forces against the cycling motion, e.g. the rolling and aerial resistive forces of a bicycle on the road, or the forces produced by an eddy-current brake in an ergometer in the lab. The net torque τ_{net} is the sum of the individual torques τ_{left} and τ_{right} applied to the left and right pedals:

$$\tau_{\text{net}} = \tau_{\text{left}} + \tau_{\text{right}}. \quad (6.6)$$

Fig. 6.5 shows an example of the net torque curve with some typical features to describe it. For example, different peak values indicate an asymmetry due to emphasizing an unequal pushing action during the downstroke or due to a problem with pulling during the upstroke.

The model of the torques acting during pedalling is shown in Fig. 6.6. In this model, the inertial mass of the cyclist and the bicycle is (partly) realised by the flywheel in the rear part of an ergometer. Assuming that the rear and frontal sprockets are connected with a chain which is not elastic and does not slip, and that the frictional forces between the chain and the sprockets are small, the physical relation between the torques τ_{net} and τ_{brake} , and the crank angular acceleration α_{crank} , is given by Eq. 6.7. In this equation, I_{flywheel} and I_{crank} represent the moments of inertia of the flywheel and

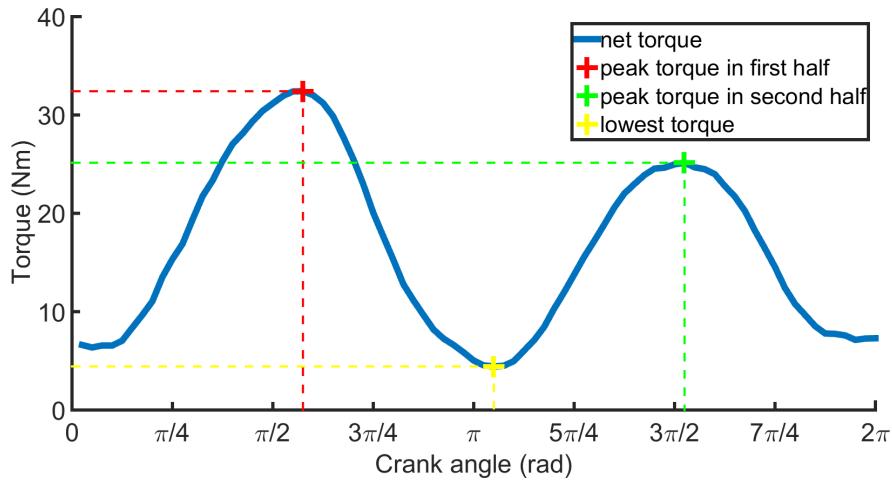


Figure 6.5: Mean of net torque τ_{net} for one pedal revolution with 60 rpm and 100 N ergometer brake force (17.6 Nm).

the crank, $R_{\text{fs}}/R_{\text{rs}}$ is the ratio of the radii of the chain wheel and the rear sprocket, i.e. the gear ratio, and L_c is the length of the crank.

$$\tau_{\text{net}} = \left(I_{\text{crank}} + \left(\frac{R_{\text{fs}}}{R_{\text{rs}}} \right)^2 I_{\text{flywheel}} \right) \cdot \alpha_{\text{crank}} + \tau_{\text{brake}} . \quad (6.7)$$

$$\tau_{\text{net}} = K \cdot \alpha_{\text{crank}} + \tau_{\text{brake}} , \quad (6.8)$$

with

$$K = \left(I_{\text{crank}} + \left(\frac{R_{\text{fs}}}{R_{\text{rs}}} \right)^2 I_{\text{flywheel}} \right) . \quad (6.9)$$

For the complete deduction of the torque equation for the pedal, see Section 6.3.1.

6.3.1 Deduction of torque equation for the pedal

This section shows how to deduce the formula that relates the net torque τ_{net} with the crank acceleration α_{front} . To derive this equation, we used a crank length of L_c , moments of inertia I_{flywheel} and I_{crank} , and a gear ratio of $\frac{R_{\text{fs}}}{R_{\text{rs}}}$. We will formulate the equations for the rear (i.e. flywheel side) and front (i.e. crank side) of the bicycle. For the front, the net torque is the sum of the torques produced by the right and left legs when force is applied to the pedals:

$$\tau_{\text{net}} = \tau_{\text{left}} + \tau_{\text{right}} . \quad (6.10)$$

Due to the torque equilibrium, the sum of the torques $\tau_{\text{net}} - \tau_{\text{brake}}$ and the torque on the border of the frontal sprocket are equivalent. The frontal sprocket has a radius of R_{fs} . Thus, the force F_{front} acting on the border of the sprocket can be written as

$$F_{\text{front}} = \frac{\tau_{\text{net}} - \tau_{\text{brake}}}{R_{\text{fs}}} . \quad (6.11)$$

The angular acceleration of the crank α_{front} , the force on the chain F_{chain} , the applied net torque τ_{net} , and the brake torque τ_{brake} are related by

$$I_{\text{crank}} \cdot \alpha_{\text{front}} = \tau_{\text{net}} - F_{\text{chain}} \cdot R_{\text{fs}} - \tau_{\text{brake}} . \quad (6.12)$$

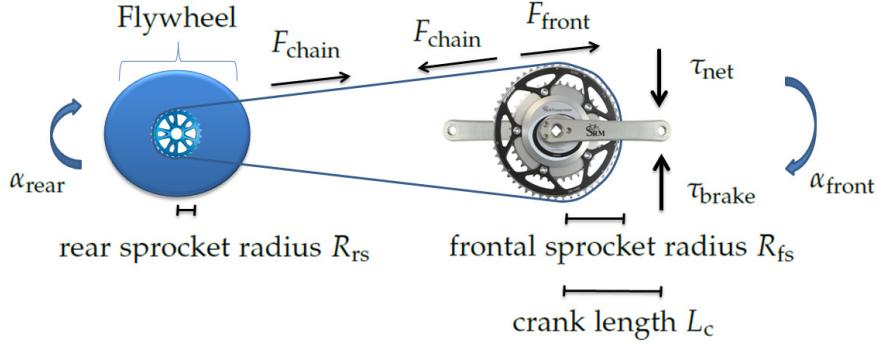


Figure 6.6: Torque and forces present during pedalling.

Regarding the rear, the rear sprocket has a radius of R_{fs} . The force on the chain F_{chain} and the acceleration of the flywheel α_{rear} are related by

$$I_{flywheel} \cdot \alpha_{rear} = F_{chain} \cdot R_{rs} \quad (6.13)$$

Thus, the force on the chain F_{chain} is

$$F_{chain} = \frac{I_{flywheel} \cdot \alpha_{rear}}{R_{rs}}. \quad (6.14)$$

Isolating F_{chain} in Eq. 6.12 yields

$$F_{chain} = -\frac{I_{crank} \cdot \alpha_{front} - \tau_{net} + \tau_{brake}}{R_{fs}}, \quad (6.15)$$

and combining Equations (6.14) and (6.15), we get

$$\frac{I_{crank} \cdot \alpha_{front} - \tau_{net} + \tau_{brake}}{R_{fs}} = -\frac{I_{flywheel} \cdot \alpha_{rear}}{R_{rs}}. \quad (6.16)$$

We know that the acceleration of the flywheel (α_{rear}) and the crank (α_{front}) are related by

$$\alpha_{rear} = \left(\frac{R_{fs}}{R_{rs}} \right) \cdot \alpha_{front}. \quad (6.17)$$

Combining the last two equations and isolating τ_{net} , we obtain

$$\tau_{net} = I_{crank} \cdot \alpha_{front} + \left(\frac{R_{fs}}{R_{rs}} \right)^2 \cdot I_{flywheel} \cdot \alpha_{front} + \tau_{brake} \quad (6.18)$$

$$\tau_{net} = \left(I_{crank} + \left(\frac{R_{fs}}{R_{rs}} \right)^2 \cdot I_{flywheel} \right) \cdot \alpha_{front} + \tau_{brake}. \quad (6.19)$$

Furthermore, we know that

$$\alpha_{front} = \alpha_{crank}, \quad (6.20)$$

and so, combining the last two equations and isolating τ_{net} , we obtain

$$\tau_{net} = \left(I_{crank} + \left(\frac{R_{fs}}{R_{rs}} \right)^2 I_{flywheel} \right) \cdot \alpha_{crank} + \tau_{brake}. \quad (6.21)$$

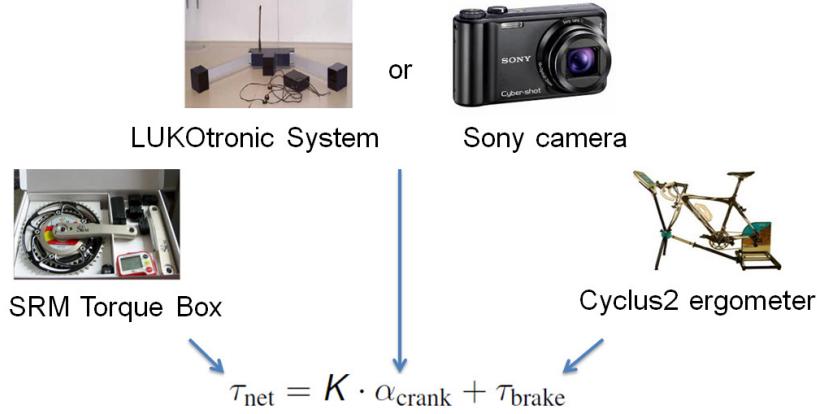


Figure 6.7: Equipment used for the experiment and their relation with the equation to validate. The angular crank acceleration was calculated from the data captured by the LUKOtronic System and the Sony camera.

Denoting by K a constant equal to

$$K = \left(I_{\text{crank}} + \left(\frac{R_{\text{fs}}}{R_{\text{rs}}} \right)^2 I_{\text{flywheel}} \right), \quad (6.22)$$

we finally obtain

$$\tau_{\text{net}} = K \cdot \alpha_{\text{crank}} + \tau_{\text{brake}}. \quad (6.23)$$

6.4 Torque validation

Our goal was to determine to what extent the net torques computed using Eq. 6.7, based on numerically approximated crank acceleration from pedal motion measurements, are close to the directly measured (assumed to be true) net torques. For this purpose, we recorded the data simultaneously obtained from the motion capture system (i.e. MoCap data), the video camera recorder (i.e. video data), the bicycle ergometer, and the torque sensor, see Fig. 6.7.

The validation is based on the fact that the angular crank acceleration α_{crank} is equal to the second derivative of the angular position of the pivot of rotation of the pedal θ . To calculate the right-hand side of Eq. 6.7, we used a crank length $L_c = 176$ mm, and moments of inertia $I_{\text{flywheel}} = 0.6576 \text{ kg} \cdot \text{m}^2$ and $I_{\text{crank}} = 0.02 \text{ kg} \cdot \text{m}^2$. The gear ratio $\frac{R_{\text{fs}}}{R_{\text{rs}}}$ was calculated directly from the number of teeth on the gears in the gear train with $R_{\text{fs}} = 50$ and $R_{\text{rs}} = 13$. The inertia of the crank is composed of the inertia of the pedals, of the crank arm, of the chain rings, and of the SRM Power Meter. Each component can be approximated by a primitive geometric form rotating around the centre of the crank axis (pedal – point mass, crank arm – solid cylinder, chain rings – rings, SRM – solid disc). We measured the weights and the sizes of each component and computed the individual moments of inertia. Their sum yields the total inertia of the crank $I_{\text{crank}} = 0.02 \text{ kg} \cdot \text{m}^2$ (for more details see Dahmen & Saupe, 2011). The moment of inertia of the flywheel was found empirically under the assumption that the friction force is an affine function of the velocity, which was verified through a fitting procedure. For this, we pedalled to accelerate the flywheel until it had a high speed. Then, we stopped pedalling and we counted the time until the flywheel did not move anymore.

