Recurrence Quantification Analysis on Gait Reaction Forces of Elderly Adults for Determination of Pathological States

Özgür Afşar*

Department of Physics, Faculty of Science, Ege University, 35100 Izmir, Turkey
*ozgur.afsar@ege.edu.tr

Abstract

A better classification between patients with parkinson disease and healthy adults is of great importance for clinicians and directly affects the selection of treatment method, the adjustment of medication dose, or even the decision about a dopaminergic therapy. Clinicians widely use semi-objective/subjective assessments in order to be able to differ patients from healthy adults. Here, to make an objective classification between two distinct groups (healthy/patient), we apply a powerful method, recurrence quantification analysis, on data including trajectory behavior of gait reaction forces with long length collected from elderly patients with Parkinson disease and healthy adults as they walk. We show that the complexity measures of the quantification analysis, determinism, entropy and divergence, behave different for two distinct groups (healthy/patients) and may be used for an objective classification.

Keywords: Complexity measures, Recurrence Plot, Gait Reaction Force, Parkinson Disease.

1. Introduction

Parkinson Disease (PD) is a common, debilitating neurodegenerative disease. PD occurs in about 1% of the population over the age of 60 and being widespread increases with age (about 20% of people over the age of 80). The disease affects gait and mobility related to the motor functions and causes functional disorder and death of vital nerve cells producing dopamine being a chemical messenger that sends messages to the part of brain that controls movement and coordination such as gait ability and balance [1].

Doing a better discrimination between patients with PD and healthy adults is important for clinicians so that it directly affects the selection of treatment method, adjustment of medication dose, or even the decision about dopaminergic therapy [2]. Clinicians often use semi-objective/subjective methods such as Hoehn-Yahr staging, which is the most popular subjective scaling used wordwide, Unified Parkinson's Disease Rating Scale (UPDRS), Short Falls Efficacy Scale-International (Short FES-I), etc [3, 4, 5]. Due to the subjectivism of the clinical methods, in the literature, there are many efforts in order to be able to make an optimal classification using objective ones [6].

In this sense of the objective methods, gait variables of patients with PD and healthy adults as control group have been widely studied to differ patients and healthy adults using some analysis regarding informatics approaches or linear/nonlinear methods on gait reaction force (GRF) collected from adults as they were walk normally [7]. It has been reported in the literature that fluctuations in stride times for healthy adults have self-similar fractal structures and they are correlated variables possessing some memory effects and hidden temporal structures, However, for a patient with PD, these self-similar structures disappear and fluctuations with short magnitude about the mean in stride times become completely random (i.e., statistically independent) [8]. Hausdorff and co-workers originally showed for the these data sets using Detrended Fluctuation Analysis (DFA) that the data from healthy subjects generally possess fractal scaling indices of around $0.8 - 1.0$, while the DFA scaling exponent becomes close to 0.5 for gait rhythm of a patient with PD. Any value between 0.8 and 1.0 represents long-range correlations, whereas values close to 0.5 will be representative of uncorrelated, random occurrences [9]. In our previous work, we showed the relation between success rate of classification of adults as healthy/patient and Hoehn-Yahr stages using entropy-based complexity measures, Shannon, Kullback-Leibler and Renormalized entropies [10]. Bernard-Elezari et al. discriminated patients with PD from healthy older adults with a high accuracy by applying a machine learning algorithm on the data from patients with PD [11]. Recently, we made a classification between healthy adults and patients

with PD (not elderly) applying recurrence quantification analysis on gait data with short length [12].

In clinical studies, it is not usual and easy to collect large number of strides from patients with PD [13]. Therefore, here, we apply a powerful method, Recurrence Plot [14], for analysis of both short/long time series in clinical conditions and classification of distinct regimes (healthy/patient, high risk/low risk, periodic/chaotic, etc).

Recurrence is one of the fundamental features of dynamical systems and introduced by Poincare in 1980 [15]. Poincare recurrence theorem states that almost all trajectories of dynamical systems will return very close to their previous positions after a sufficiently long time. Investigation of dynamical properties systems by recurrences can be visualized in two dimensional plot which is called Recurrence Plot (RP) and be derived quantification analysis, RQA, are based on point density and line structures visible in the RP. This technique shows robustness and stability, and provides an alternative for quantifying order and disorder of physical systems. In order to understand the underlying dynamics of complex systems, this method has been applied to various real-world systems in neuroscience, financial science to geophysical and climate systems as well as model systems [16].

