Plagiat

Michele Bacchereti:
Sviluppo di modelli non lineari per l'analisi di segnali Elettroencefalografici
Universita' di Pisa, Italy, February 2004
http://etd.adm.unipi.it/theses/available/etd-02052004-123236/

Original

Norbert Marwan:
Encounters with neighbours -
Current developments of concepts based on recurrence Plots and their applications
Universität Potsdam, Germany, May 2003
http://opus.kobv.de/ubp/volltexte/2005/89/

Plagiat

Si deve precisare che un singolo punto della RP *(i,j)* non contiene nessuna informazione sullo stato del sistema agli istanti *i* e *j*, ma è dalla totalità della RP che è possibile ricostruire alcune proprietà del sistema. La ricostruzione della traiettoria nello spazio delle fasi è possibile a partire dal RP binario, dove però è stata persa completamente l'informazione sulla lunghezza dei singoli vettori.

In pratica non molto utile, e solitamente impossibile, trovare ricorrenze esatte nel senso che $y_i = y_j$ ma ciò che si ricava è una informazione a proposito di stati che sono *sufficientemente* vicini.

Ciò che noi possiamo conoscere è che lo stato y_i cade in uno *spazio vicino* (secondo una specifica norma definita ad esempio una sfera definita dalla norma L_2) di raggio ε e centrato in y_j in maniera *ricorrente*. Definiamo questi y_j ***punti ricorrenti***. Tutto questo è stato espresso nella Eq. 3.3.

Definiamo ***movimenti tangenziali*** gli stati appartenenti a un corridoio intorno alla LOI[1].

In generale per il calcolo degli indici che verranno definiti, come per l'integrale di correlazione è bene escludere insieme alla LOI anche tali ***movimenti tangenziali.***

Nella definizione originale di RP, la porzione di spazio delle fasi *vicina* era una sfera il cui raggio è definito in modo che questa contenga un numero fisso di stati y_i [8]. In questo modo il raggio ε_i dipende dallo stato y_i *(i= 1,...,N)* e quindi $R_{i,j} \neq R_{j,i}$, perché lo *spazio vicino* di y_j non è detto che sia uguale a quello di y_j. Questa definizione porta a RPs asimmetriche. Questa definizione è stata sostituita da un'altra, che è diventata successivamente di uso comune, nella quale si sceglie un raggio fisso $\varepsilon_i = \varepsilon, \forall i$, utilizzata per primi da [12]. Un raggio fisso porta come conseguenza ad avere RPs simmetriche $R_{i,j} = R_{j,i}$. Lo scegliere l'una o l'altra definizione dipende dal tipo di applicazione usata. Per quanto riguarda lo studio svolto si è scelto di lavorare con uno *spazio vicino* fisso.

[1] Line of identity (diagonale principale per cui $R_{i,i}$=1)

Original

gle recurrence point at (i, j) does not contain any information about the current states at the times i and j. However, from the totality of all recurrence points it is possible to reconstruct the properties of the data. McGuire et al. (1997) have shown the preservation of the dynamical properties for the distance matrix (2.11). However, the phase space trajectory can also be reconstructed from the binary RP, where the information about the absolute length of the phase space vectors is lost. The RP provides information for reordering the indices of the phase space vectors so that the vectors are sorted by their norm. If the cumulative distribution of the lengths of the phase space vectors is known, the restored phase space trajectory will recover its amplitude by equating the sorted indices with this distribution (Thiel, 2003).

In practice it is not useful and largely impossible to find complete recurrences in the sense $\vec{x}_i = \vec{x}_j$ (e. g. the state of a chaotic system would not recur exactly to the initial state but approaches the initial state arbitrarily close). Therefore, a recurrence is defined as a state $\vec{x}_j$ is sufficiently close to $\vec{x}_i$. This means that those states $\vec{x}_j$ that fall into an m-dimensional neighbourhood (e. g. a ball for the L_2-norm or a box for the L_∞ -norm) with a radius ε_i centered at $\vec{x}_i$ are recurrent. These $\vec{x}_j$ are called *recurrence points*. In (2.6), this is simply expressed by the Heaviside function and its argument ε_i.

In the original definition of the RPs, the neighbourhood is a ball (i. e. L_2-norm is used) and its radius is chosen in such a way that it contains a fixed amount of states $\vec{x}_j$ (Eckmann et al., 1987). With such a neighbourhood, the radius ε_i changes for each $\vec{x}_i$ $(i = 1 \ldots N)$ and $\mathbf{R}_{i,j} \neq \mathbf{R}_{j,i}$ because the neighbourhood of $\vec{x}_i$ does not have to be the same as that of $\vec{x}_j$. This property leads to an asymmetric RP, but all columns of the RP have the same recurrence density (Fig. 2.5D). Using this neighbourhood criterion we will use the parameter ε for the predefinition of the recurrence density. This means that with a given $\varepsilon = 0.15$ the real, locally chosen ε_i is adjusted in such a way that the neighbourhood covers 15% of all phase space vectors, and thus the recurrence density is 0.15. We denote this neighbourhood as *fixed amount of nearest neighbours (FAN)*. However, the most commonly used neighbourhood is that with a fixed radius $\varepsilon_i = \varepsilon, \forall i$. For RPs this neighbourhood was firstly used by Zbilut et al. (1991). A fixed radius means that $\mathbf{R}_{i,j} = \mathbf{R}_{j,i}$ resulting in a symmetric RP. The type of neighbourhood that should be used depends on the application. Especially in applications of the later introduced cross recurrence plots, the neighbourhood with a FAN will play an important role.

In order to compute an RP, a norm has to be chosen. The most known norms are the L_1-norm, the L_2-norm (Euclidean norm) and the L_∞ -norm (Maximum or Supremum norm). The neighbourhoods of these norms have differ-

Plagiat

3.3 Ulteriori definizioni di RP

In letteratura sono state proposte diverse definizioni di RP :

- Iwanski e Bradley [13] hanno dato una definizione che usa una soglia con un range $[\varepsilon_{in}, \varepsilon_{out}]$:

$$R_{i,j}^{m,[\varepsilon_{in},\varepsilon_{out}]} = \theta(\|y_i - y_j\| - \varepsilon_{in}) * \theta(\varepsilon_{out} - \|y_i - y_j\|)$$

Eq. 3.4

I punti y_j sono *ricorrenti* quando cadono nell'intervallo spaziale fra il raggio interno ε_{in} e quello esterno ε_{out}. Il vantaggio di usare una RP con soglia tipo range è quello di avere una maggiore robustezza contro i *punti ricorrenti* che derivano traiettorie tangenziali. Questa costruzione delle RPs non è però utilizzabile in una analisi quantitativa.

- Choi et al [14] hanno introdotto la RP perpendicolare:

$$R_{i,j}^{m,\varepsilon} = \theta(\varepsilon - \|y_i - y_j\|) * \delta(y_i * \|y_i - y_j\|)$$

Eq. 3.5

dove δ è la delta di Dirach. Questa RP contiene tutti i punti *vicini* y_j che cadono nello *spazio vicino* di y_i, che però giacciono nel sottospazio $\Re^{m-1}$ di $\Re^m$ e sono perpendicolari alla traettoria nello spazio delle fasi di y_i.

Original

as artifacts.

Some authors exclude the LOI from the RP. This may be useful for the quantification of RPs, which we will discuss later. It can also be motivated by the definition of the Grassberger-Procaccia correlation sum (Grassberger and Procaccia, 1983) which was introduced for the determination of the correlation dimension D_2 and is closely related to RPs:

$$\mathbf{C}^{m,\varepsilon} = \frac{1}{N^2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \Theta\left(\varepsilon - \|\vec{x}_i - \vec{x}_j\|\right). \quad (2.7)$$

The correlation integral excludes the tests of $\vec{x}_i$ with itself. Nevertheless, since the threshold value ε is finite (and normally about 10% of the mean phase space radius), further long diagonal lines can occur directly below and above the LOI for smooth or high resolution data. Therefore, the diagonal lines in a small corridor around the LOI correspond to the *tangential motion* of the phase space trajectory, but not to different orbits. Thus, for the estimation of invariants it is better to exclude this entire predefined corridor and not only the LOI. This step corresponds with suggestions to exclude the tangential motion as it is done for the computation of the correlation dimension (known as Theiler correction or Theiler window; Theiler, 1986) or for the alternative estimators of Lyapunov exponents (Gao and Zheng, 1994) in which only those phase space points are considered that fulfil the constraint $|j - i| \geq w$. Theiler (1986) has suggested using the autocorrelation time as an appropriate value for w, and Gao and Zheng (1994) state that $w = (m-1)\tau$ would be a sufficient approach. However, in a representation of an RP it is better to keep the LOI. Later, this LOI will gain more importance when extensions of the recurrence plot strategies will be discussed.