The pipeline for calculating the angular acceleration from the positional data of the pedal is shown in Fig. 6.8. The input is either the 3D MoCap data projected onto a two-dimensional plane

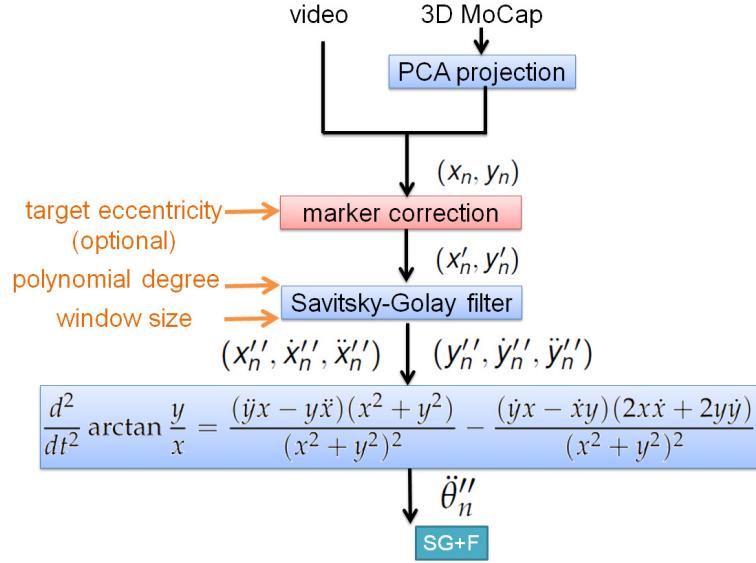


Figure 6.8: Pipeline for calculating angular crank acceleration from positional data. Parameters are given in orange. We assumed that the angular acceleration of the crank α_{crank} is equal to $\ddot{\theta}_n$.

using principal component analysis (PCA), described in detail in (Smith, 2002), or the 2D positional video data. The correction of the marker position, explained in Section 6.2, is a part of the preprocessing step (the red module in Fig. 6.8). We compared the results without this correction and with it, based on two different methods, i.e. minimizing the RMS distance between a fitted ellipse and the corrected data among all possible corrections and fitted ellipses, and with prescribed eccentricities. We tested the approach SG+F shown in Fig. 6.8 consisted of the filtering and the calculation the zeroth, first and second derivatives of x and y coordinate sequences using the corresponding Savitzky–Golay filters for each case, and with this data then calculating the second derivative of θ using Eq. 6.1.

Finally, the effects of the correction of the marker position and filtering with different parameters were compared using the SNR, where the true signal was assumed to be the SRM Torque Box data, and the noise was assumed to be the difference between the SRM Torque Box data and the right side of Eq. 6.7.

6.5 Experimental setup

6.5.1 Data acquisition and processing

We recorded the pedalling motion using simultaneously a commercial camera (Sony Optical Steady shot DSC–H55) and a motion capture system (LUKOtronic-Steinbichler Optotechnik GmbH, Neubeuern, Germany). The camera recorded videos with a sampling frequency of 29.97 Hz and a resolution of 1280×720 pixels. The video camera was placed at a distance of 270 cm from the bicycle.

The motion capture system consisted of a beam with three integrated infrared cameras. As a result of the fixed positions of the cameras within the beam, the system was pre-calibrated and did not require any separate calibration procedure. The motion capture system provided three-dimensional positional data of the infrared active markers. The sampling rate was 240 Hz. The positional accuracy provided by LUKOtronic for the distance used in our experiments was 1–3 mm. Experimentally, the positional accuracy was 1.5 mm with a precision of 0.9 mm calculated from

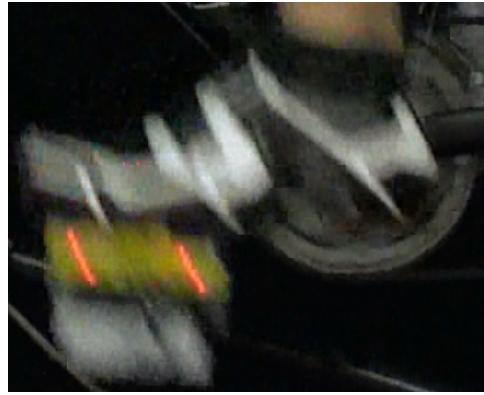


Figure 6.9: Example of LED trajectories as red streaks on yellow pedal in one video frame.

two markers fixed on the pedal during 180 crank rotations.

Using the motion capture system we recorded the position of two markers as shown in Fig. 6.1. One marker was placed on the pivot point of the pedal rotation (i.e. a point representing the angular position of the crank during the motion of the pedal) and the other marker was placed on the right side next to the first marker. During the recording, we used three reference markers attached to the bicycle frame in order to define the coordinate system, which was re-calculated for each measurement. This has the advantage that oscillations and deflections of the bicycle frame due to the pedalling motion and bike sway are taken into account during the recording of the pedal motion. The pedal orientation for the motion data was calculated from the positional data of both markers with respect to the horizontal axis. We projected the three-dimensional coordinates of the marker position to a two-dimensional plane using the PCA. Then, we rotated the coordinate system using the information of additional measurements of the pedal in the lowest position.

Using the video camera we recorded two red LEDs placed on both sides of the pedal. See Fig. 6.1 for an example. Each LED in a frame was captured as a streak due to blurring by the rapid movement of the pedal during the exposure time of the camera. See Fig. 6.9 for an example. The LEDs were placed with enough space between them to avoid a possible overlapping of the streaks. The coordinates of the heads of each streak were used in each frame as positional data of LEDs. The pivot point of the pedal was calculated as the average of both LED positions and the pedal orientation was calculated from the angle of the line between both LEDs with respect to the horizontal axis of the frame. For the detection of the streaks, each frame was converted from RGB format (i.e. red, green, and blue channels) to HSV format (i.e. hue H, saturation S, and value V). Then, a threshold on the value V was applied to find the blob s corresponding to the LED trajectories.

The LEDs are followed in our videos instead of the painted pedal for the calculation of the pedal position because the coordinates obtained from the painted pedal were not precise enough for the calculation of the angular acceleration. One main problem was the change of colour intensity of the painted pedal with pedal position. With the lamps on the roof illuminating the lab, at the lowest position of the crank rotation the colour of the pedal was darker than at the highest position. A direct illumination with an extra lamp did not solve the problem because in some video frames the pedal was as bright as the crank set. Another problem was the blurry images of the painted pedal at the places where the pedal rotated quickly. See a comparison of the data obtained from LEDs and painted pedal in (Y. Zhang, 2015).

To control the pedal brake force, we used the Cyclus2 ergometer (RBM Elektronik-Automation

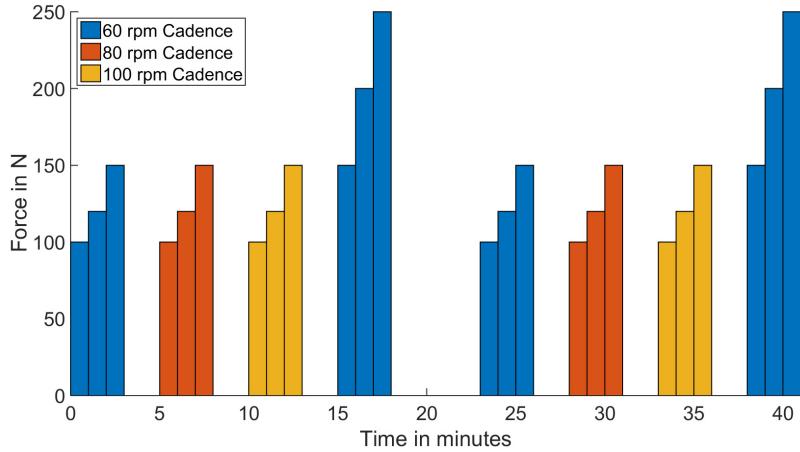


Figure 6.10: Experimental protocol. Eight test of three minutes each are performed with breaks of 2 resp. 5 minutes in between.

GmbH, Leipzig, Germany). During the ergometer operation, the flywheel on the rear part is supplied with kinetic energy which maintains its angular momentum. The angular speed of the flywheel is decelerated by means of an eddy current brake. The eddy current brake guarantees a non-slipping transmission of the braking resistance. Operating the Cyclus2 in pedal force mode, a constant pedal brake force ($\pm 5\%$ error) is imposed.

Our SRM Torque Box (Schoberer Rad Messtechnik, Welldorf, Germany) gave an instantaneous torque signal with a sampling rate of 200 Hz. This torque corresponds to the net torque, as explained in Eq. 6.6. Furthermore, when the crank has completed one pedal revolution (i.e. when the crank has crossed the sensor of the SRM Torque Box attached to the bicycle frame), this event was reported. SRM claims an accuracy of 2% for power and torque measurements.

The laboratory conditions were optimized for recording the pedal motion: we darkened the room (i.e. we covered the window to avoid sunlight), used low infrared emitting light bulbs, and a special carpet to avoid infrared reflections from the ground.

Before the validation of the physical relation between the angular acceleration of the crank and the tangential force applied to the pedals, using Eq. 6.7, we preprocessed the data obtained from all devices (see Fig. 6.7). First, we applied linear interpolation when loss of data or outliers were found. Then, we resampled all data to 200 Hz.

The correction of the marker positions was performed using a quasi-Newton strategy, minimizing either the difference between the prescribed eccentricity and the eccentricity of an ellipse fitted on the corrected positional data or the RMS distance between a fitted ellipse and the corrected data among all possible corrections and fitted ellipses. The ellipse fitting was performed by minimizing the squared sum of orthogonal distances from the points to the fitted ellipse described by Gander, Golub, and Strelbel (1994).

6.5.2 Test design

Ten cyclists participated in this study (male, 31.45 ± 9.9 years). Each participant was asked to ride at a fixed cadence using his preferred technique and cycling shoes with cleats locked onto the pedal interface. A continuous feedback of the cadence was given on a projected video image, positioned in front of the cyclist. Before the measurements, cyclists performed a warm up session of 5 minutes

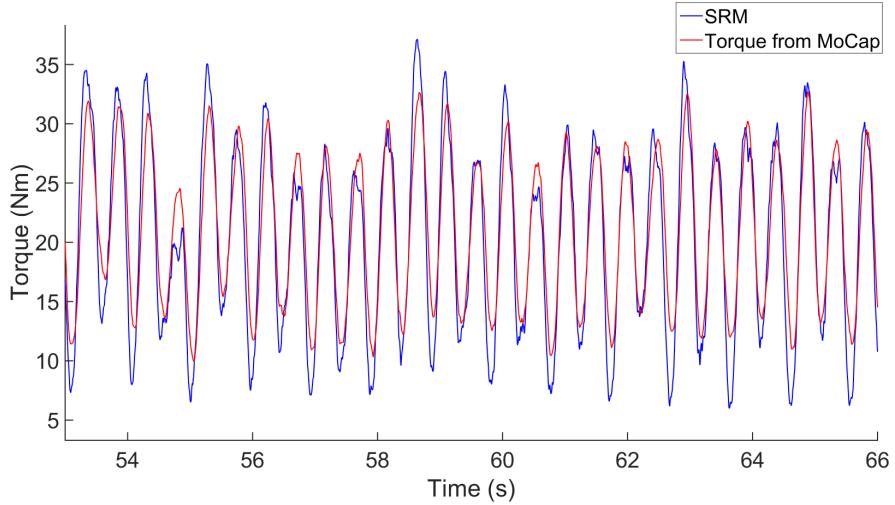


Figure 6.11: Result after alignment of SRM Torque Box data and MoCap data for the validation of the torque equation for the test with 60 rpm and 120 N (21.1 Nm) using the SG+F approach with window size 121, polynomial degree 2, and marker correction based on M.E.

at a power output of 100–140 W.

Each subject rode the bike simulator performing 8 tests of three minutes each. In each test, the cadence was held constant and the brake force increased every minute. For the first three tests the cadences were 60 rpm, 80 rpm, and 100 rpm, and for each test the brake forces were 100 N, 120 N, and 150 N (i.e. brake torques of 17.6 Nm, 21.1 Nm and 26.4 Nm). For the fourth test, the cadence was again 60 rpm but with forces of 150 N, 200 N, and 250 N (i.e. brake torques of 26.4 Nm, 35.2 Nm and 44.0 Nm). This last test was done to check the effects of large forces on the SNR results, which can vary due to crank deformation and brake force fluctuation. All four tests were repeated once after a pause of 5 minutes. See Fig. 6.10 for an illustration.

6.6 Results and discussion

Our goal was to determine to what extent the net torques computed using Eq. 6.7, based on numerically approximated crank accelerations from pedal motion measurements, are close to the directly measured (assumed to be true) net torques. We employed the following parameters for the Savitzky–Golay filter: polynomial degrees (2, 3) and window sizes ($10n + 1$, $n = 3, \dots, 22$). We applied the Savitzky–Golay filter to the resampled data at 200 Hz. We also considered the results with and without the correction of marker positions, i.e. minimizing the error between the fitted ellipse and corrected positional data (i.e. the minimum error eccentricity method) and with prescribed eccentricities ($e = 0.05i$, $i = 0, \dots, 6$). See subsection 6.2 for a description of the methods for the correction of the marker positions. Thus, we applied altogether for each test 360 different combinations of parameters and compared their SNR results. For an example of the resulting torques, see Fig. 6.11.

Table 6.1 shows the average of the best SNR results for MoCap and video of all tests with and without marker correction among all parameter combinations. The SNR results with a marker correction and eccentricity 0 (i.e. a perfect circle) were the lowest for both MoCap and video data (see Table 6.1, column 4). This confirms that the crank and the bearings are elastic to a small degree and allowed some play.

Table 6.2 shows the average results for each test without marker correction and with marker correction based on M.E. for each combination of force and cadence. The marker correction based on M.E. improved the SNR results of the video data. These results were expected because in our video data recordings the position of the midpoint between LEDs cannot be precisely at the pivot point. The correction of the marker position based on M.E. did not improve the SNR results of the MoCap data. This indicates that the active infrared marker of the motion capture device was well placed on the pedal pivot point.