We organized the letter as follows:

i. We show efficiency of RP for distinct systems, periodic, quasi-periodic, Brownian and random (stochastic) after technical details of RP and RQA.
ii. Using a windowing procedure, we apply the recurrence technique on the data including swing forces in GRF from elderly patients with PD and healthy adults and show a classification between adults as healthy/patient in RP as a visualization.
iii. We present the complexity measures, determinism, entropy and divergence, of the RQA for a distinction of the groups which consist of elderly patients and healthy adults.

2. Materials and Methods

2.1 Database and pre-processing

In this work, the data set collected by Hausdorff's group is used [17]. They developed a footswitch system including gait variables (GRF) of patients with PD and healthy adults and used a force-measuring unit [18] consisting of a pair of shoes including a total 16 force transducers with a surface of about 10 cm^2 each (Figure 1). The database which contains gait variables of patients with PD (mean age: 66.3 years, 63% men) and 73 healthy controls (mean age: 66.3 years, 55% men) is available in 'PhysioBank Database' [19].

For the database, vertical ground reaction forces (in Newton) were recorded using 8 force sensors underneath each foot when the subjects walked normally (no tasking conditions) on a self-selected pace for approximately 120 seconds in a

well-lit, obstacle free, 25-m long corridor. The output of 16 sensors were digitized and recorded with $\Delta t = 0.01$ sec.

In Figure 2, it can be seen that the total GRF is a time series during successive strides of an adult. In this research, we use the data of total GRF under the left foot (Swing Force), which is strongly correlated with swing of the right foot, whenever total GRF under the right foot equals zero during successive swings. The corresponding data that we have used in this research has length N = 2000. We apply the RQA for each
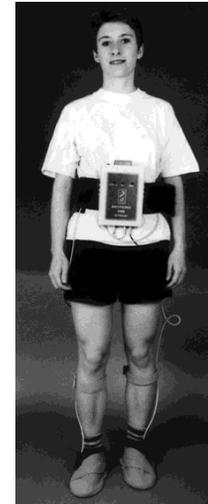
Figure 1. Gait analysis system based on a force-measuring unit consisting of a pair of shoes including force transducers and a measuring device attached to the patient's body through a cable [18].

swing using a windowing technique (totally 10 window with window length $N_W = 200$) and calculate the complexity measures of RQA for patients with PD and healthy adults, separately. We totally use 9 elderly patients (Ages$\geq$ 80, Mean
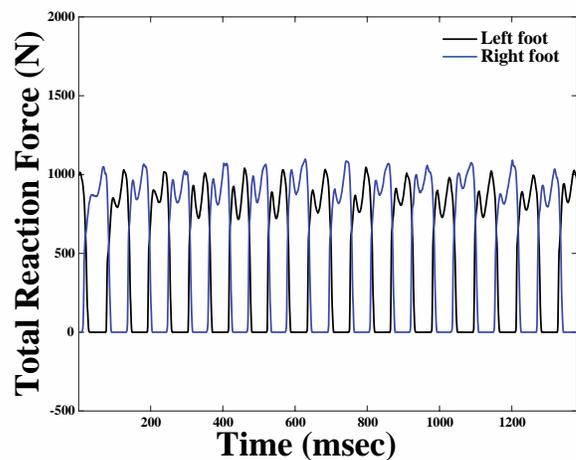
Figure 2. Total ground reaction force (GRF) underneath left and right foot during successive strides of an adult.

age $= 81.4 \pm 1.2$) with PD, since the database only have 9 elderly patients in this age range, and 9 elderly healthy adults

(Mean age $= 76.8 \pm 5.4$) as control group in accordance with the number of elderly patients with PD.

2.2 Recurrence Plot and Quantification Analysis

In 1987, Eckmann et al. proposed the recurrence plots (RP) to describe the recurrence property of a dynamical system and to visualize the time dependent behaviour of orbit x_i in a phase space [20]. The RP can be expressed by the N x N matrix

$$R_{i,j}(\varepsilon) = \Theta\left(\varepsilon - \|\overrightarrow{x_i} - \overrightarrow{x_j}\|\right), \quad i,j = 1, \dots, N, \quad (2.1)$$

where $N = L - (m-1)\tau$, ε is a threshold distance, $\Theta(x)$ is

the choice of embedding dimension and time delay, similarly, if m is too small, the geometry is not entirely unfolded, on the other hand, it leads to a round off error if m is too large. Moreover, choosing a proper time delay τ is based on the detection of the first local minimum of the mutual information as the most common method in the literature [14].