In the literature further variations of the recurrence plots can be found:

- Iwanski and Bradley (1998) have defined a variation of an RP with a corridor threshold $[\varepsilon_{in}, \varepsilon_{out}]$ (Fig. 2.5E),

$$\mathbf{R}_{i,j}^{m,[\varepsilon_{in},\varepsilon_{out}]} = \Theta\left(\|\vec{x}_i - \vec{x}_j\| - \varepsilon_{in}\right) \cdot \Theta\left(\varepsilon_{out} - \|\vec{x}_i - \vec{x}_j\|\right). \quad (2.8)$$

Those points $\vec{x}_j$ are considered to be recurrent that fall into the shell with the inner radius ε_{in} and the outer radius ε_{out}. The authors have suggested to use this kind of RPs in order to study "interesting structures" in the RP. An advantage of such a *corridor thresholded recurrence plot* is its increased robustness against recurrence points coming from the tangential motion. However, the threshold corridor removes the inner points in broad diagonal lines, which results in two lines instead of one. These

Plagiat

3.3 Ulteriori definizioni di RP

In letteratura sono state proposte diverse definizioni di RP :

- Iwanski e Bradley [13] hanno dato una definizione che usa una soglia con un range $[\varepsilon_{in}, \varepsilon_{out}]$:

$$R_{i,j}^{m,[\varepsilon_{in},\varepsilon_{out}]} = \theta\left(\left\|y_i - y_j\right\| - \varepsilon_{in}\right) * \theta\left(\varepsilon_{out} - \left\|y_i - y_j\right\|\right)$$

Eq. 3.4

I punti y_j sono *ricorrenti* quando cadono nell'intervallo spaziale fra il raggio interno ε_{in} e quello esterno ε_{out}. Il vantaggio di usare una RP con soglia tipo range è quello di avere una maggiore robustezza contro i *punti ricorrenti* che derivano traiettorie tangenziali. Questa costruzione delle RPs non è però utilizzabile in una analisi quantitativa.

- Choi et al [14] hanno introdotto la RP perpendicolare:

$$R_{i,j}^{m,\varepsilon} = \theta\left(\varepsilon - \left\|y_i - y_j\right\|\right) * \delta\left(y_i * \left\|y_i - y_j\right\|\right)$$

Eq. 3.5

dove δ è la delta di Dirach. Questa RP contiene tutti i punti *vicini* y_j che cadono nello *spazio vicino* di y_i, che però giacciono nel sottospazio $\Re^{m-1}$ di $\Re^m$ e sono perpendicolari alla traettoria nello spazio delle fasi di y_i.

Original

RPs are, therefore, not suitable for a quantification analysis. The usage of a shell as a neighbourhood can be found in an algorithm for computing Lyapunov exponents from experimental time series (Eckmann et al., 1986).

- Choi et al. (1999) have introduced the *perpendicular recurrence plot* (Fig. 2.5F)

$$\mathbf{R}_{i,j}^{m,\varepsilon} = \Theta\left(\varepsilon - \left\|\vec{x}_i - \vec{x}_j\right\|\right) \cdot \delta\left(\dot{\vec{x}}_i \cdot (\vec{x}_i - \vec{x}_j)\right). \tag{2.9}$$

Here, δ is the Delta function. This recurrence plot contains only those points $\vec{x}_j$ that fall into the neigbourhood of $\vec{x}_i$ and lie in the $(m-1)$-dimensional subspace of $\mathfrak{R}^m$ that is perpendicular to the phasespace trajectory at $\vec{x}_i$. These points correspond locally to those lying on a Poincaré section. This criterion cleans up the RP more from recurrence points based on the tangential motion than the previous corridor thresholded RPs. The authors have shown the increased efficiency of the perpendicular RPs for their application on estimation of the largest Lyapunov exponent. Using this kind of an RP, the finding of unstable periodic orbits (if they exist) is more robust.

- The RP contains, finally, tests of all states with each other, which results in N^2 tests for N considered states. Still, it is also possible to test each state with a predefined amount k of subsequent states (Zbilut et al., 1991; Koebbe and Mayer-Kress, 1992; Atay and Altıntaş, 1999)

$$\mathbf{R}_{i,j}^{m,\varepsilon} = \Theta\left(\varepsilon - \left\|\vec{x}_i - \vec{x}_{i+i_0+j-1}\right\|\right), \quad i = 1 \dots N-k,\, j = 1 \dots k. \tag{2.10}$$

This reveals an $(N-k) \times k$-matrix which does not have to be square (Fig. 2.5H). The y-axis represents the time distances to the following recurrence points but not their absolute time. All diagonal oriented structures in the common RP are now projected to the horizontal orientation. For $i_0 = 0$, the LOI, which was the diagonal line in the common RP, is now the horizontal line on the x-axis. With non-zero i_0 the RP contains recurrences of a certain state only in the predefined time interval after time i_0 (Koebbe and Mayer-Kress, 1992).

This representation of recurrences may be more intuitive than the RPs usually are because the consecutive states are not oriented diagonally. However, such an RP represents only the first $(N-k)$ states. Mindlin and Gilmore (1992) have proposed the *close returns plot* which is, in fact, such an RP exactly for one dimension. Using this kind of RP, a first quantification approach of RPs (or "close returns plots") can be found ("close

Plagiat

individuare la tipologia del RP sotto esame e su cui si possono calcolare degli indici quantitativi di notevole interesse.

In particolare si possono avere matrici delle ricorrenze con una prevalenza di:

- *Omogeneità:* tipica di sistemi stazionari nei quali tutti i tempi caratteristici sono piccoli rispetto alla lunghezza della serie.
- *Drift:* se il sistema presenta un parametro lentamente variabile lungo l'arco di osservazione della serie.
- La presenza di oscillazioni ricorsive danno RP con orientazioni diagonali e evidenziano *periodicità'* nella RP. Nella Figura 3-3c'è un chiaro esempio di un sistema periodico con 2 frequenze in rapporto 1 a 4 (le linee diagonali principali sono divide da 4 linee più corte; *rapporti irrazionali forniscono strutture più complesse*).
- Bruschi cambiamenti nella dinamica causano *aree bianche o bande* nella RP. Può essere utile per rilevare eventi rari andando a vedere la frequenza di alcuni stati.

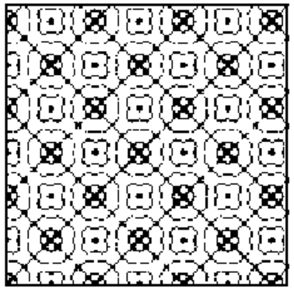

Figura 3-3 Recurrence Plot di un segnale periodico

Original

according to Eckmann et al. (1987) in this work.

It should be emphasized again that the recurrence of states is an important feature. Beside the recurrence plots, there are some other methods that use recurrences. For example, the recurrence in the phase space is used for the recurrence time statistics (Kac, 1947; Gao, 1999; Balakrishnan et al., 2000), first return map (Lathrop and Kostelich, 1989), space time separation plot (Provenzale et al., 1997) or as a measure for nonstationarity (Kennel, 1997; Rieke et al., 2002, closely related to the recurrence time statistics).

2.2.2 Structures in Recurrence Plots

The initial purpose of RPs is the visual inspection of higher dimensional phase space trajectories. The view on RPs gives hints about the time evolution of these trajectories. The advantage of RPs is that they can also be applied to rather short and even nonstationary data.

The RPs exhibit characteristic large scale and small scale patterns. The first patterns were denoted by Eckmann et al. (1987) as *typology* and the latter as *texture*. The typology offers a global impression which can be characterized as *homogeneous*, *periodic*, *drift* and *disrupted*.

- *Homogeneous* RPs are typical of stationary and autonomous systems in which relaxation times are short in comparison with the time spanned by the RP. An example of such an RP is that of a random time series (Fig. 2.6A).
- Oscillating systems have RPs with diagonal oriented, *periodic* recurrent structures (diagonal lines, checkerboard structures). The illustration in Fig. 2.6B is a rather clear periodic system with two frequencies and a frequency ratio of four (the main diagonal lines are divided by four crossing short lines; irrational frequency ratios cause more complex periodic recurrent structures). However, even for those oscillating systems whose oscillations are not easily recognizable, the RPs can be used in order to find their oscillations (an example can be found in Eckmann et al., 1987, cp. unstable periodic orbits).
- The *drift* is caused by systems with slowly varying parameters. Such slow (adiabatic) change brightens the RP's upper-left and lower-right corners (Fig. 2.6C).
- Abrupt changes in the dynamics as well as extreme events cause *white areas or bands* in the RP (Fig. 2.6D). RPs offer an easy possibility to find

Michele Bacchereti: Sviluppo di modelli non lineari per l’analisi di segnali Elettroencefalografici

Plagiat

individuare la tipologia del RP sotto esame e su cui si possono calcolare degli indici quantitativi di notevole interesse.