The SNR results using the marker correction based on M.E. were lower than those obtained when fitting to an ellipse with prescribed eccentricities. For example, the marker correction based on a prescribed eccentricity of $e = 0.25$ improved the SNR results for video (+3.2 dB) and slightly for MoCap (+0.3 dB), see Table 6.1. Fig. 6.12 shows the effects on MoCap results with different combinations of parameters of Savitsky–Golay filters and prescribed eccentricities for the tests with forces 100 N–150 N. The best results for both MoCap and video data were obtained with prescribed eccentricity $e = 0.25$, window size 121, and polynomial degree 2. For forces larger than 150 N another set of parameters provided the best results. These parameters are given in Table 6.3.

To obtain the deformation of the trajectory of the point representing the angular position of the crank for the bicycle is difficult, but possible, e.g. with torsional strain gauges at the crank. However, the strain gauges are the main component of several commercial torque meters that would allow direct torque measurements. In any case, the marker correction based on M.E. can improve the SNR results for the proposed video-based indirect torque measurements.

The differences between the torque based on the angular acceleration (obtained from the MoCap and video data) and the measured torque with SRM Torque Box might have arisen from the eddy current brake of the bicycle simulator. This brake is less than ideal, so that fluctuations in the pedal brake force could occur during our test. Furthermore, we assumed that the chain, the crank and the pedal are completely inelastic and non-slipping, and that the frictional forces between the chain and the gears are negligible compared with the pedal brake force produced by the eddy current brake of the Cyclus2 ergometer.

Table 6.1: Average SNR for MoCap and video. Columns correspond to the type of input data, the SNR results without marker correction, the SNR results with M.E., and the results with different prescribed eccentricities for marker correction.

| Type | SNR | SNR $e = \text{M.E.}$ | SNR $e = 0$ | SNR $e = 0.15$ | SNR $e = 0.2$ | SNR $e = 0.25$ | SNR $e = 0.3$ |
|-------|------------|--------------------------|----------------|-------------------|------------------|-------------------|------------------|
| MoCap | 14.82 | 14.67 | 9.74 | 13.28 | 14.74 | 15.12 | 14.80 |
| | ± 1.86 | ± 1.75 | ± 1.97 | ± 2.04 | ± 1.94 | ± 1.96 | ± 2.11 |
| video | 10.81 | 12.30 | 9.61 | 11.60 | 13.42 | 14.02 | 14.18 |
| | ± 2.11 | ± 2.47 | ± 2.18 | ± 1.84 | ± 1.75 | ± 1.78 | ± 2.11 |

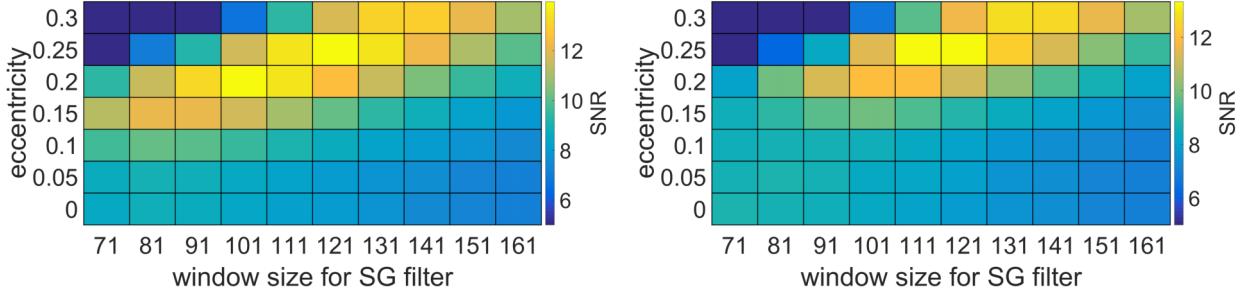


Figure 6.12: Mean SNR result of MoCap data (left) and video data (right) with correction of the marker position and different window sizes for the SG filter with degree 2 polynomials in the SG+FF approach.

Table 6.2: Average SNR of each test. The table shows the results for each combination of force and cadence. Columns correspond to brake force F_{brake} , cadence, the SNR results without marker correction and the results with marker correction based on M.E.

| Test | | MoCap | | video | |
|--------------------|---------------|---------------------|--------------------------|---------------------|--------------------------|
| F_{brake} | cad. N rpm | SNR | SNR $e = \text{M.E.}$ | SNR | SNR $e = \text{M.E.}$ |
| 100 | 60 | 15.22 ± 1.24 | 15.27 ± 1.27 | 11.22 ± 1.90 | 13.13 ± 2.13 |
| | 80 | 13.55 ± 1.29 | 13.34 ± 0.89 | 10.24 ± 1.41 | 11.03 ± 1.65 |
| 100 | 100 | 12.39 ± 0.98 | 12.30 ± 1.00 | 9.74 ± 1.75 | 10.90 ± 2.04 |
| | 60 | 16.28 ± 1.46 | 16.22 ± 1.40 | 12.18 ± 2.57 | 14.01 ± 2.89 |
| 120 | 80 | 14.71 ± 1.18 | 14.49 ± 0.80 | 12.03 ± 2.43 | 12.27 ± 2.22 |
| | 100 | 13.88 ± 1.44 | 13.95 ± 1.37 | 10.31 ± 1.89 | 12.08 ± 2.68 |
| 150 | 60 | 17.16 ± 1.31 | 16.88 ± 1.21 | 10.94 ± 1.99 | 13.63 ± 2.86 |
| | 80 | 15.20 ± 1.24 | 14.85 ± 1.30 | 11.34 ± 2.10 | 12.42 ± 1.86 |
| 150 | 100 | 14.96 ± 1.70 | 14.69 ± 1.41 | 9.29 ± 1.54 | 11.20 ± 2.33 |
| | 60 | 17.10 ± 2.06 | 17.19 ± 1.78 | 11.85 ± 2.33 | 15.18 ± 3.51 |
| 250 | 60 | 17.55 ± 2.77 | 17.30 ± 2.26 | 10.76 ± 1.63 | 14.79 ± 2.67 |

Table 6.3: Parameters for the best average SNR of the SG+F approach with marker correction based in two different methods. Columns correspond to type of data, brake force F_{brake} , cadence, polynomial degree and window size of the Savitsky–Golay filter employed for the method with prescribed eccentricities e , polynomial degree and window size of the Savitsky–Golay filter employed for the M.E. method.

| Type | F_{brake} N | cadence rpm | prescribed eccentricity | | | M.E. | | |
|-------|-------------------------|----------------|-------------------------|----------------|--------------|----------------------|----------------|--|
| | | | polynomial degree | window size | eccentricity | polynomial degree | window size | |
| MoCap | 100–150 | 60–100 | 2 | 121 | 0.25 | 2 | 91 | |
| MoCap | 200 | 60 | 2 | 121 | 0.3 | 2 | 81 | |
| MoCap | 250 | 60 | 3 | 191 | 0.2 | 3 | 191 | |
| video | 100–150 | 60–100 | 2 | 121 | 0.25 | 2 | 81 | |
| video | 200 | 60 | 3 | 171 | 0.25 | 3 | 201 | |
| video | 250 | 60 | 3 | 181 | 0.2 | 3 | 201 | |

For the calculation of the crank acceleration, it was assumed that the motion of the marker on the pedal pivot point lies in a two-dimensional plane. This plane can be easily calculated from MoCap data projecting the 3D data onto 2D by using PCA but with video data, an additional reference to align the camera plane with the pedal motion plane is necessary to have better results. The positional data extracted from each video frame corresponded to the 2D projection of the crank motion plane to the camera plane. If these planes are not aligned, the projection of the circular trajectory of the crank motion to the camera plane results in an oval trajectory. Such misalignment of both planes may happen during the pedalling motion due to bike sway. This undesirable distortion could be minimized in each frame by a perspective correction based on some reference marker points on the bike frame. See Section 7.2.7 for an idea of perspective correction based on a disk reference.

6.7 Conclusions

The physical relation between the angular acceleration of the crank and the tangential force applied to the pedals was validated with measured data of the net torque, brake force, video and motion capturing data (MoCap) data. These variables are related by an affine linear equation. The variation of the angular acceleration is proportional to the force variation when the pedal brake force is constant.

The proposed methods provided a valid calculation of the variation of the pedal acceleration (i.e. the torque variation). Our results show that optical motion tracking of the crank rotation can be used to estimate the tangential force variation applied to the pedal. The results obtained using a commercial video camera were close to the results achieved with a motion capture system provided a correction of the marker position was applied. Thus, an expensive optical device is not necessary to estimate the torque variation. We proposed values of the parameters for the Savitsky–Golay filter that can be used for pedal brake forces ranging between 100 and 250 N with cadences between 60 and 100 rpm for MoCap and video data. The training of a particular pedalling technique can be performed by providing athletes real-time performance feedback based on our proposed calculation of pedal torque variation. The only prerequisite is the availability of an ergometer delivering an adjustable constant pedal brake force, two LEDs on the pedal and a video camera.

For future research, we intend to further improve our results with high-speed camcorders (e.g. 59.94 fps) with a high image resolution. Further, asymmetry indexes for torque variation from

pedal motion in cycling can be extracted from the motion curves. The latter idea is explained in Section 7.2.6. Furthermore, we plan to compensate for the distortion due to bike sway by performing a perspective correction with parameters calculated from the distortion of a reference attached to the bike frame, as explained in Section 7.2.7. Finally, we want to test a prototype system adapted to the bicycle simulator which can register the crossing of each tooth of the rear sprocket over the sensor. From this information we will calculate the crank acceleration and compare it with our methods.

Chapter 7

Conclusions and future research

7.1 Conclusions

Cyclic motion is at the core of many sports, such as running, swimming, or cycling. The study of the corresponding kinematic variables is fundamental for the evaluation of training routines and the assessment of performance. This dissertation has concentrated on the non-linear analysis of single time series of knee motion data during cycling ergometry using the state-of-the-art tools, the multidimensional analysis of the acceleration data of gait motion by means of a novel registration method called equalized DBA (eDBA), the improvement of the classification results of self-organizing maps SOMs with eDBA in the preprocessing step, as well as the estimation of the torque variation in cycling ergometry from noisy 3D motion capturing data (MoCap) or 2D video data with a novel method for the correction of the marker position.

In Chapter 2, motion capturing and the properties human motion data were described. Further, the concepts of variation, stability, end-point and coordinative variability were clarified. In the literature, cyclic motion and movement repetitions are analysed based on these concepts but they are not used uniformly. Thus, a direct comparison of existing results is not always possible, lessening the possibility of following related results. Finally, some reasons to analyse and model the human motion as a dynamical system were listed.

In Chapter 3, the reconstruction theorem (i.e. Takens' theorem) for dynamical systems and the embedding space reconstruction of a multidimensional attractor from time series were reviewed. Further, the calculation of some dynamic invariants from embedding space with related state-of-the-art methods, some problems and limitations of numerical calculation of the embedding space, and the definition of chaos were discussed.

In Chapter 4, non-linear methods and hypothesis testing based on surrogate data were used to analyse whether the aperiodic behaviour of noisy data captured from cycling motion could be considered as chaotic. The critical sensitivity of chaotic systems to both the initial conditions and perturbations may explain the irregular behaviour of the observable signals. However, the time series recorded from knee motion did not show a distinct scaling region in the calculation of the maximal Lyapunov exponent that is typical in chaotic time series. Neither was any strong evidence of chaos based on surrogate data found. In most of the cases, the hypotheses that the time series have features similar to time series generated by (non-stationary) random processes could not be rejected. Nevertheless, we shall also note the possibility that our discriminating statistic could not be powerful enough to detect the difference between the surrogates and the original time series. Future work must seek more powerful statistics in order to check the null hypotheses. Further, an extension of the method for comparison between the discriminating statistics of surrogate and

original data was discussed. This extension included the standard deviation of the discriminating statistic based on the estimation of dynamic invariants for the comparison in order to avoid wrong hypothesis rejections.

In Chapter 5, a novel tool for cycle registration was proposed which facilitates the analysis of movement repetitions and cycles in cyclic motion data. The equalized DBA (eDBA) method calculates the average of a set of cycles based on dynamic time warping (DTW) and a modification of DTW barycentric averaging (DBA). The eDBA algorithm allows the study of features of kinematic variables in cyclic motion depending on the phase using the eDBA average cycle as the reference for phase registration. The parametrization allows studying how the kinematic variables change with the phase angle. From this research, it will be possible to determine the range of motion variability that can be considered optimal for healthy subjects. This optimal range must be defined with the help of motion scientists to test their potential for distinguishing groups of subjects and their suitability for application in training and rehabilitation. Further, a novel definition of the quality of cycle intersections with transversal sections was given providing a criterion for the right choice of the section. Finally, the effects of phase registration using the eDBA method on self-organizing map (SOM) were described. With acceleration data from gait motion, it was shown that the quality of the SOM and the classification rate improved when the phase registration was applied in the preprocessing step.