To show efficiency of the RP method, in Figure 3, we plot distinct time series and corresponding RPs regarding periodic, quasi-periodic, Brownian and random motions simulating from the sine function, the quasi-periodic circle map, the cumulative sums of random displacemets of single particle in one dimension and the white noise, respectively. The diogonal lines emerge in RP for both periodic and quasi-periodic motions so that these are long diogonals and with equal distances for periodic signal although they are shorter and with

Figure 3. The distinct time series (upper) and corresponding recurrence plots of periodic, quasi-periodic, brownian and random motions from the left to right one, respectively (m=3, τ=2, N_n/N=0.05).

the Heaviside function (i.e. $\Theta(x) = 0$ if $x < 0$, and $\Theta(x) = 1$ if $x > 0$) and $\|\cdot\|$ is a norm (i.e. Euclidean norm) [14]. Phase space vector $\overrightarrow{x_i}$ can be reconstructed using Taken's time delay method, $x_i = (u_i, u_{i+\tau}, \dots, u_{i+(m-1)\tau})$, based on the observations u_i where m is embedding dimension and τ is time delay. The recurrence plot is obtained by plotting the recurrence matrix in Eq. (1). Recurrence states which are in ε-neighbourhood (i.e. $\left(\varepsilon - \|\overrightarrow{x_i} - \overrightarrow{x_j}\|\right)>0$) state that $R_{i,j}(\varepsilon) = 1$ and it is plotted a black dot at the coordinates (i,j), otherwise states that $R_{i,j}(\varepsilon) = 0$ which leads to a white dot at the coordinates (i,j) in RP.

Several criteria for the choice of the threshold value have been advocated in the literature [14]. One of these uses a fixed number of neighbours (FAN) N_n for every point of trajectory that the threshold value ε_i changes for each state to set up all columns of RP have the same point density the way that recurrence rate adjusted a fixed value with $RR = N_n/N$ [20]. It should be noted that there may be almost no recurrence point if ε is chosen too small, but almost every point is a neighbour of the others in phase space if ε is too large. For

different distances for quasiperiodic one. For the Brownian motion, band structured patterns with single isolated points occur although it can be only seen isolated points for random (uncorrelated) motion.

Several measures can be calculated using RQA to identify dynamics and the degree of order/disorder of systems. The simplest measure of RQA is the recurrence rate (RR) or per cent recurrence

$$RR(\varepsilon) = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j}(\varepsilon), \quad (2.2)$$

which is a measure of density of recurrence points in RP. The average number of neighbours that each point on the trajectory has in its ε-neighbourhood is defined by

$$N_n(\varepsilon) = \frac{1}{N} \sum_{i,j=1}^{N} R_{i,j}(\varepsilon). \quad (2.3)$$

The next measures are based on the histogram $P(\varepsilon, l)$ of diagonal lines of length l, i.e.

$$P(\varepsilon, l) = \sum_{i,j=1}^{N} \left(1 - R_{i-1,j-1}(\varepsilon)\right)\left(1 - R_{i+l,j+l}(\varepsilon)\right) \dots$$
$$\prod_{k=0}^{l-1} R_{i+k,j+k}(\varepsilon) \quad (2.4)$$

The ratio of recurrence points that form diagonal structures (of at least length $l_{min} = 2$ in this research) to all recurrence points

$$DET = \frac{\sum_{l=l_{min}}^{N} lP(\varepsilon, l)}{\sum_{l=1}^{N} lP(\varepsilon, l)} \qquad (2.5)$$

is a measure of determinism (or predictability) of system. It should be noted that deterministic systems possess longer diagonals and less single isolated points in RP whereas stochastic or chaotic systems have none or very short diagonals.

The Shannon entropy of the probability distribution of the diagonal line lengths $P(\varepsilon, l)$ is defined by

$$ENT = -\sum_{l=l_{min}}^{N} P(\varepsilon, l) \ln P(\varepsilon, l) \qquad (2.6)$$

Another RQA measure is about length of longest diagonal L_{max} in the RP, or its inverse, the measure of divergence

$$L_{max} = \max(\{l_i\}_{i=1}^{N_l}), \text{ respectively } DIV = \frac{1}{L_{max}}, \qquad (2.7)$$

where $N_l = \sum_{l \geq l_{min}} P(\varepsilon, l)$ is the total number of diagonal lines. The measure of the divergence is related to the exponential divergence of the phase space trajectory [14, 20].