In particolare si possono avere matrici delle ricorrenze con una prevalenza di:

- *Omogeneità:* tipica di sistemi stazionari nei quali tutti i tempi caratteristici sono piccoli rispetto alla lunghezza della serie.
- *Drift:* se il sistema presenta un parametro lentamente variabile lungo l’arco di osservazione della serie.
- La presenza di oscillazioni ricorsive danno RP con orientazioni diagonali e evidenziano *periodicità’* nella RP. Nella Figura 3-3c’è un chiaro esempio di un sistema periodico con 2 frequenze in rapporto 1 a 4 (le linee diagonali principali sono divide da 4 linee più corte; *rapporti irrazionali forniscono strutture più complesse*).
- Bruschi cambiamenti nella dinamica causano *aree bianche o bande* nella RP. Può essere utile per rilevare eventi rari andando a vedere la frequenza di alcuni stati.

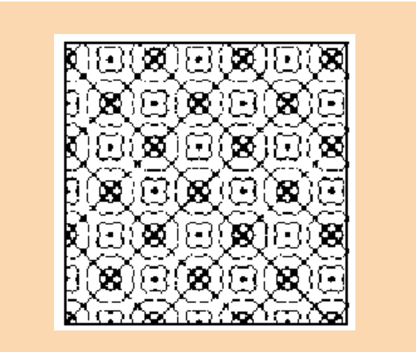

Figura 3-3 Recurrence Plot di un segnale periodico

Norbert Marwan: Encounters with neighbours -
Current developments of concepts based on recurrence Plots and their applications

Original

and to assess extreme and rare events by using the frequency of their recurrences.

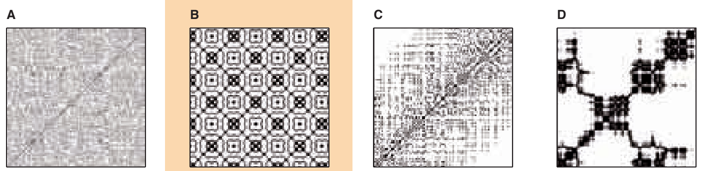

Figure 2.6: Characteristic typology of recurrence plots: (A) homogeneous (uniformly distributed noise), (B) periodic (super-positioned harmonic oscillations), (C) drift (logistic map $x_{i+1} = 4x_i(1 - x_i)$ corrupted with a linearly increasing term, cp. Fig. 2.9D) and (D) disrupted (Brownian motion). These examples illustrate how different RPs can be. The used data have the length 400 (A, B, D) and 150 (C), respectively; no embeddings are used; the thresholds are $\varepsilon = 0.2$ (A, C, D) and $\varepsilon = 0.4$ (B).

The closer inspection of the RPs reveals small scale structures (the texture) which are *single dots*, *diagonal lines* as well as *vertical* and *horizontal lines* (the combination of vertical and horizontal lines obviously forms rectangular clusters of recurrence points).

- *Single, isolated recurrence points* can occur if states are rare, if they do not persist for any time or if they fluctuate heavily. However, they are not a unique sign of chance or noise (for example in maps).
- A *diagonal line* $\mathbf{R}_{i+k,j+k} = 1$ (for $k = 1 \ldots l$, where l is the length of the diagonal line) occurs when a segment of the trajectory runs parallel to another segment, i. e. the trajectory visits the same region of the phase space at different times. The length of this diagonal line is determined by the duration of such similar local evolution of the trajectory segments. The direction of these diagonal structures can differ. Diagonal lines parallel to the LOI (angle $\pi/4$) represent the parallel running of trajectories for the same time evolution. The diagonal structures perpendicular to the LOI represent the parallel running with contrary times (mirrored segments; this is often a hint for an inappropriate embedding). Since the definition of the Lyapunov exponent uses the time of the parallel running of trajectories, the relationship between the diagonal lines and the Lyapunov exponent is obvious (further explanation in Subsec. 2.2.3).

Michele Bacchereti: Sviluppo di modelli non lineari per l'analisi di segnali Elettroencefalografici

Plagiat

3.4.2 *Definizione del "Small scale texture"*

Per quanto riguarda la small-scale texture questa può evidenziare:

- ***Punti ricorrenti singoli ,isolati*** : questi appaiono se sono presenti stati del sistema rari, che ricorrono raramente, che non persistono se non per un istante o fluttuano rapidamente.
- ***Linee diagonali*** (a cui abbiamo fatto frequentemente accenno) definite

  $R_{i+k,j+k} = 1$ per k=1,…,l dove l è la lunghezza delle linee diagonali.

  queste compaiono quando segmenti di traiettorie diverse sono paralleli, ad esempio quando le traiettorie visitano la stessa regione dello spazio delle fasi in tempi diversi. La lunghezza di queste diagonali è determinata dalla durata di tali evoluzioni locali adiacenti della traiettoria. La direzione di queste strutture può differire. Quelle prima definite sono strutture diagonali parallele alla LOI (a 45°) che corrispondono a traiettorie che sono parallele con la stessa evoluzione temporale.

  Esistono anche strutture parallele alla diagonale a 135° :

  $R_{i+k,j-k} = 1$ per k=1,…,l dove l è la lunghezza delle linee diagonali.

  rappresentanti invece di segmenti di traiettoria con evoluzione temporale contraria.
- ***Linee verticali (orizzontali)*** definite:

  $R_{i,j+k} = 1$ per k=1,…,h dove h è la lunghezza delle linee verticali

  denotano periodi temporali dove lo stato del sistema non cambia o cambia molto lentamente.

Le strutture di piccola scala sono alla base dell'analisi quantitativa delle RPs.
L'interpretazione qualitativa delle RPs basata su una visione d'insieme necessita una certa esperienza e può essere utile a fornire informazioni di massima sul sistema. In ogni caso una analisi quantitativa offre un'interpretazione più oggettiva del sistema. Questo tipo di analisi ha fatto sì che le RPs acquisissero maggiore popolarità ed un certo accredito nella comunità scientifica. Andando a riassumere le considerazioni possibili in una analisi qualitativa nella interpretazione delle RPs:

Norbert Marwan: Encounters with neighbours -
Current developments of concepts based on recurrence Plots and their applications

Original

and to assess extreme and rare events by using the frequency of their recurrences.

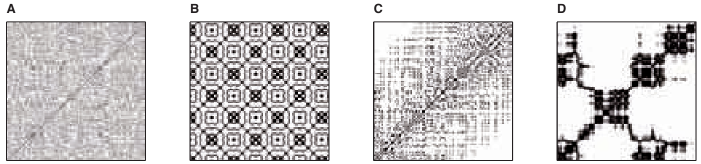

Figure 2.6: Characteristic typology of recurrence plots: (A) homogeneous (uniformly distributed noise), (B) periodic (super-positioned harmonic oscillations), (C) drift (logistic map $x_{i+1} = 4x_i(1 - x_i)$ corrupted with a linearly increasing term, cp. Fig. 2.9D) and (D) disrupted (Brownian motion). These examples illustrate how different RPs can be. The used data have the length 400 (A, B, D) and 150 (C), respectively; no embeddings are used; the thresholds are $\varepsilon = 0.2$ (A, C, D) and $\varepsilon = 0.4$ (B).

The closer inspection of the RPs reveals small scale structures (the texture) which are *single dots*, *diagonal lines* as well as *vertical* and *horizontal lines* (the combination of vertical and horizontal lines obviously forms rectangular clusters of recurrence points).

- *Single, isolated recurrence points* can occur if states are rare, if they do not persist for any time or if they fluctuate heavily. However, they are not a unique sign of chance or noise (for example in maps).
- A *diagonal line* $\mathbf{R}_{i+k,j+k} = 1$ (for $k = 1 \dots l$, where l is the length of the diagonal line) occurs when a segment of the trajectory runs parallel to another segment, i. e. the trajectory visits the same region of the phase space at different times. The length of this diagonal line is determined by the duration of such similar local evolution of the trajectory segments. The direction of these diagonal structures can differ. Diagonal lines parallel to the LOI (angle $\pi/4$) represent the parallel running of trajectories for the same time evolution. The diagonal structures perpendicular to the LOI represent the parallel running with contrary times (mirrored segments; this is often a hint for an inappropriate embedding). Since the definition of the Lyapunov exponent uses the time of the parallel running of trajectories, the relationship between the diagonal lines and the Lyapunov exponent is obvious (further explanation in Subsec. 2.2.3).