In Chapter 6, a novel method was developed and validated to compute the variations in the torque applied to the pedals during crank rotation from pedalling motion using an ergometer. The key problem was to estimate reliably the angular pedal acceleration from noisy 3D MoCap or 2D video data. We processed the positional data with a novel automatic method for the correction of the marker position with the help of least squares fitted ellipses. Then, a bandwidth differentiator with special characteristics was applied for the calculation of the angular acceleration. We tested the Savitsky–Golay filter (Savitzky & Golay, 1964) from which we obtained the filtered components required for the calculation of the second derivative of the angular position of the crank. Finally, we solved the torque equilibrium equation for the pedal torque. For validation of the method, we used a direct pedal torque measurement. The results from video became close to the MoCap results when the correction of the marker position was applied. Thus, an expensive optical device is not necessary to estimate the torque variation. We proposed values of the parameters for the Savitsky–Golay filter that can be used for pedal brake forces ranging between 100 and 250 N with cadences between 60 and 100 rpm for MoCap and video data. The training of a particular pedalling technique can be performed by providing athletes real-time performance feedback based on our proposed calculation of pedal torque variation. The only prerequisite is the availability of an ergometer delivering an adjustable constant pedal brake force, two LEDs on the pedal and a video camera.

7.2 Future research

7.2.1 Non-uniform and multidimensional embedding

Since the discovery of Takens (1981) of a topologically faithful reconstruction of the state space of a dynamical system in a lower dimensional space, the so-called embedding space, it became possible to estimate and to study the dynamics from noisy measurements of only a few state variables of the underlying motion. An adequate selection of the embedding parameters minimizes the effects of the presence of noise and the finite amount of data. In previous papers about motion, e.g. see (Small, 2005; Stergiou, 2004), a uniform delay embedding defined by Eq. 3.2 was assumed to be

sufficient. However, recently it was shown that the estimates of dynamic invariants calculated from the numerical reconstruction of well-known attractors using non-uniform embedding defined by Eq. 3.3 are closer to the theoretical values than when using a uniform embedding, especially for quasiperiodic and multiple time-scale time series (Schreiber, 1999; Ragulskis & Lukoseviciute, 2009; Uzal et al., 2011).

The best way of calculating the invariants and the proper embedding space remains under research. However, the optimal parameter values for the embedding dimension and time lag can be simultaneously selected by a global optimization of the proposed cost function of Uzal and Verdes (2010). Different reconstructions (multivariate, time-delayed univariate, uniform or non-uniform) can be directly compared through this method, and thereby the suitability of different embedding settings can be assessed. An advantage of this approach is its fully automatic and objective character, in contrast to, for example, the subjective practitioner-dependent choices of the location of the first local minimum of the mutual information or the value of a threshold characterizing a negligible fraction of false nearest neighbours.

Dynamic invariants in robust embedding spaces for the analysis of noisy data are not sufficiently formalized yet. One option is to observe more than one variable of the ‘presumed’ dynamical system and create a single embedding space joining the embedding spaces of the variables. See (Barnard et al., 2001) for an example. Another option, proposed in (Vlachos & Kugiumtzis, 2010), consists of multidimensional embedding combined with non-uniform embedding. More research must be done about the advantages of these methods and their usability for analysing motion data.

7.2.2 Multifractality

The fractal dimension and the multifractal descriptor are mathematical concepts which have been used for the analysis of signals recorded from the human body. A multifractal system is a generalization of a fractal system in which a single exponent (i.e. the fractal dimension) is not enough to describe its dynamics: instead, a continuous spectrum of exponents (the so-called singularity spectrum) is needed (Balocchi, 2011). The output of a monofractal system has the same fractal properties regardless of initial conditions or of driving forces. In contrast, the output of a multifractal system has different fractal properties that depend on the input conditions and over extended periods of time displays various types of singularities.

Many papers have been devoted to the modelling and analysis of motion over long time scales considering only monofractal properties. Recently, physiological signals, such as the heart rate and the EMG fluctuations of healthy individuals, have been reported to be multifractal (Ivanov et al., 1999; J. Wang et al., 2005; G. Wang, Ren, Li, & Wang, 2007). The finding of multifractality in signals related to the human body has posed new challenges to our understanding of the regulation of human systems (e.g. motion control system). Hence, it would be beneficial to analyse multifractality in cyclic motion. For instance, one software package for the analysis of multifractal time series is described in (Goldberger et al., 2000).

7.2.3 Chaos control theory

The dynamics observed in diverse physiological systems may be associated with the chaotic dynamics arising from the interaction of multiple feedback loops. These feedback mechanisms have several time delays, calculated as the time between the sensing of some disturbance and the generation of an appropriate physiological response. The feedbacks maintain key physiological variables within their normal limits. For details, see (Glass et al., 1988).

Motion variability is an essential feature of the human motor system attributed to its ability to coordinate many different physiological systems over many different timescales. The introduction of non-linear dynamics and chaos theory to the study of human motion has revealed a chaotic structure in its variation. For this reason, the human motor control behaviour has been related to chaotic dynamical systems. The evolution of the output in time of such systems can be completely different even for very close initial conditions and tiny perturbations, making the predictability of these systems tough and even impossible in the long term.

In addition to the interaction of multiple feedback loops, a chaotic system can produce a significant number of dynamical behaviours (e.g. periodic and not periodic behaviours) with the help of tiny perturbations chosen properly. For instance, when a trajectory approaches a desired periodic orbit embedded in the attractor, one applies small perturbations to switch the trajectory to the neighbourhood of the desired periodic orbit that can now be stabilized. The right choice of a tiny perturbation can direct the trajectory of the solution in the state space to wherever one wants in the attractor. Future research must include the theory of control of chaos in how the feedback mechanisms in the human body produce one desired movement through small disturbances, especially for cyclic motion and repetitive tasks. If the tiny perturbations occur very often, more elaborate algorithms are required for the calculation of dynamical invariants and to check the evidence of chaos. For an introduction about chaos control, see (Boccaletti et al., 2000).

7.2.4 Selection of a suitable Poincaré section in an embedding space

The placement of the Poincaré sections, for reconstructed embedding spaces from noisy data as in Chapter 3 as well as for observed spaces as in Chapter 5, is of high relevance for the usefulness of the results. An ‘optimal’ Poincaré section contains intersections with the (reconstructed) trajectory of the attractor which represent better the dynamic properties of the system under study than the others. However, the qualitative assessment of the Poincaré sections has not been well discussed in the literature.

Future research must include the modification of the method for quality of transversal sections presented in Chapter 5 in order to select the ‘optimal’ Poincaré section for embedding data. The embedding can be either uniform described by Eq. 3.2 or non-uniform described by Eq. 3.3. The eDBA average should be calculated from the reconstructed embedding space for the placement of the Poincaré section. Then, the selection of the Poincaré section should be guided by both the SNR of the signal inside the section and the number of intersections, so the SNR of each embedding point must be calculated as the sum of all SNR coordinate points. Finally, the ground truth for comparison can be the dynamical properties of well-known dynamical systems.

7.2.5 Curve registration and curve averaging for continuous time

Two approaches for cycle registration and cycle averaging were proposed in Chapter 5. These methods were proposed for discrete signals but the analysed time series were captured from continuous motion. The discretization may affect the calculation of the average cycle, especially for low sampling rates. Future research should extend our approach to curve registration and averaging to continuous time. Some related research was done in (Boßhammer, 2015), where implementations of both continuous averaging and continuous curve registration methods proposed in (Ramsay, 1997) were compared with the standard DBA and DTW respectively.

7.2.6 Asymmetry indices for torque variation from pedal motion in cycling

The interaction between the cyclist and the bicycle is given by the body position and the pedalling technique. However, the latter has received less attention due to the misconception that the bike goes ahead regardless of how the force on the pedals is applied. A good pedalling technique allows the subject to minimize the energy needed to perform the motion and to avoid the risk of injuries.

Leg motion is produced by muscle synergies which allow several combinations of extensions and flexions of hip, knee, and ankle joints (Hug et al., 2010). These synergies characterize the motion patterns and reflect the individual motor abilities. Thus, the optimal technique depends on how the body combines different muscle synergies, which can produce the same set of performance parameter values.

Bilateral cycling motion has usually been assessed assuming symmetry in force production and kinematics of lower limbs. Yet differences in power output and mechanical work of the legs from 5% to 20% have been reported in uninjured cyclists and non-cyclists (F. P. Carpes, Mota, & Faria, 2010). It is still not clear how pedalling asymmetry affects the performance, but there is a consensus that a reduction of asymmetry may reduce the risk of premature fatigue and overuse injuries. Only a few studies have defined pedalling asymmetry based on measures of (net) torque (F. Carpes et al., 2007; Bertucci, Arfaoui, & Polidori, 2012), work (Sanderson, 1990; Hunt, Sanderson, Moffet, & Inglis, 2004), and force (Smak et al., 1999). For a review of asymmetry in cycling, see (F. Carpes et al., 2007).

Future research must include the analysis of the asymmetry of the crank acceleration curves given in Chapter 6, to improve the pedal technique. For instance, different peak values in the acceleration curves indicate an asymmetry due to emphasizing an unequal pushing action during the downstroke or due to a problem with pulling during the upstroke. Definitions for the asymmetry adapted from the mentioned publications to acceleration curves are given in the following. In a crank rotation, two peaks are visible in the acceleration curves, each one due to the pushing of one of the legs mostly.

- Peak acceleration asymmetry: the (percentage) difference between the values of two peaks in a cycle. See Fig 7.1a.
- Positive acceleration asymmetry: the (percentage) difference between the positive areas of two peaks in a cycle. See Fig 7.1b.
- Negative acceleration asymmetry: the (percentage) difference between the negative areas of two peaks in a cycle. See Fig 7.1c.
- Left-Right acceleration asymmetry: the (percentage) difference between the area of the first and the second half of a cycle. See Fig 7.1d.

7.2.7 Compensation for the bike sway in videos of pedalling motion

For the calculation of the crank acceleration in Chapter 6, it was assumed that the pedal motion lies in a two-dimensional plane. This plane can be easily calculated from the motion capturing data (MoCap) projecting the 3D data onto 2D by using principal component analysis (PCA). However, the motion data in our videos were the result of the 2D projection of the crank motion plane to the camera plane. When the two planes were not aligned, the projection of the circular trajectory of the crank motion to the camera plane resulted in an undesirable oval trajectory. The misalignment

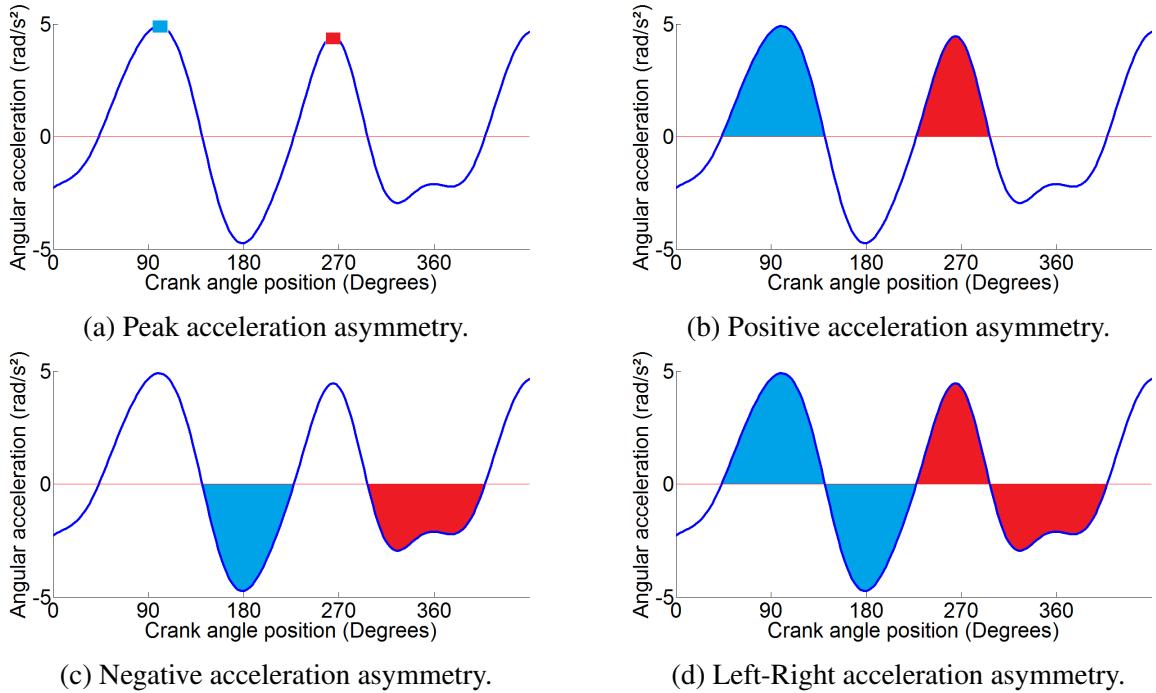


Figure 7.1: Examples of possible asymmetry definitions. The colors indicate the peak values or the areas that are compared.

of both planes happened very often during the pedalling motion due to bike sway, but this problem can be minimized in each frame by performing a perspective correction.