In this letter, as a first step of RQA, we reconstruct the m-dimensional phase space trajectory. To define appropriate delay time τ, mutual information as a function of the time delay is calculated for each swing forces from gaits of the healthy adults and the patients with PD. The optimal values of τ, based on the detection of the first local minimum of the mutual information function, vary around 10.22 ∓ 2.41 (mean and standard deviation) for the data sets. The optimal delay time, $\tau = 10$, is selected in this range for reconstruction of phase space. To determine proper embedding dimension, the percentage of false nearest neighbours for replaced values of the embedding dimension is calculated for the data sets regarding swing forces. The optimal embedding dimension m, based on the criteria of the percentage of false nearest neighbours being less than 1%, ranges in 5.17 ∓ 1.34 (mean and standard deviation) for the data sets. Therefore, we choose $m = 5$ for the topologically proper reconstruction of the phase space.

3. Results and Discussion

Firstly, we show the differences between visualisations of RPs obtained from gait variables (swing forces) of an elderly healthy adult and an elderly patient with PD randomly chosen from two distinct groups.

In Figure 4, it can be seen that the portrait of the RP from swing forces collected from the healthy adult is very similar to those of from the time series of quasiperiodic motion due to long diagonal lines with different distances. The diagonals start to break down in the RP of the data from the elderly patients with PD as can be seen in Figure 5. This case points

out that there is a kind of transition in gait variables from the healthy adults to the patients.

Secondly, in order to define type of the transition between data of the adults, we apply the RQA using a windowing technique on the data coming from swing forces and we show
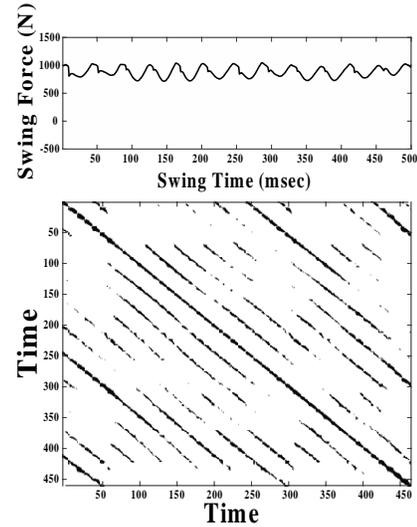
Figure 4. Swing forces as a time series (upper) and corresponding RP (lower) of an elderly healthy adult ($m=3$, $\tau=2$, $N_n/N=0.05$).

tendencies of the complexity measures of RQA (*DET*, *ENT* and *DIV*) through each window for the patient with PD and the healthy adult in Figure 6, seperately. It can be clearly seen that the healthy adult and the patient with PD have different tendencies through the windows. For the heallthy adult, *DIV*
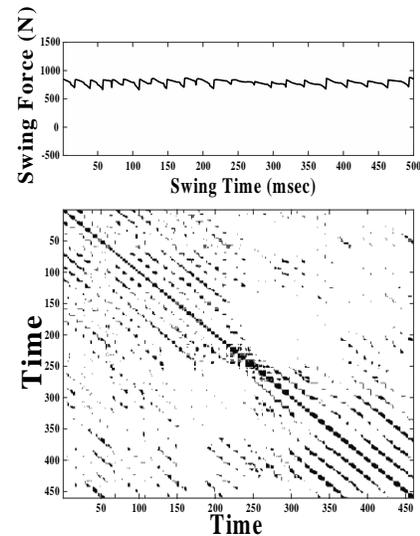
Figure 5. Swing forces as a time series (upper) and corresponding RP (lower) of an elderly patient with PD ($m=3$, $\tau=2$, $N_n/N=0.05$).

measure posssess lower tendency than those of the patient with PD as *DET* and *ENT* measures have higher tendencies.
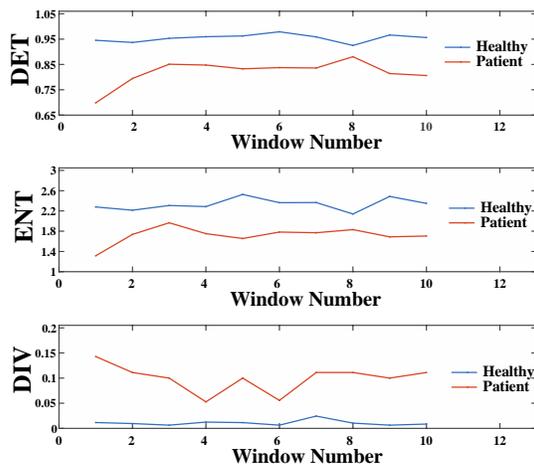
Figure 6. From top to bottom, *DET*, *ENT* and *DIV* as the complexity measures from RQA on the data from swing forces of the elderly healthy adult and the elderly patient with PD ($m=3$, $\tau=2$, $N_n/N=0.05$).