3.4.2 Definizione del "Small scale texture"

Per quanto riguarda la small-scale texture questa può evidenziare:

- ***Punti ricorrenti singoli ,isolati*** : questi appaiono se sono presenti stati del sistema rari, che ricorrono raramente, che non persistono se non per un istante o fluttuano rapidamente.
- ***Linee diagonali*** (a cui abbiamo fatto frequentemente accenno) definite

  $R_{i+k,j+k} = 1$ per k=1,…,l dove l è la lunghezza delle linee diagonali.

  queste compaiono quando segmenti di traiettorie diverse sono paralleli, ad esempio quando le traiettorie visitano la stessa regione dello spazio delle fasi in tempi diversi. La lunghezza di queste diagonali è determinata dalla durata di tali evoluzioni locali adiacenti della traiettoria. La direzione di queste strutture può differire. Quelle prima definite sono strutture diagonali parallele alla LOI (a 45°) che corrispondono a traiettorie che sono parallele con la stessa evoluzione temporale.

  Esistono anche strutture parallele alla diagonale a 135° :

  $R_{i+k,j-k} = 1$ per k=1,…,l dove l è la lunghezza delle linee diagonali.

  rappresentanti invece di segmenti di traiettoria con evoluzione temporale contraria.

- ***Linee verticali (orizzontali)*** definite:

  $R_{i,j+k} = 1$ per k=1,…,h dove h è la lunghezza delle linee verticali

  denotano periodi temporali dove lo stato del sistema non cambia o cambia molto lentamente.

Le strutture di piccola scala sono alla base dell'analisi quantitativa delle RPs.

L'interpretazione qualitativa delle RPs basata su una visione d'insieme necessita una certa esperienza e può essere utile a fornire informazioni di massima sul sistema. In ogni caso una analisi quantitativa offre un'interpretazione più oggettiva del sistema. Questo tipo di analisi ha fatto sì che le RPs acquisissero maggiore popolarità ed un certo accredito nella comunità scientifica. Andando a riassumere le considerazioni possibili in una analisi qualitativa nella interpretazione delle RPs:

- A *vertical (horizontal) line* $\mathbf{R}_{i,j+k} = 1$ (for $k = 1 \dots v$, with v the length of the vertical line) marks a time length in which a state does not change or changes very slowly. It seems, that the state is trapped for some time. This is a typical behaviour of laminar states (intermittency).

These small scale structures are the base of a quantitative analysis of the RPs.

The examples in Fig. 2.6 illustrate how different the small scale patterns can be. A large amount of single points and the vanishing amount of lines are caused by heavy fluctuation, which is characteristic for uncorrelated noise (Fig. 2.6A). The relationship between periodically recurrent structures and oscillators is obvious (Fig. 2.6B). The exact recurrent dynamics cause long diagonal lines separated by a fixed distance. The nonregular occurrence of short as well as of long diagonal lines is characteristic for chaotic processes (Fig. 2.6C), whereas the nonregular occurrence of extended black clusters and extended white areas corresponds with a nonregular behaviour in the system, which could be, for example, correlated noise (Fig. 2.6D).

In a more general sense the line structures in an RP exhibit locally the time relationship between the current trajectory segments. A line structure in an RP of length l corresponds to the closeness of the segment $f(T_1(t))$ to another segment $f(T_2(t))$, where $T_1(t)$ and $T_2(t)$ are the local time scales (or transformations of an imaginary absolute time scale t) which preserve that $f(T_1(t)) \approx f(T_2(t))$ for some time $t = 1 \dots l$. Under some assumptions (e. g. piecewise existence of an inverse of the transformation $T(t)$) the local slope $\mathfrak{m}(t)$ of a line in an RP represents the local time derivative of the product of the inverse second time scale $T_2^{-1}(t)$ and the first time scale $T_2(t)$

$$\mathfrak{m}(t) = \partial_t T_2^{-1}(T_1(t)) . \qquad (2.14)$$

We will consider here an illustrative example. A further explanation of the relationship between the slope of the lines and the trajectories is given in the Subsec. about cross recurrence plots (2.3.2). Let us consider a function $f(T) = T(t)$ with a section of a monotonical, linear increase $T_{lin} = t$ and another (hyperbolic) section which follows $T_{hyp} = -\sqrt{r^2 - t^2}$ (Fig. 2.7A) and both sections visit the same area in the phase space. Since the inverse of the hyperbolic section is $T_{hyp}^{-1} = \sqrt{r^2 - t^2}$, the derivative

$$\mathfrak{m} = \partial_t T_{lin}^{-1}\left(T_{hyp}(t)\right) = \frac{t}{\sqrt{r^2 - t^2}} \qquad (2.15)$$

corresponds to the derivative of a circle line with a radius r, a bowed line structure with the form of a circle occurs in the RP (Fig. 2.7C).

Summarizing the last mentioned points we can establish the following list of observations and give the corresponding qualitative interpretation:

3.4.2 Definizione del "Small scale texture"

Per quanto riguarda la small-scale texture questa può evidenziare:

- ***Punti ricorrenti singoli ,isolati*** : questi appaiono se sono presenti stati del sistema rari, che ricorrono raramente, che non persistono se non per un istante o fluttuano rapidamente.
- ***Linee diagonali*** (a cui abbiamo fatto frequentemente accenno) definite

  $\boldsymbol{R_{i+k,j+k} = 1}$ per k=1,…,l dove l è la lunghezza delle linee diagonali.

  queste compaiono quando segmenti di traiettorie diverse sono paralleli, ad esempio quando le traiettorie visitano la stessa regione dello spazio delle fasi in tempi diversi. La lunghezza di queste diagonali è determinata dalla durata di tali evoluzioni locali adiacenti della traiettoria. La direzione di queste strutture può differire. Quelle prima definite sono strutture diagonali parallele alla LOI (a 45°) che corrispondono a traiettorie che sono parallele con la stessa evoluzione temporale.

  Esistono anche strutture parallele alla diagonale a 135° :

  $\boldsymbol{R_{i+k,j-k} = 1}$ per k=1,…,l dove l è la lunghezza delle linee diagonali.

  rappresentanti invece di segmenti di traiettoria con evoluzione temporale contraria.

- ***Linee verticali (orizzontali)*** definite:

  $\boldsymbol{R_{i,j+k} = 1}$ per k=1,…,h dove h è la lunghezza delle linee verticali

  denotano periodi temporali dove lo stato del sistema non cambia o cambia molto lentamente.

Le strutture di piccola scala sono alla base dell'analisi quantitativa delle RPs.
L'interpretazione qualitativa delle RPs basata su una visione d'insieme necessita una certa esperienza e può essere utile a fornire informazioni di massima sul sistema. In ogni caso una analisi quantitativa offre un'interpretazione più oggettiva del sistema. Questo tipo di analisi ha fatto sì che le RPs acquisissero maggiore popolarità ed un certo accredito nella comunità scientifica. Andando a riassumere le considerazioni possibili in una analisi qualitativa nella interpretazione delle RPs:

(if, in addition, these diagonal lines are periodic, the considered system contains unstable periodic orbits)

7. Diagonal lines (orthogonal to the LOI) → the evolution of states is similar at different times but with inverse time; sometimes this is a sign for an insufficient embedding
8. Vertical and horizontal lines/ clusters → some states do not change or change slowly for some time (laminar states)

The visual interpretation of RPs requires some experience. The study of RPs from paradigmatic systems gives a good introduction into characteristic typology and texture. However, their quantification offers a more objective way for the investigation of the considered system. With this quantification, the RPs have become more and more popular within a growing group of scientists who use RPs and their quantification techniques for data analysis (a search with the Scirus search engine reveals over 200 journal published works and approximately 700 web published works about RPs).

2.2.3 The Quantitative Analysis of Recurrence Plots

Zbilut and Webber have developed a tool which quantifies the mentioned structures in the RPs, the *recurrence quantification analysis (RQA)* (Zbilut and Webber Jr., 1992; Webber Jr. and Zbilut, 1994). They define measures of complexity using the recurrence point density and diagonal structures in the recurrence plot: the *recurrence rate* (or per cent recurrences), the *determinism* (or per cent determinism), the *divergence* (the inverse of the maximal length of diagonal structures), the *entropy* and the *trend* (or drift). A computation of these measures in small windows (sub-matrices) of the RP moving along the LOI yields the time dependent behaviour of these variables. Some studies based on these RQA measures show that these measures are able to find bifurcation points, especially chaos-order transitions (Trulla et al., 1996). The RQA is based on RPs gained by using a fixed threshold ε, hence the RPs are symmetric. In the following, these RQA measures are introduced. In the Subsec. 2.2.4 we will adopt this concept in order to quantify the vertical structures in the RP.