Such a perspective correction requires additional information extracted from a reference in each video image. For example, one or more disk references can be placed on the bicycle frame. When a disk is photographed obliquely, it will appear in photos as an ellipse. The ellipse parameters provide enough information to determine the angle between the camera view and the axis of the disc, e.g. the ratio between the major and minor axes and the tilt angle of the major axis with respect to the vertical in the image. Thus, one or more disk references can provide information about the position of the crank motion plane relative to the camera plane by assuming that angles between the bicycle frame plane and the crank motion plane remain constant. See (Elder, 2014) for one method of determining the rotations between the disc axis and the line of sight from the camera.

Chapter 8

Appendix

8.1 Types of dynamical systems

In the following, some terminology about the types of dynamical system is given, based on (Strogatz, 2000). Two main types of dynamical systems exist: differential equations and iterated maps (also known as difference equations). Differential equations describe the evolution of systems in continuous time, whereas iterated maps arise in problems where time is discrete. Now confining our attention to differential equations, the main distinction is between ordinary and partial differential equations. For instance, the equation for a damped harmonic oscillator

$$m \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx = 0, \quad (8.1)$$

is an ordinary differential equation, because there is only one independent variable, the time t . In contrast, the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad (8.2)$$

is a partial differential equation and it has both time t and space x as independent variables.

Another kind of dynamical system is a system with delayed feedback, for example of the form $\dot{x}(t) = F(x(t - \tau))$. The derivative of the variable at a certain time is given in terms of the values of the function at previous times. These systems can create arbitrarily complex dynamics already with very few variables and rather simple equations of motion. Note that discrete dynamics of the form $x_{k+1} = F(x_k)$ is not equivalent to delay differential equations of the form $\dot{x}(t) = F(x(t - \tau))$. The former is 1-D (precise knowledge of x_k is sufficient to predict the future) whereas the latter is infinite dimensional (you need knowledge of the entire trajectory $x(t)$ over an interval of length τ)¹.

A very general framework for ordinary differential equations is provided by the system

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, \dots, x_n) \\ &\vdots \\ \dot{x}_n &= f_n(x_1, \dots, x_n). \end{aligned} \quad (8.3)$$

Here the overdots denote differentiation with respect to t . Thus $\dot{x}_i \equiv \frac{dx_i}{dt}$ and x_1, \dots, x_n represent the variables. A system is said to be linear if all the variables x_i on the right side appear to the first

¹https://www.researchgate.net/post/Do_we_have_chaotic_behaviour_in_two_dimensional_continuous_dynamical_systems

power only. Otherwise, the system would be non-linear. Typically non-linear terms are products, powers, and functions of the variables x_i , such as x_1x_2 , x_1^3 , or $\cos(x_2)$. Systems that do not include any explicit time dependence are called autonomous. Nonautonomous systems have an explicit time dependence, e.g. the force harmonic oscillator

$$m \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx = F \cos(t). \quad (8.4)$$

However, the time dependence can be removed by adding an extra dimension to the system:

$$\dot{x}_1 = x_2, \quad (8.5)$$

$$\dot{x}_2 = \frac{(-kx_1 - bx_2 + F \cos x_3)}{m}, \quad (8.6)$$

$$\dot{x}_3 = 1. \quad (8.7)$$

This system would traditionally be regarded as a second order linear equation, whereas other authors will regard it as third-order non-linear system (the system is non-linear thanks to the \cos term). Similarly, an n th-order time-dependent equation is a special case of an $(n+1)$ dimensional system.

8.2 Non-linear filtering

Traditional linear filters are based on the assumption that the signal and noise components can be distinguished in the spectrum. For coarsely sampled signals from non-linear systems this poses a problem since the signal itself can have a broadband spectrum. Linear filters may interact unfavourably with the non-linear structure but the non-linear noise reduction of Schreiber (1993) takes into account the non-linear nature of the data in order to not distort the original signal.

Suppose we have a scalar time series $\{x[n]\}$, $n = 1, \dots, l, \dots, N$ sampled signals from a non-linear system, where the $x[n]$ are composed of a clean signal $y[n]$ with some observational noise $\eta[n]$ added, $x[n] = y[n] + \eta[n]$, and $\sigma^2 = \langle \eta[n]^2 \rangle$ is called the absolute noise level.

As the first step for non-linear noise reduction, the uniform delay embedding $\mathbf{v}[n] = (x[n], x[n-\tau], x[n-2\tau], \dots, x[n-(d_e-1)\tau])$ explained in Section 3.2 is used to reconstruct the original state space. Then, each measurement $x[n]$ is replaced by the average value of the first coordinate (i.e. the coordinate without delay) of the embedding vectors $\mathbf{v}[n]$ in the neighbourhood. The neighbours are selected from the neighbourhood of reference point $U_\varepsilon(\mathbf{v}[n]) := \left\{ \mathbf{v} \mid \|\mathbf{v}[n] - \mathbf{v}[l]\| < \varepsilon \right\}$ with distance to the reference point less than a certain distance, ε . Finally, the coordinate $x[n]$ is replaced by the corrected measure $x_{\text{new}}[n]$ given by,

$$x_{\text{new}}[n] = \frac{1}{|U_\varepsilon(\mathbf{v}[n])|} \sum_j x[j], \quad (8.8)$$

with $|U_\varepsilon(\mathbf{v}[l])|$ being the number of nearest neighbours and $\{j \mid \mathbf{v}[j] \in U_\varepsilon(\mathbf{v}[n])\}$. Note that the neighbourhood of each point $x[n]$ at least contains the point itself. If that is the only member, the average Eq. 8.8 is simply the uncorrected measurement and no change is made.

The radius of the neighbourhoods ε should be taken large enough to cover the noise extent. If the noise level is known, an ε about three times the noise level gives the best results for artificial data (Schreiber, 1993; Hegger et al., 1999). Further, Hegger et al. (1999) recommended that the embedding dimension should be larger than the required by the embedding theorems (e.g. two

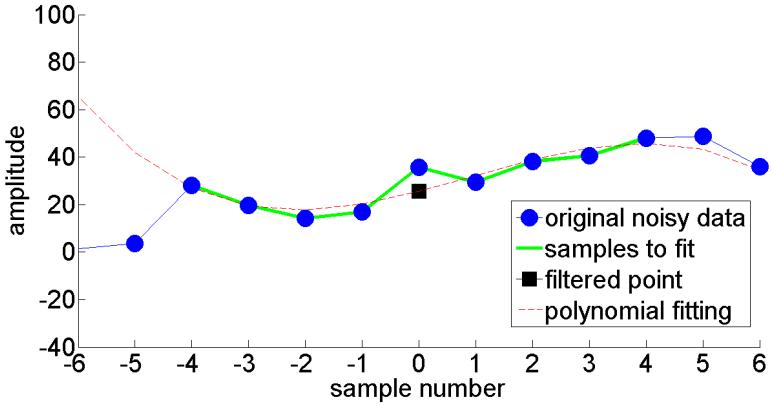


Figure 8.1: Example of polynomial smoothing. The smoothed output value (black square) is obtained by evaluating a polynomial of order $N = 3$ (red line) fitted on a group of $2M + 1$ samples (on the green line), here with half-length $M = 4$, at $n = 0$.

more dimensions) to capture the noisy components better and to make sure that the attractor is well reconstructed despite of the noise.

8.3 Savitzky–Golay smoothing filter

In order to increase the signal-to-noise ratio without greatly distorting the signal, we used the Savitzky–Golay filter (Savitzky & Golay, 1964), also known as polynomial smoothing (Hamming, 1989) or as least-squares smoothing filters (Schafer, 2011).

The Savitzky–Golay filter is a method for data smoothing based on an approximation by a local least-squares polynomial. This filter reduces the noise while maintaining the shape and height of the waveform peaks (e.g. their relative widths and heights), (Press, Teukolsky, Vetterling, & Flannery, 2007). This filter is a generalization of the FIR averaging filter. It can preserve the high frequency content of the desired signal, at the expense of not removing as much noise as the averager, (Orfanidis, 2010).

The basic idea of this filter is to fit a polynomial to a set of consecutive input samples and then evaluate the resulting polynomial at the central point. Fig. 8.1 shows an example of polynomial smoothing for the central point $n = 0$. The input $x[k]$, with $k \in \mathbb{Z}$, is a sequence of discrete points (solid blue dots). A window (i.e. points on the green line) with a window size of $2M + 1$ data points, with $M \geq 1$, is used to calculate a least-squares polynomial fit (here $2M + 1 = 9$). A polynomial $p[k]$ of degree $N = 3$ (red dashed line) is fitted to the data $x[k]$, with $k = n - M, \dots, n + M$ and the smoothed output value (black square) is obtained by evaluating $p[k]$ for $k = n$. We obtain the coefficients a_i of the polynomial p of order N ,

$$p[k] = \sum_{i=0}^N a_i k^i, \quad (8.9)$$

that minimizes the mean-squared approximation error ϵ_n for the group of input samples centered on n ,

$$\epsilon_n = \sum_{m=-M}^M (p[n+m] - x[n+m])^2. \quad (8.10)$$

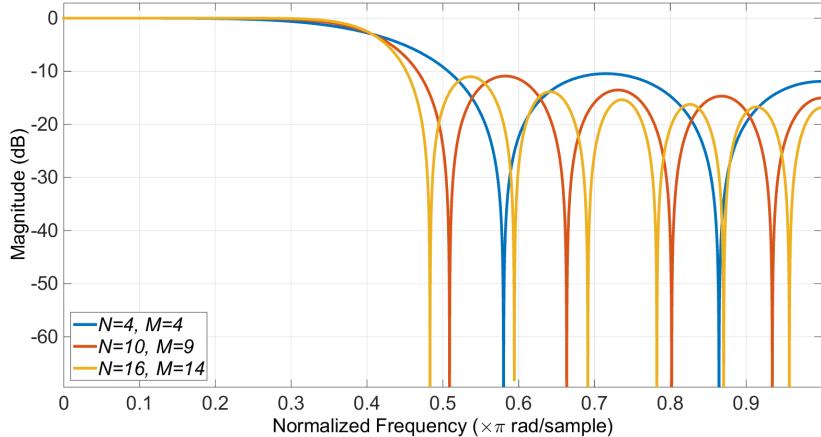


Figure 8.2: The normalized cutoff frequency $f_c = 0.4$ can be achieved by different combinations of the order of the polynomial N and the half-window size M .

It can be shown that this is equivalent to discrete convolution with a fixed impulse response, (Schafer, 2011). Savitzky and Golay (1964) published tables of filter coefficients for combined smoothing and differentiation. These tables are given for two parameters: the half-length of the fitting window, M , and the order of the fitted polynomial, N . The dependence of the cutoff frequency f_c on N and M is given by

$$f_c = \frac{N + 1}{3.2M - 4.6}, \text{ for } M \geq 25 \text{ and } N \leq M. \quad (8.11)$$

Sometimes the same cutoff frequency can be achieved using different combinations of N and M . However, the stopband rejection is different for each combination. Fig. 8.2 shows some frequency responses of filters with a normalized cutoff frequency of 0.4. Fig. 8.3 shows the normalized cutoff frequency for different combinations of these parameters.