Decreasing determinism and increasing divergence may characterize three kind of transitions between distinct regimes, quasiperiodic-periodic, periodic-chaotic or quasiperiodic-chaotic. However, also decreasing entropy quarantees that the transition is from the quasiperiodic regime to the periodic one which leads to increasing degree of order. It is well known that this feature was shown to be related to the ability of quasiperiodic system to encode more information, in the Shannon sense, than periodic ones.
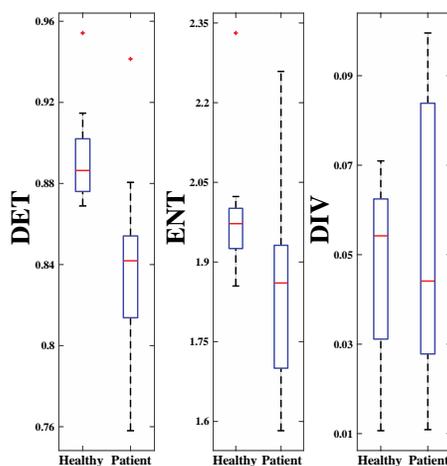
Figure 7. Box plot of the average values of the RQA measures over windows for two distinct group including elderly healthy adults and patient with PD ($m=3$, $\tau=2$, $N_n/N=0.05$). The lower and upper lines of the box are the 25th and 75th percentiles of the sample, the distance between the top and bottom of the box is the inter quartile range and the line in the middle of the box is the sample median. Outliers (plus sign) are cases with values that are more than 1.5 times the interquartile range.

Thirdly, we present box plot of the average values of the RQA measures over windows for two distinct group including elderly healthy adults and patient with PD (totally 18 adults) in Figure 7. Although two distinct groups as healthy and patient can be classified according to their median andmean values, if the standard deviations in the values of measures is considered, *DET* measure of RQA makes a better classification than the others between the distinct groups. In the light of these results, specially, *DET* measure can be propose as a method for an objective classification between healthy/patient adults.

Lastly, we summarize the results in Table 1 including median, mean and standard deviations of the groups, which regards measures of the RQA for the healthy/patient adults.

Table 1. The values of median, mean and standard deviations of the groups regarding healthy/patient adults.

RQA Measures	Health Group		Patient Group	
	Median	Mean±Std	Median	Mean±Std
DET	0.886	0.89±0.03	0.842	0.84±0.05
ENT	1.972	1.99±0.14	1.860	1.84±0.20
DIV	0.054	0.046±0.02	0.044	0.051±0.03

4. Conclusions

We have shown the efficiency of the RP method using different short time series including periodic, quasi-periodic, brownian and random dynamics, and have displayed differences of dynamics as the RP visualitions. Later, we have applied the recurrence technique on gait reaction force database and have shown the differences in the RPs for an elderly healthy control and a patient with PD. Using a windowing method on the time series, we have also shown that RQA measures of the subjects differ from the healthy adult to the patient one in terms of their tendencies. From the healthy adult to the patient one, trends of the measures lead to a scale-down for the *DET* and *ENT* measures as the behavior has a scale-up for the *DIV* measure. This kind of changings of the measures may characterize that there is a transition from quasi-periodic regime to periodic one in the the gait data. Lastly, we have apllied the analysis (RQA) on all adults in two distict group which include healthy adults and patients with PD, separately, and we have shown that *DET* measure from the RQA is very convenient to make an objective classification between healthy and patient adults.

References

1. Hausdorff, J.M, Gait Dynamics in Parkinson disease: Common and distinct behavior among stride length, gait variability, and fractal like scaling, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2009, 19(2), 026113.

2. Whittle, M.W, Clinical gait analysis: A review. *Human Movement Science*, 1996, 15, 369–387.

3. Simon, S.R, Quantification of human motion: gait analysis-benefits and limitations to its application to clinical problems, *Journal of biomechanics*, 2004, 37, 1869–1880.