The first measure of the RQA is the *recurrence rate* or *per cent recurrences (REC)*

$$RR = \frac{1}{N^2} \sum_{i,j=1}^{N} \mathbf{R}_{i,j}^{m,\varepsilon}, \qquad (2.16)$$

which simply counts the black dots in the RP. It is a measure of the density of recurrence points and corresponds to the definition of the correlation sum (2.7)

Plagiat

Omogeneità	*il processo è stazionario*
Fading dall'angolo in alto a sinistra all'angolo in basso a destra	*non stazionarietà, il sistema contiene drift o trend*
Bande bianche	*non stazionarietà; ci sono stati del sistema rari o lontano da quella che è l'evoluzione media del sistema; possono esserci transizioni*
Patterns periodici	*ci sono evoluzioni del sistema cicliche; la distanza dei vari patterns periodici corrispondono al periodo.*
Punti isolati singoli	*sono presenti fluttuazioni rapide del sistema; se ci sono solo punti isolati il sistema può essere interpretato come random*
Linee diagonali (parallele alla LOI)	*l'evoluzione del sistema è simile in tempi diversi; il sistema può essere deterministico; se le linee diagonali pervengono a lato di singoli punti isolati, il sistema può avere origine caotica; se poi le linee diagonali sono anche periodiche allora il sistema in questione contiene orbite periodiche instabili.*
Linee diagonali (perpendicolari alla LOI)	*l'evoluzione del sistema è simile in tempi diversi con scala dei tempi inversi; può essere segno di una dimensione di embedding sottostimata.*
Linee verticali e orizzonatali	*ci sono stati del sistema che non variano o lo fanno molto lentamente*

Tabella 3-1 Si riassume le informazioni che è possibile estrarre da una analisi qualitativa delle RPs

Original

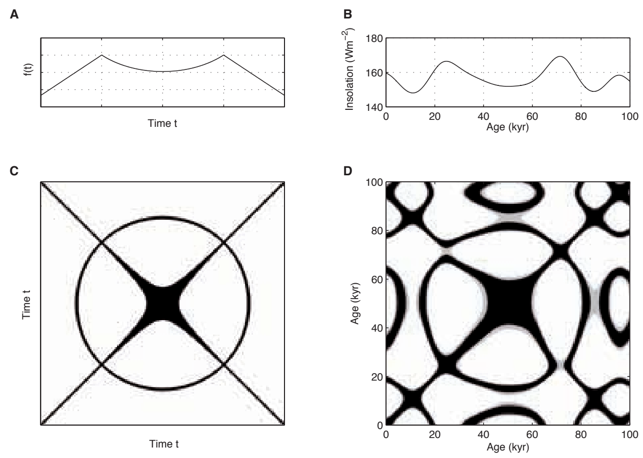

Figure 2.7: (A, C) Illustrative example of the relationship between the slope of lines in an RP and the local derivatives of the involved trajectory segments. Since the local derivative of the transformation the time scales of the linear and the hyperbolic sections corresponds to the derivative of a circle line, a circle occurs in the RP. (B, D) A corresponding structure found in nature: the solar insolation on the latitude 44°N for the last 100 kyr (data from Berger and Loutre, 1991). RPs created without embedding.

1. Homogeneity → the process is obviously stationary
2. Fading to the upper left and lower right corners → nonstationarity; the process contains a trend or drift
3. Disruptions (white bands) occur → nonstationarity; some states are rare or far from the normal; transitions may have occurred
4. Periodic patterns → cyclicities in the process; the time distance between periodic patterns (e. g. lines) corresponds to the period
5. Single isolated points → heavy fluctuation in the process; if only single isolated points occur, the process may be a random process
6. Diagonal lines (parallel to the LOI) → the evolution of states is similar at different times; the process could be deterministic; if these diagonal lines occur beside single isolated points, the process can be from chaos

Plagiat

Omogeneità	*il processo è stazionario*
Fading dall'angolo in alto a sinistra all'angolo in basso a destra	*non stazionarietà, il sistema contiene drift o trend*
Bande bianche	*non stazionarietà; ci sono stati del sistema rari o lontano da quella che è l'evoluzione media del sistema; possono esserci transizioni*
Patterns periodici	*ci sono evoluzioni del sistema cicliche; la distanza dei vari patterns periodici corrispondono al periodo.*
Punti isolati singoli	*sono presenti fluttuazioni rapide del sistema; se ci sono solo punti isolati il sistema può essere interpretato come random*
Linee diagonali (parallele alla LOI)	*l'evoluzione del sistema è simile in tempi diversi; il sistema può essere deterministico; se le linee diagonali pervengono a lato di singoli punti isolati, il sistema può avere origine caotica; se poi le linee diagonali sono anche periodiche allora il sistema in questione contiene orbite periodiche instabili.*
Linee diagonali (perpendicolari alla LOI)	*l'evoluzione del sistema è simile in tempi diversi con scala dei tempi inversi; può essere segno di una dimensione di embedding sottostimata.*
Linee verticali e orizzonatali	*ci sono stati del sistema che non variano o lo fanno molto lentamente*

Tabella 3-1 Si riassume le informazioni che è possibile estrarre da una analisi qualitativa delle RPs

Original

(if, in addition, these diagonal lines are periodic, the considered system contains unstable periodic orbits)

7. Diagonal lines (orthogonal to the LOI) → the evolution of states is similar at different times but with inverse time; sometimes this is a sign for an insufficient embedding
8. Vertical and horizontal lines/ clusters → some states do not change or change slowly for some time (laminar states)

The visual interpretation of RPs requires some experience. The study of RPs from paradigmatic systems gives a good introduction into characteristic typology and texture. However, their quantification offers a more objective way for the investigation of the considered system. With this quantification, the RPs have become more and more popular within a growing group of scientists who use RPs and their quantification techniques for data analysis (a search with the Scirus search engine reveals over 200 journal published works and approximately 700 web published works about RPs).

2.2.3 The Quantitative Analysis of Recurrence Plots

Zbilut and Webber have developed a tool which quantifies the mentioned structures in the RPs, the *recurrence quantification analysis (RQA)* (Zbilut and Webber Jr., 1992; Webber Jr. and Zbilut, 1994). They define measures of complexity using the recurrence point density and diagonal structures in the recurrence plot: the *recurrence rate* (or per cent recurrences), the *determinism* (or per cent determinism), the *divergence* (the inverse of the maximal length of diagonal structures), the *entropy* and the *trend* (or drift). A computation of these measures in small windows (sub-matrices) of the RP moving along the LOI yields the time dependent behaviour of these variables. Some studies based on these RQA measures show that these measures are able to find bifurcation points, especially chaos-order transitions (Trulla et al., 1996). The RQA is based on RPs gained by using a fixed threshold ε, hence the RPs are symmetric. In the following, these RQA measures are introduced. In the Subsec. 2.2.4 we will adopt this concept in order to quantify the vertical structures in the RP.

The first measure of the RQA is the *recurrence rate* or *per cent recurrences (REC)*

$$RR = \frac{1}{N^2} \sum_{i,j=1}^{N} \mathbf{R}_{i,j}^{m,\varepsilon}, \qquad (2.16)$$

which simply counts the black dots in the RP. It is a measure of the density of recurrence points and corresponds to the definition of the correlation sum (2.7)

Plagiat

4 Analisi Quantitativa delle Recurrence Plots: RQA[2]

4.1 Definizioni delle statistiche

Per una valutazione più quantitativa di questi grafici, si sono scelti (seguendo l'approccio di Trulla, Giuliani et al. [15]) alcuni parametri da valutare su epoche consecutive della serie. In particolare sono state considerate sei statistiche *non lineari,* ovvero (1) %ricorrenza REC, (2) %determinismo DET, (3) entropia ENTR, (4) line max e DIV (5) Trend, (6) RATIO. Esistono studi che dimostrano il possibile utilizzo di alcuni parametri per individuare i punti di biforcazione soprattutto nelle transizioni fra caos e determinismo [15]. La RQA si basa sulla costruzione delle RPs con soglia fissa, cioè su RPs simmetriche. Andiamo a definire uno per uno i vari indici non lineari :

- **Percentuale delle Ricorrenze REC:**

$$REC = \frac{1}{N^2}\sum_{i,j=1}^{N} R_{i,j}^{m,\varepsilon}$$

Eq. 4.1

Rappresenta la percentuale di punti ricorrenti evidenziati sulla matrice e che hanno una distanza inferiore alla soglia prescelta;

- **Percentuale di Determinismo DET:**

Definiamo la frequenza di distribuzione delle strutture diagonali di lunghezza l nelle RPs $P^{\varepsilon}(l) = \{l_i; i = 1...N_l\}$, dove N_l è il numero di linee diagonali. Sistemi con comportamento stocastico non presentano linee diagonali, mentre sistemi deterministici presentano molte linee diagonali e pochi punti isolati singoli.