8.4 Relaxation method

We recorded the positions of the pivot point of the pedal during crank rotation in Chapter 6. Ideally, after projecting the pedal trajectory on a two-dimensional fitting plane, we should obtain a perfect (sampled) circle. However, the measurements contain noise in the radial and angular directions. In order to simultaneously deal with both types of noise adaptively, one alternative is to jointly project the points onto a fitting circle (or ellipse) and relax the resulting angular values θ_k , $k = 1, \dots, J$ as follows (see Fig. 8.4). We minimize the cost function

$$C(\theta_1, \dots, \theta_J) = \sum_{k=1}^J \Delta q_k^2 + \lambda \sum_{k=2}^{J-1} (\theta_{k+1} - 2\theta_k + \theta_{k-1})^2, \quad (8.12)$$

where Δq_k^2 is the square of the Euclidean distance between the data point q_k and its corresponding projected point p_k , λ is a non-negative relaxation factor, and θ_k is the angular position of the corresponding projection p_k . The relaxation factor λ controls the amount of angular smoothing relative to the radial distortion of the data. Without relaxation, e.g. at $\lambda = 0$, the cost function is minimized for phase angles θ_k corresponding to the orthogonal projections of the data points q_k onto the circle

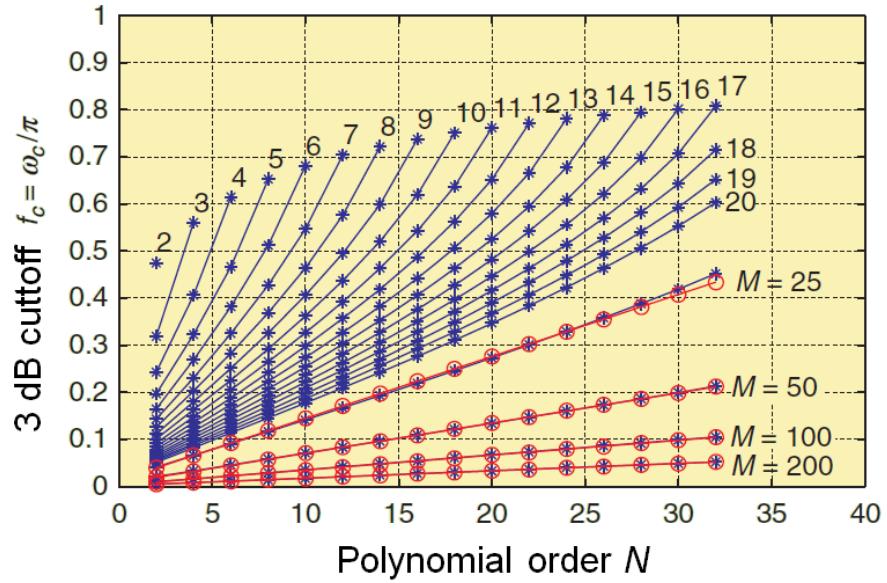


Figure 8.3: Cutoff frequency of Savitsky–Golay filters given the order of the polynomial N and the half-window size M . Figure adapted from Schafer, 2011.

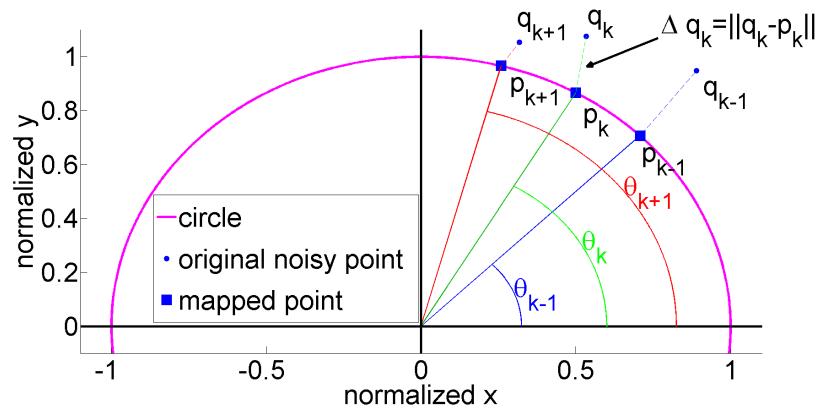


Figure 8.4: Relaxation example: The measured points q are projected onto points p (on the ideal circle or ellipse) minimizing the cost function C .

(or ellipse). As $\lambda \rightarrow \infty$, the relaxation yields an equidistant distribution of samples, i.e. constant angular speed.

We implemented a second approach with the relaxation method with a fitting circle as pre-processing step and then we used the Savitzky–Golay coefficients for the calculation of the angular acceleration. Different parameter combinations were tried for the Savitzky–Golay filters with window sizes $F = 2, 3$ and the polynomial orders $N = 10n + 1, n = 3, \dots, 11$, and for relaxation factors with $\lambda = 10^j, j = 0, \dots, 9$. The best results were obtained with $\lambda = 10^6$. However, this method did not perform well for video data and the results of this approach did not outperform the results obtained in Chapter 6. Future research must study the results using the relaxation method with a fitting ellipse instead of a fitting circle.

8.5 List of acronyms

BMU best matching unit

CNS central nervous system

DBA DTW barycentric averaging

DFA detrended fluctuation analysis

DOF degrees of freedom

DTW dynamic time warping

ECG electrocardiogram

eDBA equalized DBA

EEC electroencephalogram

FIR finite response filter

FNN false nearest neighbor

IIR infinite impulse response filter

LED light-emitting diode

MoCap motion capturing data

PCA principal component analysis

PPS pseudo periodic surrogate

RMS root mean square

RPDE recurrence period density entropy

SNR signal-to-noise ratio

SOM self-organizing map

SSS small-shuffle surrogate

TFTS truncated Fourier transform surrogate

TSS temporal shift surrogate

Bibliography

- Aeyels, D. (1981). Generic observability of differentiable systems. *SIAM Journal on Control and Optimization*, 19(5), 595–603.
- Arikan, O. (2006). Compression of motion capture databases. *Proceedings of ACM SIGGRAPH 2006*, 25(3), 890–897.
- Badii, R., Broggi, G., Derighetti, B., Ravani, M., Ciliberto, S., Politi, A., & Rubio, M. (1988). Dimension increase in filtered chaotic signals. *Physical Review Letters*, 60(11), 979–982.
- Balocchi, R. (2011). Fractal dimension: From geometry to physiology. In *Advanced methods of biomedical signal processing* (pp. 327–346). John Wiley & Sons, Inc.
- Barbivc, J., Safonova, A., Pan, J. Y., Faloutsos, C., Hodgins, J. K., & Pollard, N. S. (2004). Segmenting motion capture data into distinct behaviors. In *Gi '04: Proceedings of graphics interface 2004* (pp. 185–194). Waterloo, Ontario, Canada: Canadian Human-Computer Communications Society.
- Barnard, J. P., Aldrich, C., & Gerber, M. (2001). Embedding of multidimensional time-dependent observations. *Physical Review E*, 64(4), 046201-1–4.
- Bartlett, R., Wheat, J., & Robins, M. (2007). Is movement variability important for sports biomechanists? *Sports Biomechanics*, 6(2), 224–243.
- Barton, G., Lees, A., Lisboa, P., & Attfield, S. (2006). Visualisation of gait data with Kohonen self-organising neural maps. *Gait & Posture*, 24(1), 46–53.
- Beaudoin, P., Poulin, P., & van de Panne, M. (2007). Adapting wavelet compression to human motion capture clips. In *Proceedings of graphics interface 2007* (pp. 313–318). New York, NY, USA: ACM.
- Beek, P., Peper, C., & Stegeman, D. (1995). Dynamical models of movement coordination. *Human Movement Science*, 14(4–5), 573–608.
- Begg, R., & Kamruzzaman, J. (2003). A comparison of neural networks and support vector machines for recognizing young-old gait patterns. In *TENCON 2003. Conference on convergent technologies for the asia-pacific region* (Vol. 1, pp. 354–358).
- Bernstein, N. A. (1996). *Dexterity and its development*. Lawrence Erlbaum.
- Bertucci, W. M., Arfaoui, A., & Polidori, G. (2012). Analysis of the pedaling biomechanics of master's cyclists: A preliminary study. *Journal of Science and Cycling*, 1(2), 42–46.
- Böhm, H., Siebert, S., & Walsh, M. (2008). Effects of short-term training using SmartCrank on cycle work distribution and power output during cycling. *European Journal of Applied Physiology*, 103, 225–232.
- Bini, R., Diefenthäler, F., & Mota, C. (2010). Fatigue effects on the coordinative pattern during cycling: kinetics and kinematics evaluation. *Journal of Electromyography and Kinesiology*, 20(1), 102–107.
- Black, A. (1994). *The effect of steady rate exercise on the pattern of force production of the lower limbs in cycling* (Doctoral dissertation). University of British Columbia, Canada.

- Boccaletti, S., Grebogi, C., Lai, Y. C., Mancini, H., & Maza, D. (2000). The control of chaos: theory and applications. *Physics reports*, 329(3), 103–197.
- Boßhammer, M. (2015). *Aligning & averaging time-dependent data* (Bachelor thesis). Universität Konstanz.
- Broer, H., & Takens, F. (2010). *Dynamical systems and chaos*. Springer.
- Broer, H., Takens, F., & Hasselblatt, B. (2010). *Handbook of dynamical systems* (Vol. 3). North Holland.
- Broomhead, D. S., Huke, J. P., & Muldoon, M. R. (1992). Linear filters and non-linear systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 54(2), 373–382.
- Broomhead, D. S., & King, G. P. (1986). Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena*, 20(2), 217–236.
- Buzug, T., & Pfister, G. (1992). Optimal delay time and embedding dimension for delay-time coordinates by analysis of the global static and local dynamical behavior of strange attractors. *Physical Review A*, 45(10), 7073–7088.
- Cannon, D. T., Kolkhorst, F. W., & Cipriani, D. J. (2007). Effect of pedaling technique on muscle activity and cycling efficiency. *European Journal of Applied Physiology*, 99(6), 659–664.
- Cao, L. (1997). Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D: Nonlinear Phenomena*, 110(1–2), 43–50.
- Cappellini, G., Ivanenko, Y. P., Poppele, R. E., & Lacquaniti, F. (2006). Motor patterns in human walking and running. *Journal of Neurophysiology*, 95(6), 3426–3437.
- Carpes, F., Rossato, M., Faria, I., & Bolli Mota, C. (2007). Bilateral pedaling asymmetry during a simulated 40-km cycling time-trial. *Journal of Sports Medicine and Physical Fitness*, 47(1), 51–57.
- Carpes, F. P., Mota, C. B., & Faria, I. E. (2010). On the bilateral asymmetry during running and cycling - a review considering leg preference. *Physical therapy in sport*, 11(4), 136–142.
- Casaleggio, A., Corana, A., & Ridella, S. (1995). Correlation dimension estimation from electrocardiograms. *Chaos, Solitons & Fractals*, 5(3–4), 713–726.
- Casdagli, M. (1992). A dynamical systems approach to modeling input-output systems. In *- proceedings of santa fe institute studies in the sciences of complexity* (Vol. 12, pp. 265–265).
- Casdagli, M., Eubank, S., Farmer, J., & Gibson, J. (1991). State space reconstruction in the presence of noise. *Physica D: Nonlinear Phenomena*, 51(1–3), 52–98.
- Chau, T., Young, S., & Redekop, S. (2005). Managing variability in the summary and comparison of gait data. *Journal of Neuroengineering and Rehabilitation*, 2(1), 1–22.
- Christensen, L., Johannsen, P., Sinkjær, T., Petersen, N., Pyndt, H., & Nielsen, J. (2000). Cerebral activation during bicycle movements in man. *Experimental Brain Research*, 135(1), 66–72.
- Costa, M., Goldberger, A., & Peng, C. (2005). Multiscale entropy analysis of biological signals. *Physical Review E*, 71(2), 021906-1–18.
- Dahmen, T., & Saupe, D. (2011). Calibration of a power-speed-model for road cycling using real power and height data. *International Journal of Computer Science in Sport*, 10(2), 18–36.
- Davids, K., Bennett, S., & Newell, K. (2006). *Movement system variability*. Human Kinetics Champaign, IL.
- Davids, K., Glazier, P., Araujo, D., & Bartlett, R. (2003). Movement systems as dynamical systems: the functional role of variability and its implications for sports medicine. *Sports Medicine*, 33(4), 245–260.
- Devaney, R. L. (1989). *An introduction to chaotic dynamical systems* (2nd ed.). Addison-Wesley Reading.

- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1), 269–271.
- Ding, M., Grebogi, C., Ott, E., Sauer, T., & Yorke, J. A. (1993). Plateau onset for correlation dimension: When does it occur? *Physical Review Letters*, 70(25), 3872–3875.
- Dingwell, J. B. (2006). Lyapunov exponents. In M. Akay (Ed.), *Wiley encyclopedia of biomedical engineering*. Hoboken, NJ, USA: John Wiley & Sons, Inc.
- Dingwell, J. B., & Cusumano, J. P. (2000). Nonlinear time series analysis of normal and pathological human walking. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 10(4), 848–863.
- Dolan, K., & Spano, M. (2001). Surrogate for nonlinear time series analysis. *Physical Review E*, 64(4), 046128-1–6.
- Elder, J. (2014). *Determining rotations between disc axis and line of sight*. <http://web.ncf.ca/jim/scale/ellipse.html>.
- Ettema, G., & Lorås, H. (2009). Efficiency in cycling: A review. *European Journal of Applied Physiology*, 106(1), 1–14.
- Faria, M. (2009). Recent advances in specific training for cycling. *International Sport Medicine Journal*, 10(1), 16–32.
- Fraser, A. M., & Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical Review A*, 33, 1134–1140.
- Gander, W., Golub, G. H., & Strebel, R. (1994). Least-squares fitting of circles and ellipses. *BIT Numerical Mathematics*, 34(4), 558–578.
- Gandevia, S. C., & Burke, D. (1992). Does the nervous system depend on kinesthetic information to control natural limb movements? *Behavioral and Brain Sciences*, 15(4), 614–632.
- Ghez, J., & Vaienti, S. (1992). Integrated wavelets on fractal sets. *Nonlinearity*, 5, 777–790.
- Gibson, J. F., Farmer, J. D., Casdagli, M., & Eubank, S. (1992). An analytic approach to practical state space reconstruction. *Physica D: Nonlinear Phenomena*, 57(1–2), 1–30.
- Glass, L. (2001). Synchronization and rhythmic processes in physiology. *Nature*, 410(6825), 277–284.
- Glass, L., Beuter, A., & Larocque, D. (1988). Time delays, oscillations, and chaos in physiological control systems. *Mathematical Biosciences*, 90(1), 111–125.
- Glass, L., & Mackey, M. C. (1988). *From clocks to chaos: The rhythms of life*. Princeton University Press.
- Glazier, P. S., Wheat, J. S., Pease, D. L., & Bartlett, R. M. (2006). The interface of biomechanics and motor control. In *Movement system variability* (pp. 49–69). Human Kinetics Champaign, IL.
- Gleick, J. (1987). *Chaos: Making a new science*. New York, NY: Viking.
- Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., ... Stanley, H. E. (2000). Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals. *Circulation*, 101(23), e215–e220.
- Goldsmith, M. (2009). *The maximal Lyapunov exponent of a time series* (Doctoral dissertation). Concordia University.
- Grassberger, P., & Procaccia, I. (1983a). Characterization of strange attractors. *Physical Review Letters*, 50, 346–349.
- Grassberger, P., & Procaccia, I. (1983b). Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, 9(1–2), 189–208.
- Grimaldi, G., & Manto, M. (2012). *Mechanisms and emerging therapies in tremor disorders*. Springer Science & Business Media.