4. Kempen, G.I, et al., The Short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling, *Age and Ageing*, 2008, 37, 45-50.

5. Goetz, C.G, Tilley, B.C, Shaftman, S.R, Stebbins, G.T, Fahn, S, Martinez-Martin, P, Poewe, W, Sampaio, C, Stern, M.B, Dodel, R, Dubois, B, Holloway, R, Jankovic, J, Kulisevsky, J, Lang, A.E, Lees, A, Leurgans, S, LeWitt, P.A, Nyenhuis, D, Olanow, C.W, Rascol, O, Schrag, A, Teresi, J.A, van Hilten, J.J, LaPelle, N, Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorder*, 2008, 23(15), 2129-2170.

6. Rabey, J. M, Korczyn, A. D. The Hoehn and Yahr rating scale for Parkinson's disease Instrumental Methods and Scoring in Extrapyramidal Disorders, Springer, 1995.

7. Toosizadeh. N, Mohler, J, Lei, H, Parvaneh, S, Sherman, S, Najafi, B, Motor Performance Assessment in Parkinson's Disease: Association between Objective In-Clinic, Objective In-Home, and Subjective/Semi-Objective Measures, *Plos One*, 2015, 10(4), e0124763.

8. Hausdorff, J.M, Lertratanakul, A, Cudkowicz, M.E, Peterson, A.L, Kaliton, D, Goldberger, A.L, Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis, *Journal of applied physiology*, 2000, 88, 2045–2053.

9. Bartsch, R, Plotnik, M, Kantelhardt, J.W, Havlin, S, Giladi, N, Hausdorff, J.M, Fluctuation and synchronization of gait intervals and gait force profiles distinguish stages of parkinson's disease, *Physica A: Statistical Mechanics and its Applications*, 2007, 383, 455–465.

10. Hausdorff, J.M, Peng, C.K, Ladin, Z, Wei, J.Y, Goldberger, A.L, Is walking a random walk? Evidence for long-range correlations in stride interval of human gait, *Journal of Applied Physiology*, 1995, 383, 349-358.

11. Afsar, O, Tirnakli, U, Kurths, J, Entropy-based complexity measures for gait data of patients with parkinson's disease. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2016, 26, 023115.

12. Bernad-Elazari, H, Herman, T, Mirelman, A, Gazit, E, Giladi, N, Hausdorff, J.M, Objective characterization of daily living transitions in patients with parkinson's disease using a single body-fixed sensor, *Journal of neurology*, 2016, 263, 1544-1551.

13. Afsar, O, Tirnakli, U, Marwan, N, Recurrence quantification analysis at work: Quasi-periodicity based interpretation of gait force profiles for patients with Parkinson disease, *Nature-Scientific Reports*, 2018, in press.

14. Kirchner, M, Schubert, P, Liebherr, M, Haas, C.T, Detrended Fluctuation Analysis and Adaptive Fractal Analysis of Stride Time Data in Parkinson's Disease: Stitching Together Short Gait Trials, *Plos One*, 2014, 9, 1-6.

15. Marwan, N., Romano, M. C., Thiel, M. & Kurths, J. Recurrence plots for the analysis of complex systems, *Physics Reports*, 2007, 438, 237–329.

16. Poincare, H, Sur le probleme des trois corps et les equations de la dynamique, *Acta mathematica*, 1890, 13, 1-270.

17. Marwan, N, Schinkel, S, Kurths, J, Recurrence plots 25 years later-gaining confidence in dynamical transitions, *EPL (Europhysics Letters)*, 2013, 101, 20007.

18. Hausdorff, J.M, Ladin, Z, Wei, J.Y, Footswitch system for measurement of the temporal parameters of gait, *Journal of biomechanics*, 1995, 28, 347–351.

19. Bazner, H, Oster, M, Daffertshofer, M, Hennerici M, Assessment of gait in subcortical vascular encephalopathy by computerized analysis: a cross-sectional and longitudinal study, *Journal of Neurology*, 2000, 415, 841-849.

20. Goldberger, A.L, Amaral, L.A, Glass, L, Hausdorff, J.M, Ivanov, P.C, Mark, R.G, Mietus, J.E, Moody, G.B, Peng, C.K, Stanley, H.E, PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals, *Circulation*, 2000, 101, e215-e220.

21. Eckmann, J.P, Kamphorst, S.O, Ruelle, D, Recurrence plots of dynamical systems, *EPL (Europhysics Letters)*, 1987, 4, 973-977.