Si definisce l'indice come il rapporto fra il numero di punti che formano strutture diagonali e il numero totale di punti:

[2] RQA: Recurrence Quantification Analysis

Original

except that the LOI is included. Nevertheless, the constraint for the correlation sum, that a large amount of data points are needed, also applies to the RR when used as an estimation of the correlation sum. In the limit

$$P_{\bullet} = \lim_{N\to\infty} \frac{1}{N^2}\sum_{i,j=1}^{N} \mathbf{R}_{i,j}^{m,\varepsilon}, \tag{2.17}$$

this measure becomes the probability of finding a recurrence point in the RP (probability that a state will recur). With the knowledge of the probability $\rho(x)$ of the states where dimension $m = 1$ (or the maximum norm) the recurrence rate can be analytically computed by using the convolution (Thiel et al., 2003a)

$$P_{\bullet}^{\varepsilon} = \int_{-\varepsilon}^{\varepsilon} \rho(x) * \rho(x)\, dx. \tag{2.18}$$

This probability $P_{\bullet}^{\varepsilon}$ can be used to analytically describe the RQA measures for some systems (Thiel et al., 2002, 2003a).

The next measures consider the diagonal lines. The frequency distribution of the lengths l of the diagonal structures in the RP is $P^{\varepsilon}(l) = \{l_i;\ i = 1 \ldots N_l\}$, where N_l is the absolute number of diagonal lines (each line is counted only once in contrast to the cumulative distribution[1]). Processes with stochastic behaviour cause none or very short diagonals, whereas deterministic processes cause longer diagonals and less single, isolated recurrence points. Therefore, the ratio of recurrence points that form diagonal structures to all recurrence points

$$DET = \frac{\sum_{l=l_{min}}^{N} l\, P^{\varepsilon}(l)}{\sum_{i,j}^{N} \mathbf{R}_{i,j}^{m,\varepsilon}}, \tag{2.19}$$

is introduced as a measure for the *determinism* (or predictability) in the system. However, this measure does not have the real meaning of the determinism of a process. The threshold l_{min} excludes the diagonal lines which are formed by the tangential motion of the phase space trajectory. For $l_{min} = 1$ the determinism is equal to the recurrence rate. The choice of l_{min} could be made in a similar way as the choice of the size for the Theiler window (cf. remark on p. 13), but one has to take into account that a too large l_{min} can worsen the histogram $P(l)$ and thus the reliability of the measure DET.

[1] The cumulative distribution for the line length

$$P_c^{\varepsilon}(l) = \sum_{i=1}^{N_l} (i - l + 1)\, P^{\varepsilon}(i)$$

counts each diagonal line several times, in the sense that a line of length l contains l lines of length one, $(l-1)$ lines of length two, $(l-2)$ lines of length three ... one line of length l.

$$DET = \frac{\sum_{l=l_{min}}^{N} l * P^{\varepsilon}(l)}{\sum_{i,j}^{N} R_{i,j}^{m,\varepsilon}}$$

Eq. 4.2

Questo indice individua in qualche modo il *determinismo* presente nel sistema anche se non lo misura in senso stretto. La soglia che si va a scegliere l_{min} sarà tale da escludere le linee diagonali che vengono a formarsi a causa dei ***movimenti tangenziali***. Per $l_{min} = 1$ DET è uguale a 1. Nella scelta di l_{min} dobbiamo tener conto che valori troppo alti possono degradare l'istogramma $P^{\varepsilon}(l)$.

- **Entropia ENTR:**

  Questa statistica misura l'entropia di Shannon della distribuzione dei segmenti di linee paralleli alla diagonale. In genere è alta nelle finestre periodiche (grande diversità di linee diagonali) e bassa in quelle caotiche (bassa diversità di linee diagonali).

  La sua definizione è:

$$ENTR = -\sum_{l=l_{min}}^{N} p(l) ln[p(l)] \quad \text{con} \quad p(l) = \frac{P^{\varepsilon}(l)}{\sum_{l=l_{min}}^{N} P^{\varepsilon}(l)}$$

Eq. 4.3

L'ENTR dà una misura della complessità delle strutture deterministiche del sistema.

- **MaxLine e DIV**

except that the LOI is included. Nevertheless, the constraint for the correlation sum, that a large amount of data points are needed, also applies to the RR when used as an estimation of the correlation sum. In the limit

$$P_{\bullet} = \lim_{N\to\infty} \frac{1}{N^2} \sum_{i,j=1}^{N} \mathbf{R}_{i,j}^{m,\varepsilon}, \qquad (2.17)$$

this measure becomes the probability of finding a recurrence point in the RP (probability that a state will recur). With the knowledge of the probability $\rho(x)$ of the states where dimension $m = 1$ (or the maximum norm) the recurrence rate can be analytically computed by using the convolution (Thiel et al., 2003a)

$$P_{\bullet}^{\varepsilon} = \int_{-\varepsilon}^{\varepsilon} \rho(x) * \rho(x)\, dx. \qquad (2.18)$$

This probability $P_{\bullet}^{\varepsilon}$ can be used to analytically describe the RQA measures for some systems (Thiel et al., 2002, 2003a).

The next measures consider the diagonal lines. The frequency distribution of the lengths l of the diagonal structures in the RP is $P^{\varepsilon}(l) = \{l_i;\ i = 1 \ldots N_l\}$, where N_l is the absolute number of diagonal lines (each line is counted only once in contrast to the cumulative distribution[1]). Processes with stochastic behaviour cause none or very short diagonals, whereas deterministic processes cause longer diagonals and less single, isolated recurrence points. Therefore, the ratio of recurrence points that form diagonal structures to all recurrence points

$$DET = \frac{\sum_{l=l_{min}}^{N} l\, P^{\varepsilon}(l)}{\sum_{i,j}^{N} \mathbf{R}_{i,j}^{m,\varepsilon}}, \qquad (2.19)$$

is introduced as a measure for the *determinism* (or predictability) in the system. However, this measure does not have the real meaning of the determinism of a process. The threshold l_{min} excludes the diagonal lines which are formed by the tangential motion of the phase space trajectory. For $l_{min} = 1$ the determinism is equal to the recurrence rate. The choice of l_{min} could be made in a similar way as the choice of the size for the Theiler window (cf. remark on p. 13), but one has to take into account that a too large l_{min} can worsen the histogram $P(l)$ and thus the reliability of the measure DET.

[1]The cumulative distribution for the line length

$$P_c^{\varepsilon}(l) = \sum_{i=1}^{N_l} (i - l + 1)\, P^{\varepsilon}(i)$$

counts each diagonal line several times, in the sense that a line of length l contains l lines of length one, $(l-1)$ lines of length two, $(l-2)$ lines of length three … one line of length l.

Plagiat

$$DET = \frac{\sum_{l=l_{min}}^{N} l * P^{\varepsilon}(l)}{\sum_{i,j}^{N} R_{i,j}^{m,\varepsilon}}$$

Eq. 4.2

Questo indice individua in qualche modo il *determinismo* presente nel sistema anche se non lo misura in senso stretto. La soglia che si va a scegliere $\boldsymbol{l_{min}}$ sarà tale da escludere le linee diagonali che vengono a formarsi a causa dei ***movimenti tangenziali***. Per $\boldsymbol{l_{min} = 1}$ DET è uguale a 1. Nella scelta di $\boldsymbol{l_{min}}$ dobbiamo tener conto che valori troppo alti possono degradare l'istogramma $\boldsymbol{P^{\varepsilon}(l)}$.

- **Entropia ENTR:**

  Questa statistica misura l'entropia di Shannon della distribuzione dei segmenti di linee paralleli alla diagonale. In genere è alta nelle finestre periodiche (grande diversità di linee diagonali) e bassa in quelle caotiche (bassa diversità di linee diagonali).

  La sua definizione è:

$$ENTR = -\sum_{l=l_{min}}^{N} p(l)\, ln[p(l)] \quad \text{con} \quad p(l) = \frac{P^{\varepsilon}(l)}{\sum_{l=l_{min}}^{N} P^{\varepsilon}(l)}$$

Eq. 4.3

L'ENTR dà una misura della complessità delle strutture deterministiche del sistema.