- Gu, Q., Peng, J., & Deng, Z. (2009). Compression of human motion capture data using motion pattern indexing. *Computer Graphics Forum*, 28(1), 1–12.
- Hamill, J., Palmer, C., & Van Emmerik, R. E. (2012). Coordinative variability and overuse injury. *Sports Medicine, Arthroscopy, Rehabilitation, Therapy & Technology*, 4:45(1), 1–9.
- Hamill, J., van Emmerik, R. E., Heiderscheit, B. C., & Li, L. (1999). A dynamical systems approach to lower extremity running injuries. *Clinical Biomechanics*, 14(5), 297–308.
- Hamming, R. (1989). *Digital filters*. Dover Publications.
- Harbourne, R., & Stergiou, N. (2009). Movement variability and the use of nonlinear tools: principles to guide physical therapist practice. *Physical Therapy*, 89(3), 267–282.
- Hausdorff, J. M. (2007). Gait dynamics, fractals and falls: finding meaning in the stride-to-stride fluctuations of human walking. *Human Movement Science*, 26(4), 555–589.
- Hausdorff, J. M., Peng, C. K., Ladin, Z., Wei, J. Y., & Goldberger, A. L. (1995). Is walking a random walk? Evidence for long-range correlations in stride interval of human gait. *Journal of applied physiology*, 78(1), 349–358.
- Hautamäki, V., Nykänen, P., & Fränti, P. (2008). Time-series clustering by approximate prototypes. In *19th international conference on pattern recognition, ICPR 2008* (pp. 1–4).
- Höchtl, F., Böhm, H., & Senner, V. (2010). Prediction of energy efficient pedal forces in cycling using musculoskeletal simulation models. *Procedia Engineering*, 2(2), 3211–3215.
- Hegger, R., Kantz, H., & Schreiber, T. (1999). Practical implementation of nonlinear time series methods: The TISEAN package. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 9(2), 413–435.
- Hug, F., Turpin, N., Guével, A., & Dorel, S. (2010). Is interindividual variability of EMG patterns in trained cyclists related to different muscle synergies? *Journal of Applied Physiology*, 108(6), 1727–1736.
- Hull, M. L., Kautz, S., & Beard, A. (1991). An angular velocity profile in cycling derived from mechanical energy analysis. *Journal of Biomechanics*, 24(7), 577–586.
- Hunt, M. A., Sanderson, D. J., Moffet, H., & Inglis, J. T. (2004). Interlimb asymmetry in persons with and without an anterior cruciate ligament deficiency during stationary cycling. *Archives of physical medicine and rehabilitation*, 85(9), 1475–1478.
- Ivancevic, T. T., & Djukic, S. (2009). *Complex sports biodynamics*. Berlin, Heidelberg: Springer.
- Ivanov, P. C., Amaral, L. A. N., Goldberger, A. L., Havlin, S., Rosenblum, M. G., Struzik, Z. R., & Stanley, H. E. (1999). Multifractality in human heartbeat dynamics. *Nature*, 399(6735), 461–465.
- Janssen, D., Schöllhorn, W. I., Newell, K. M., Jäger, J. M., Rost, F., & Vehof, K. (2011). Diagnosing fatigue in gait patterns by support vector machines and self-organizing maps. *Human Movement Science*, 30(5), 966–975.
- Jirsa, V., Fink, P., Foo, P., & Kelso, J. (2000). Parametric stabilization of biological coordination: A theoretical model. *Journal of Biological Physics*, 26(2), 85–112.
- Joglekar, S. (2015, November). *Self-organizing maps with google's tensorflow*. Retrieved from <https://codesachin.wordpress.com/2015/11/28/self-organizing-maps-with-googles-tensorflow/>
- Kantz, H. (1994). A robust method to estimate the maximal lyapunov exponent of a time series. *Physics Letters A*, 185(1), 77–87.
- Kantz, H., & Schreiber, T. (2004). *Nonlinear time series analysis* (Vol. 7). Cambridge University Press.
- Kaski, S., & Lagus, K. (1996). Comparing self-organizing maps. In *Artificial neural networks ICANN 96* (pp. 809–814). Springer.

- Kautz, S., & Hull, M. (1993). A theoretical basis for interpreting the force applied to the pedal in cycling. *Journal of Biomechanics*, 26(2), 155–165.
- Kennel, M., Brown, R., & Abarbanel, H. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A*, 45(6), 3403–3411.
- Kneip, A., & Gasser, T. (1992). Statistical tools to analyze data representing a sample of curves. *The Annals of Statistics*, 1266–1305.
- Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9), 1464–1480.
- Kohonen, T., Nieminen, I., & Honkela, T. (2009). On the quantization error in SOM vs. VQ: A critical and systematic study. In J. Príncipe & R. Miikkulainen (Eds.), *Advances in self-organizing maps* (Vol. 5629, p. 133-144). Springer Berlin Heidelberg.
- Koktas, N., Yalabik, N., & Yavuzer, G. (2006). Combining neural networks for gait classification. In J. F. Martinez-Trinidad, J. A. Carrasco-Ochoa, & J. Kittler (Eds.), *Progress in pattern recognition, image analysis and applications* (Vol. 4225, p. 381-388). Springer Berlin Heidelberg.
- Korff, T., Romer, L. M., Mayhew, I., & Martin, J. C. (2007). Effect of pedaling technique on mechanical effectiveness and efficiency in cyclists. *Medicine & Science in Sports & Exercise (MSSE)*, 39(6), 991–995.
- Lathrop, D., & Kostelich, E. (1989). Characterization of an experimental strange attractor by periodic orbits. *Physical Review A*, 40(7), 4028–4031.
- Li, S., Okuda, M., & Takahashi, S. (2008). Compression of human motion animation using the reduction of interjoint correlation. *Journal on Image and Video Processing*, 2008(1), 1–15.
- Lin, Y., & McCool, M. (2007). Nonuniform segment-based compression of motion capture data. In G. Bebis et al. (Eds.), *Advances in visual computing* (pp. 56–65). Berlin, Heidelberg: Springer.
- Lipsitz, L. A. (2002). Dynamics of stability: the physiologic basis of functional health and frailty. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 57(3), B115–125.
- Little, M., McSharry, P., Roberts, S., Costello, D., & Moroz, I. (2007). Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *BioMedical Engineering OnLine*, 6(1), 23–27.
- Liu, G., & McMillan, L. (2006). Segment-based human motion compression. In *Proceedings of the 2006 ACM SIGGRAPH* (pp. 127–135). Aire-la-Ville, Switzerland, Switzerland: Eurographics Association.
- Lopes, R., & Betrouni, N. (2009). Fractal and multifractal analysis: A review. *Medical Image Analysis*, 13(4), 634–649.
- Luo, X. (2005). *Characterising nonlinear determinism from experimental time series data* (Doctoral dissertation). The Hong Kong Polytechnic University.
- Luo, X., Nakamura, T., & Small, M. (2005). Surrogate test to distinguish between chaotic and pseudoperiodic time series. *Physical Review E*, 71, 026230-1–8.
- Luo, X., Zhang, J., & Small, M. (2005). Nonparametric rank test for nonlinearity detection. *International Symposium on Nonlinear Theory and its Applications (NOLTA2005)*, 642–645.
- Maiwald, T., Mammen, E., Nandi, S., & Timmer, J. (2008). Surrogate data — a qualitative and quantitative analysis. In R. Dahlhaus, J. Kurths, P. Maass, & J. Timmer (Eds.), *Mathematical methods in signal processing and digital image analysis* (pp. 41–74). Springer.
- Mayer, R., Neumayer, R., Baum, D., & Rauber, A. (2009). Analytic comparison of self-organising maps. In J. Príncipe & R. Miikkulainen (Eds.), *Advances in self-organizing maps* (Vol.

- 5629, p. 182-190). Springer Berlin Heidelberg.
- Merkwirth, C., Parlitz, U., Wedekind, I., Engster, D., & Lauterborn, W. (2009). *OpenTSTOOL user manual*. (Drittes Physikalisches Institut Universität Göttingen)
- Miller, D., Stergiou, N., & Kurz, M. (2006). An improved surrogate method for detecting the presence of chaos in gait. *Journal of Biomechanics*, 39(15), 2873–2876.
- Mitschke, F., Möller, M., & Lange, W. (1988). Measuring filtered chaotic signals. *Physical Review A*, 37(11), 4518–4521.
- Moeslund, T. B., Hilton, A., & Kruger, V. (2006). A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104(2), 90–126.
- Monroig, E. (2009). *Detection of changes in dynamical systems by nonlinear time series analysis* (Doctoral dissertation). University of Tokyo.
- Moon, F. (2008). *Chaotic and fractal dynamics: Introduction for applied scientists and engineers*. Wiley.
- Mornieux, G., Zamezati, K., Mutter, E., Bonnefoy, R., & Belli, A. (2006). A cycle ergometer mounted on a standard force platform for three-dimensional pedal forces measurement during cycling. *Journal of Biomechanics*, 39(7), 1296–1303.
- Nakamura, T., Luo, X., & Small, M. (2005). Testing for nonlinearity in time series without the Fourier transform. *Physical Review E*, 72(5), 055201-1–4.
- Nakamura, T., & Small, M. (2005). Small-shuffle surrogate data: Testing for dynamics in fluctuating data with trends. *Physical Review E*, 72, 056216-1–6.
- Nakamura, T., Small, M., & Hirata, Y. (2006). Testing for nonlinearity in irregular fluctuations with long-term trends. *Physical Review E*, 74, 026205-1–8.
- Nashner, L. M. (1980). Balance adjustments of humans perturbed while walking. *Journal of Neurophysiology*, 44(4), 650–664.
- Nessler, J., De Leone, C., & Gilliland, S. (2009). Nonlinear time series analysis of knee and ankle kinematics during side by side treadmill walking. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19(2), 026104-1–11.
- Newell, K., Deutsch, K., Sosnoff, J., & Mayer-Kress, G. (2006). Variability in motor output as noise: A default and erroneous proposition. In *Movement system variability* (pp. 3–22). Human Kinetics Champaign, IL.
- Newell, K. M., & Corcos, D. M. (1993). *Variability and motor control*. Human Kinetics.
- Niennattrakul, V., & Ratanamahatana, C. A. (2007). Inaccuracies of shape averaging method using dynamic time warping for time series data. In *Computational science ICCS 2007* (pp. 513–520). Springer.
- Orfanidis, S. (2010). *Introduction to signal processing*. Prentice Hall.
- Oullier, O., Marin, L., Stoffregen, T. A., Bootsma, R. J., & Bardy, B. G. (2006). Variability in postural coordination dynamics. In *Movement system variability* (pp. 25–47). Human Kinetics Champaign, IL.
- Ovaska, S., & Valiviita, S. (1998). Angular acceleration measurement: A review. *IEEE Transactions on Instrumentation and Measurement*, 47(5), 1211–1217.
- Packard, N. H., Crutchfield, J. P., Farmer, J. D., & Shaw, R. S. (1980). Geometry from a time series. *Physical Review Letters*, 45, 712–716.
- Parlitz, U. (1998). Nonlinear modeling - advanced black-box techniques. In J. Suykens & J. Vandewalle (Eds.), (pp. 209–239). Kluwer Academic Publisher.
- Peitgen, H. O., Jürgens, H., & Saupe, D. (2006). *Chaos and fractals: new frontiers of science*. Springer Science & Business Media.