- **MaxLine e DIV**

Original

Diagonal structures show the range in which a segment of the trajectory is rather close to another segment of the trajectory at a different time; thus these lines give a hint about the divergence of the trajectory segments. The *average diagonal line length*

$$L = \frac{\sum_{l=l_{min}}^{N} l\, P^{\varepsilon}(l)}{\sum_{l=l_{min}}^{N} P^{\varepsilon}(l)} \tag{2.20}$$

is the average time that two segments of the trajectory are close to each other, and can be interpreted as the mean prediction time. Instead of this average the RQA uses the *maximum length* of the diagonal structures or its inverse, the *divergence*,

$$L_{max} = \max(\{l_i;\ i = 1 \dots N_l\}) \quad \text{respective} \quad DIV = \frac{1}{L_{max}}. \tag{2.21}$$

Eckmann has stated that "the length of the diagonal lines is related to the largest positive Lyapunov exponent" if there is one in the considered system (Eckmann et al., 1987). Different approaches have been suggested in order to use these lengths for the estimation of the largest positive Lyapunov exponent as DIV (Trulla et al., 1996) or the average of the inverse of the half lengths of the diagonals (Choi et al., 1999, they have defined this measure for perpendicular RPs).

The measure *entropy* refers to the Shannon entropy of the frequency distribution of the diagonal line lengths

$$ENTR = -\sum_{l=l_{min}}^{N} p(l) \ln p(l) \quad \text{with} \quad p(l) = \frac{P^{\varepsilon}(l)}{\sum_{l=l_{min}}^{N} P^{\varepsilon}(l)} \tag{2.22}$$

and reflects the complexity of the deterministic structure in the system. However, this entropy depends sensitively on the bin number and, thus, may differ for different realizations of the same process, as well as for different data preparations.

The measures introduced up to now, RR, DET, L etc. can also be computed separately for each diagonal parallel to the LOI. The representation of these diagonalwise computed measures, $RR_*(t)$, $DET_*(t)$ and $L_*(t)$, over the time distance t from the LOI hints at the determination of the already mentioned Theiler window (cf. Subsec. 2.2.1). Henceforth, measures with a subscribed asterisk or index denote diagonalwise computed measures. This diagonalwise determination of the RQA measures will receive more importance in the analysis of CRPs (the definition of the diagonalwise computed measures can be found in Subsec. 2.3.1, which can be adopted for the RPs). Furthermore, the measure RR_* is closely related to the *close returns histogram* introduced

Michele Bacchereti: Sviluppo di modelli non lineari per l'analisi di segnali Elettroencefalografici

Plagiat

Lunghezza massima dei segmenti diagonali. Il suo inverso è direttamente correlato con l'esponente di Lyapunov più grande positivo.

$L_{max} = max(\{l_i; i = 1...N_l\})$ e rispettivamente $DIV = \frac{1}{L_{max}}$

Eq. 4.4

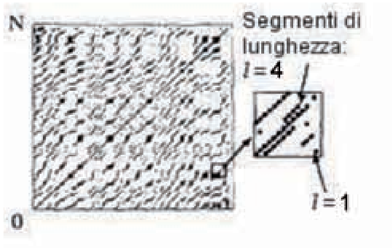

Figura 4-1 Esempio di RP dove si evidenzia la lunghezza delle strutture in questione

Eckmann afferma che la lunghezza delle linee diagonali è in relazione con il più grande esponente di Lyapunov positivo se c'è n'è uno nel sistema [8].

- **Trend**

  Misura la dispersione dei punti ricorrenti allontanandosi dalla diagonale principale.

$$TREND = \frac{\sum_{i=1}^{\overline{N}} (i - \frac{\overline{N}}{2})(REC_i - \langle REC_i \rangle)}{\sum_{i=1}^{\overline{N}} (i - \frac{\overline{N}}{2})}$$

Eq. 4.5

Fornisce informazione sulla non stazionarietà e sui drift presenti nel sistema. Nel calcolo si esclude il bordo della RP, infatti $\overline{N} < N$. La scelta di $\overline{N}$ dipende dal tipo di sistema studiato: $N - \overline{N} > 10$ può essere

Norbert Marwan: Encounters with neighbours - Current developments of concepts based on recurrence Plots and their applications

Original

by Lathrop and Kostelich (1989). This can be used to find periodic orbits in low-dimensional chaotic systems (Lathrop and Kostelich, 1989; Mindlin and Gilmore, 1992; Gilmore, 1998). Since periodic orbits are more closely related to the occurrence of longer diagonal structures, the measures DET_* and L_* are more suitable candidates for this kind of study. The measure RR_* have been already used by Eckmann et al. (1987) for the study of nonstationarity in the data.

The last RQA measure is the *trend*, which is a linear regression coefficient over the recurrence point density RR_* of the diagonals parallel to the LOI (for definition see Eq. (2.34) in Subsec. 2.3.1) as a function of the time distance between these diagonals and the LOI

$$TREND = \frac{\sum_{i=1}^{\tilde{N}} (i - \tilde{N}/2)(RR_i - \langle RR_i \rangle)}{\sum_{i=1}^{\tilde{N}} (i - \tilde{N}/2)^2}. \quad (2.23)$$

The trend gives information about a nonstationarity in the process, especially a drift. The computation excludes the edges of the RP ($\tilde{N} < N$) because the statistic lacks by reason of less recurrence points. The choice of $\tilde{N}$ depends on the studied process. Whereas $N - \tilde{N} > 10$ should be sufficient for noise, this difference should be much larger for a process with some autocorrelation (ten times the order of magnitude of the autocorrelation time should always be enough). It should be noted that if the time dependent RQA (measures are computed in shifted windows) is used, $TREND$ will depend strongly on the size of the windows and may reveal contrary results for different window sizes.

In some publications a further measure, the *ratio*, can be found. It is defined as the ratio between DET and RR (Webber Jr. and Zbilut, 1994) and can be computed from the frequency distributions of the lengths of the diagonal lines

$$RATIO = N^2 \frac{\sum_{l=l_{min}}^{N} l\, P^{\varepsilon}(l)}{\left(\sum_{l=1}^{N} l\, P^{\varepsilon}(l)\right)^2}. \quad (2.24)$$

A heuristic study of physiological systems has revealed that this ratio can be used in order to discover transitions, because it was found that during certain types of transitions the RR can decrease, whereas DET does not change at the same time (Webber Jr. and Zbilut, 1994).

Currently, a satisfying theory about the statistics of these measures of complexity has not been developed. Therefore, a reliable statement about the significance of these measures cannot be made. Nevertheless, a possibility for assessing the significance of these measures lies in applying a sufficient surrogate test (but this works only for stationary processes).

sufficiente per eliminare gli effetti del rumore. Da notare che questo indice dipende molto dalla grandezza delle finestre su cui si vanno a calcolare i vari indici.

- **Ratio**

E' definito come il rapporto fra DET e RR, può essere calcolato la distribuzione in frequenza della lunghezza delle linee diagonali:

$$RATIO = N^2 \frac{\sum_{l=l_{min}}^{N} l * P^{\varepsilon}(l)}{(\sum_{l=1}^{N} l * P^{\varepsilon}(l))^2}$$

Eq. 4.6

studi euristici di sistemi fisiologici hanno rivelato che questo indice può essere utile per evidenziare transizioni, visto che in alcune transizioni del sistema RR decresce mentre DET rimane costante [16].

4.2 Ulteriore misura della complessità del sistema: "laminarity" e "trapping time"

Consideriamo gli stati $y(i)$ della traiettoria e definiamo l'insieme dei *punti ricorrenti* $\wp_i$:

$$\wp_i = \{y(j) : R_{i,j} = 1; j \in [1 \ldots N]\}$$

Eq. 4.7

Definiamo il sottoinsieme dei *punti ricorrenti:*

by Lathrop and Kostelich (1989). This can be used to find periodic orbits in low-dimensional chaotic systems (Lathrop and Kostelich, 1989; Mindlin and Gilmore, 1992; Gilmore, 1998). Since periodic orbits are more closely related to the occurrence of longer diagonal structures, the measures DET_* and L_* are more suitable candidates for this kind of study. The measure RR_* have been already used by Eckmann et al. (1987) for the study of nonstationarity in the data.