- Peng, C., Havlin, S., Stanley, H., & Goldberger, A. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1), 82–87.
- Perc, M. (2005). The dynamics of human gait. *European Journal of Physics*, 26, 525–534.
- Petitjean, F., Forestier, G., Webb, G., Nicholson, A. E., Chen, Y., & Keogh, E. (2014). Dynamic time warping averaging of time series allows faster and more accurate classification. In *IEEE international conference on data mining ICDM 2014* (pp. 470–479).
- Petitjean, F., Ketterlin, A., & Gançarski, P. (2011, 25). A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognition*, 44(3), 678–693.
- Pincus, S. (1995). Approximate entropy (ApEn) as a complexity measure. *Chaos (Woodbury, NY)*, 5(1), 110–117.
- Pollard, C., Heiderscheit, B., van Emmerik, R., & Hamill, J. (2005). Gender differences in lower extremity coupling variability during an unanticipated cutting maneuver. *Journal of Applied Biomechanics*, 21(2), 143–152.
- Pöhlbauer, G. (2004). Survey and comparison of quality measures for self-organizing maps. In J. Paralič, G. Pöhlbauer, & A. Rauber (Eds.), *Proceedings of the fifth workshop on data analysis (WDA'04)* (pp. 67–82). Elfa Academic Press.
- Preda, M., Jovanova, B., Arsov, I., & Preeux, F. (2007). Optimized MPEG-4 animation encoder for motion capture data. In *Web3D '07: Proceedings of the twelfth international conference on 3D web technology* (pp. 181–190). New York, NY, USA: ACM.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes: The art of scientific computing* (3rd ed.). Cambridge University Press.
- Quintana-Duque, J. C. (2011). Kompression menschlicher Bewegungsdaten: Eine Übersicht. In D. Link & J. Wiemeyer (Eds.), *Sportinformatik trifft Sporttechnologie, 8. Symposium der dvs-Sektion Sportinformatik* (Vol. 217, pp. 157–162). Feldhaus Edition Czwalina.
- Quintana-Duque, J. C. (2012). Non-linear dynamic invariants based on embedding reconstruction of systems for pedaling motion. In *Sportinformatik 2012. 9. Symposium der dvs-Sektion Sportinformatik*. (pp. 28–38). Konstanz, Germany.
- Quintana-Duque, J. C., Dahmen, T., & Saupe, D. (2015). Estimation of pedal torque variation from pedal motion in cycling. *International Journal of Computer Science in Sport*, 14(1), 34–50.
- Quintana-Duque, J. C., & Saupe, D. (2012). Evidence of chaos in pedaling motion using nonlinear methods. In *Proceedings of world congress of performance analysis of sport IX* (pp. 30-1–11). Routledge.
- Quintana-Duque, J. C., & Saupe, D. (2015). Phase registration improves classification and clustering of cycles based on self-organizing maps. In *Proceedings of the 2nd international workshop on sensor-based activity recognition and interaction* (pp. 5:1–5:6). New York, NY, USA: ACM.
- Quintana-Duque, J. C., Saupe, D., & Vieten, M. (2014). Parametrization of cyclic motion and transversal sections. In *Sportinformatik X* (pp. 111–117).
- Ragulskis, M., & Lukoseviciute, K. (2009). Non-uniform attractor embedding for time series forecasting by fuzzy inference systems. *Neurocomputing*, 72(10–12), 2618–2626.
- Rakthanmanon, T., Campana, B., Mueen, A., Batista, G., Westover, B., Zhu, Q., ... Keogh, E. (2012). Searching and mining trillions of time series subsequences under dynamic time warping. In *Proceedings of the 18th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 262–270).
- Ramsay, J. (1997). *Functional data analysis*. Wiley Online Library.

- Ramsay, J., Hooker, G., & Graves, S. (2009). Registration: Aligning features for samples of curves. In J. Ramsay, G. Hooker, & S. Graves (Eds.), *Functional data analysis with R and MATLAB* (pp. 117–130). New York, NY: Springer New York.
- Ren, L. (2006). *Statistical analysis of natural human motion for animation* (Doctoral dissertation). Carnegie Mellon University. (Adviser-Hodgins, Jessica K.)
- Rosenhahn, B., Klette, R., & Metaxas, D. (2007). *Human motion: understanding, modelling, capture and animation* (Vol. 36). Springer.
- Rosenstein, M., & Collins, J. (1994). Visualizing the effects of filtering chaotic signals. *Computers & Graphics*, 18(4), 587–592.
- Rosenstein, M., Collins, J., & De Luca, C. (1993). A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena*, 65(1–2), 117–134.
- Rossato, M., Bini, R., Carpes, F., Diefenthäler, F., & Moro, A. (2008). Cadence and work-load effects on pedaling technique of well-trained cyclists. *International Journal of Sports Medicine*, 29(9), 746–752.
- Sadeghi, H., Mathieu, P., Sadeghi, S., & Labelle, H. (2003). Continuous curve registration as an intertrial gait variability reduction technique. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(1), 24–30.
- Sanderson, D. (1990). The influence of cadence and power output on asymmetry of force application during steady-rate cycling. *Journal of Human Movement Studies*, 19, 1–9.
- Sano, M., & Sawada, Y. (1985). Measurement of the Lyapunov spectrum from a chaotic time series. *Physical Review Letters*, 55, 1082–1085.
- Sauer, T., Yorke, J., & Casdagli, M. (1991). Embedology. *Journal of Statistical Physics*, 65(3–4), 579–616.
- Savitzky, A., & Golay, M. J. E. (1964, 1). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8), 1627–1639.
- Schafer, R. (1987). *Clinical biomechanics: Musculoskeletal actions and reactions*. Williams & Wilkins.
- Schafer, R. (2011). What is a Savitzky-Golay filter? *Signal Processing Magazine, IEEE*, 28(4), 111–117.
- Schreiber, T. (1993). Extremely simple nonlinear noise-reduction method. *Physical Review E*, 47, 2401–2404.
- Schreiber, T. (1999). Interdisciplinary application of nonlinear time series methods. *Physics Reports*, 308(1), 1–64.
- Schreiber, T., & Schmitz, A. (1997). Discrimination power of measures for nonlinearity in a time series. *Physical Review E*, 55(5), 5443–5447.
- Shimada, M., & Uehara, K. (2000). Discovering of correlation from multi-stream of human motion. In *DS '00: Proceedings of the third international conference on discovery science* (pp. 290–294). London, UK: Springer.
- Signorini, M. G., & Ferrario, M. (2011). Nonlinear analysis of experimental time series. In *Advanced methods of biomedical signal processing* (pp. 347–378). John Wiley & Sons, Inc.
- Slifkin, A. B., & Newell, K. M. (1998). Is variability in human performance a reflection of system noise? *Current directions in psychological science*, 7(6), 170–177.
- Smak, W., Neptune, R., & Hull, M. (1999). The influence of pedaling rate on bilateral asymmetry in cycling. *Journal of Biomechanics*, 32(9), 899–906.
- Small, M. (2005). *Applied nonlinear time series analysis: applications in physics, physiology and finance* (Vol. 52). World Scientific Pub Co Inc.

- Small, M., Judd, K., & Mees, A. (2001). Testing time series for nonlinearity. *Statistics and Computing*, 11(3), 257–268.
- Small, M., Nakamura, T., & Luo, X. (2007). Surrogate data methods for data that isn't linear noise. *Nonlinear Phenomena Research Perspectives*.
- Small, M., & Tse, C. (2002). Applying the method of surrogate data to cyclic time series. *Physica D: Nonlinear Phenomena*, 164(3–4), 187–201.
- Small, M., Yu, D., & Harrison, R. G. (2001). Surrogate test for pseudoperiodic time series data. *Physical Review Letters*, 87(18), 188101-1–4.
- Sminchisescu, C. (2008). 3d human motion analysis in monocular video: techniques and challenges. In *Human motion* (pp. 185–211). Springer.
- Smith, L. I. (2002). A tutorial on principal components analysis. *Cornell University, USA*, 51, 52.
- Stapelfeldt, B., Mornieux, G., & Gollhofer, A. (2006). Wirkung von Feedback-Training im Radsport auf physiologische und biomechanische Parameter. *BISp-Jahrbuch Forschungsförderung*, 7, 175–179.
- Stapelfeldt, B., Mornieux, G., Oberheim, R., Belli, A., & Gollhofer, A. (2007). Development and evaluation of a new bicycle instrument for measurements of pedal forces and power output in cycling. *International Journal of Sports Medicine*, 28(4), 326–332.
- Stark, J., Broomhead, D. S., Davies, M., & Huke, J. (2003). Delay embeddings for forced systems. II. stochastic forcing. *Journal of Nonlinear Science*, 13(6), 519–577.
- Stergiou, N. (2004). *Innovative analyses of human movement*. Human Kinetics.
- Stergiou, N., & Decker, L. M. (2011). Human movement variability, nonlinear dynamics, and pathology: is there a connection? *Human Movement Science*, 30(5), 869–888.
- Stergiou, N., Harbourne, R., & Cavanaugh, J. (2006). Optimal movement variability: a new theoretical perspective for neurologic physical therapy. *Journal of Neurologic Physical Therapy*, 30(3), 120.
- Strogatz, S. H. (2000). *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering* (1st ed.). Cambridge, MA: Westview Press.
- Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical systems and turbulence* (pp. 366–381). Springer.
- Theiler, J. (1995). On the evidence for low-dimensional chaos in an epileptic electroencephalogram. *Physics Letters A*, 196(5–6), 335–341.
- Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., & Farmer, J. D. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena*, 58(1), 77–94.
- Theiler, J., & Prichard, D. (1997). Using ‘surrogate surrogate data’ to calibrate the actual rate of false positives in tests for nonlinearity in time series. *Fields Institute Communications*, 11, 99.
- Theurel, J., Crepin, M., Foissac, M., & Temprado, J. J. (2011). Effects of different pedalling techniques on muscle fatigue and mechanical efficiency during prolonged cycling. *Scandinavian Journal of Medicine & Science in Sports*, 22(6), 714–721.
- Tufillaro, N. (1994). Topological organization of low-dimensional chaos. In P. Grassberger & J. P. Nadal (Eds.), *From statistical physics to statistical inference and back* (Vol. 428, p. 311–316). Springer Netherlands.
- Tykierko, M. (2007). Using invariants to determine change detection in dynamical system with chaos. *Central European Journal of Operations Research*, 15(3), 223–233.
- Uzal, L. C., Grinblat, G. L., & Verdes, P. F. (2011). Optimal reconstruction of dynamical systems: A noise amplification approach. *Physical Review E*, 84(1), 016223-1–17.

- Uzal, L. C., & Verdes, P. F. (2010). Optimal irregular delay embeddings. *International Conference on Chaos and Nonlinear Dynamics*, 1–2.
- van der Kooij, H., Koopman, B., & van der Helm, F. C. (2008). *Human motion control*.
- Vatanen, T., Osmala, M., Raiko, T., Lagus, K., Sysi-Aho, M., Oresic, M., ... Lähdesmäki, H. (2015). Self-organization and missing values in SOM and GTM. *Neurocomputing*, 147(0), 60–70. (Workshop on Self-Organizing Maps 2012 (WSOM 2012))
- Vereijken, B., van Emmerik, R. E. A., Whiting, H. T. A., & Newell, K. M. (1992). Free(z)ing degrees of freedom in skill acquisition. *Journal of Motor Behavior*, 24(1), 133–142.
- Vieten, M. M., Sehle, A., & Jensen, R. L. (2013). A novel approach to quantify time series differences of gait data using attractor attributes. *PloS one*, 8(8), e71824-1–10.
- Vlachos, I., & Kugiumtzis, D. (2010). Nonuniform state-space reconstruction and coupling detection. *Physical Review E*, 82(1), 016207-1–21.
- Voss, H. U. (2000). Anticipating chaotic synchronization. *Physical Review E*, 61(5), 5115–5119.
- Wang, G., Ren, X., Li, L., & Wang, Z. (2007). Multifractal analysis of surface EMG signals for assessing muscle fatigue during static contractions. *Journal of Zhejiang University. Science A*, 8(6), 910–915.
- Wang, J., Ning, X., Ma, Q., Bian, C., Xu, Y., & Chen, Y. (2005). Multiscale multifractality analysis of a 12-lead electrocardiogram. *Physical Review E*, 71(6), 062902-1–4.
- Winter, D. A. (1984). Kinematic and kinetic patterns in human gait: Variability and compensating effects. *Human Movement Science*, 3(1–2), 51–76.
- Wolf, A., Swift, J., Swinney, H., & Vastano, J. (1985). Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, 16(3), 285–317.
- Zhang, J. (2007). *Detecting and describing pseudo-periodic dynamics from time series* (Doctoral dissertation). The Hong Kong Polytechnic University.
- Zhang, Y. (2015). *Cycling power and crank angular acceleration from video and motion capture* (Master thesis). University of Konstanz.
- Zhao, Y., Sun, J., & Small, M. (2008). Evidence consistent with deterministic chaos in human cardiac data: Surrogate and nonlinear dynamical modeling. *International Journal of Bifurcation and Chaos*, 18(1), 141–160.