The last RQA measure is the *trend*, which is a linear regression coefficient over the recurrence point density RR_* of the diagonals parallel to the LOI (for definition see Eq. (2.34) in Subsec. 2.3.1) as a function of the time distance between these diagonals and the LOI

$$TREND = \frac{\sum_{i=1}^{\tilde{N}}(i - \tilde{N}/2)(RR_i - \langle RR_i \rangle)}{\sum_{i=1}^{\tilde{N}}(i - \tilde{N}/2)^2}. \quad (2.23)$$

The trend gives information about a nonstationarity in the process, especially a drift. The computation excludes the edges of the RP ($\tilde{N} < N$) because the statistic lacks by reason of less recurrence points. The choice of $\tilde{N}$ depends on the studied process. Whereas $N - \tilde{N} > 10$ should be sufficient for noise, this difference should be much larger for a process with some autocorrelation (ten times the order of magnitude of the autocorrelation time should always be enough). It should be noted that if the time dependent RQA (measures are computed in shifted windows) is used, $TREND$ will depend strongly on the size of the windows and may reveal contrary results for different window sizes.

In some publications a further measure, the *ratio*, can be found. It is defined as the ratio between DET and RR (Webber Jr. and Zbilut, 1994) and can be computed from the frequency distributions of the lengths of the diagonal lines

$$RATIO = N^2 \frac{\sum_{l=l_{min}}^{N} l\, P^{\varepsilon}(l)}{\left(\sum_{l=1}^{N} l\, P^{\varepsilon}(l)\right)^2}. \quad (2.24)$$

A heuristic study of physiological systems has revealed that this ratio can be used in order to discover transitions, because it was found that during certain types of transitions the RR can decrease, whereas DET does not change at the same time (Webber Jr. and Zbilut, 1994).

Currently, a satisfying theory about the statistics of these measures of complexity has not been developed. Therefore, a reliable statement about the significance of these measures cannot be made. Nevertheless, a possibility for assessing the significance of these measures lies in applying a sufficient surrogate test (but this works only for stationary processes).

Plagiat

sufficiente per eliminare gli effetti del rumore. Da notare che questo indice dipende molto dalla grandezza delle finestre su cui si vanno a calcolare i vari indici.

- **Ratio**

E' definito come il rapporto fra DET e RR, può essere calcolato la distribuzione in frequenza della lunghezza delle linee diagonali:

$$RATIO = N^2 \frac{\sum_{l=l_{min}}^{N} l * P^{\varepsilon}(l)}{(\sum_{l=1}^{N} l * P^{\varepsilon}(l))^2}$$

Eq. 4.6

studi euristici di sistemi fisiologici hanno rivelato che questo indice può essere utile per evidenziare transizioni, visto che in alcune transizioni del sistema RR decresce mentre DET rimane costante [16].

4.2 Ulteriore misura della complessità del sistema: "laminarity" e "trapping time"

Consideriamo gli stati $y(i)$ della traiettoria e definiamo l'insieme dei *punti ricorrenti* $\wp_i$:

$$\wp_i = \{y(j) : R_{i,j} = 1; j \in [1...N]\}$$

Eq. 4.7

Definiamo il sottoinsieme dei *punti ricorrenti:*

continued on following page

Original

In a more theoretical study, Thiel et al. (2003b) have revealed analytical solutions for the RQA measures of stochastic systems and maps. Gao and Cai (2000) have studied the relationship between the RQA measures and a divergence exponent which is closely related to the largest Lyapunov exponent. Furthermore, the clear relationship between the cumulative distribution (cf. footnote on p. 23) $P_c(l)$ and the second order Rényi entropy K_2 has been found (Faure and Korn, 1998; Thiel et al., 2003a). Referring to their studies Thiel et al. (2003a), have stated that the distribution $P_c(l)$ is related rather to K_2 than to the largest positive Lyapunov exponent.

An appropriate embedding of time series is motivated by the desire to avoid false nearest neighbours. However, in an RP false nearest neighbours will occur as black dots, rather short black lines or (for a specific embedding) as black lines perpendicular to the LOI (i. e. with an angle of $-\pi/4$). Whereas the estimation of some invariants of the RP (like K_2) are independent from the embedding (and consequently does not need any embedding), the estimation of the measures RR, DET, L etc. depends on the embedding and needs a sufficient choice (Thiel et al., 2003a).

All these RQA measures are based largely on the distribution of the length of the diagonal structures in the RP. Additional information about further geometrical structures as vertical and horizontal elements is not included. In the following, I will extend this quantitative view to vertical structures and propose new measures of complexity based on the distribution of the vertical line length. Since we are using symmetric RPs, we will only consider the vertical structures in the following.

2.2.4 New Measures of Complexity: Laminarity and Trapping Time

Let us consider a point $\vec{x}_i$ of the trajectory and the set of its associated recurrence points $\mathcal{R}_i$

$$\mathcal{R}_i = \{\vec{x}_j : \mathbf{R}_{i,j} = 1;\ j \in [1 \ldots N]\}. \qquad (2.25)$$

Let us denote subsets of these recurrence points

$$\mathcal{V}_{i,j} = \left\{j + 1 \ldots j + v_j : \vec{x}_j \notin \mathcal{R}_i;\ \vec{x}_{j+1} \ldots \vec{x}_{j+v_j} \forall \mathcal{R}_i;\ \vec{x}_{j+v_j+1} \notin \mathcal{R}_i\right\} \qquad (2.26)$$

which form connected vertical structures of the length v_j. In continuous time systems with high time resolution and with an adequately large threshold ε a large part of these subsets $\mathcal{V}_{i,j}$ usually corresponds to the tangential motion of the phase space trajectory (cp. Subsec. 2.2.1 on p. 13), i. e. to the sojourn points described by Gao (1999). However, not all elements of these sets are real sojourn points. Although sojourn points do not occur in maps, the subsets $\mathcal{V}_{i,j}$

Plagiat

continued from previous page

$$\chi_{i,j} = \{ j+1 \ldots j+k_j : y(j) \notin \wp_i ; y(j+1) \ldots y(j+k_j) \in \wp_i ; y(j+k_j+1) \notin \wp_i \}$$

Eq. 4.8

questo definisce strutture verticali di lunghezza k_j. In sistemi continui con risoluzione temporale alta e con una soglia ε adeguatamente alta una larga parte di questi sottoinsiemi corrisponde ai ***movimenti tangenziali*** di cui si è parlato in precedenza.

Definiamo $P_i(k) = \{ k_j ; j = 1 \ldots N_k \}$, con N_k il numero di linee verticali; $\bigcup_{i=1}^{N} P_i(k)$ è la distribuzione delle linee verticale nell'intera RP che chiamiamo $P^{\varepsilon}(k)$.

Analogamente alla definizione dell'indice DET, si definisce il rapporto fra i *punti ricorrenti* che formano linee verticali e i *punti ricorrenti* totali:

$$LAM = \frac{\sum_{k=k_{min}}^{N} k * P^{\varepsilon}(k)}{\sum_{k=1}^{N} k * P^{\varepsilon}(k)}$$

Eq. 4.9

denominata ***laminarità***. Il calcolo di LAM è realizzato per quei ***k*** che superano la lunghezza minima k_{min} in modo da minimizzare l'influenza dei ***movimenti tangenziali***. Questo indice dà una misura della presenza di stati laminari nel sistema. LAM diminuisce con il crescere dei punti singoli isolati nelle RPs.

Definiamo invece la lunghezza media delle linee verticali:

$$TT = \frac{\sum_{k=k_{min}}^{N} k * P^{\varepsilon}(k)}{\sum_{k=k_{min}}^{N} P^{\varepsilon}(k)}$$

Eq. 4.10

denominata ***trapping time***. TT misura il tempo medio in cui il sistema attende in uno specifico stato.

Original

are not necessarily empty because of laminar states. Furthermore, the finite size of the threshold ε can pretend a tangential motion, although there are rather small cycles instead of a tangential motion (e. g. Shilnikov chaos).

Next, we determine the length $v_j = |\mathcal{V}_{i,j}|$ of all subsets $\mathcal{V}_{i,j}$. $\mathcal{P}_i(v) = \{v_j ; j = 1 \ldots N_v\}$ denotes the set of all occurring subset lengths in $\mathcal{V}_i$ (N_v is the absolute number of the vertical lines), and from $\bigcup_{i=1}^{N} \mathcal{P}_i(v)$ we determine the distribution of the vertical line lengths $P^{\varepsilon}(v)$ in the entire RP.

Analogous to the definition of the determinism (2.19), we compute the ratio between the recurrence points forming the vertical structures and the entire set of recurrence points

$$LAM = \frac{\sum_{v=v_{min}}^{N} v P^{\varepsilon}(v)}{\sum_{v=1}^{N} v P^{\varepsilon}(v)}, \qquad (2.27)$$

and call it *laminarity* LAM. The computation of LAM is realized for those v that exceed a minimal length v_{min} in order to decrease the influence of sojourn points. For maps we use $v_{min} = 2$. LAM is the measure of the amount of vertical structures in the whole RP and represents the occurrence of laminar states in the system without, however, describing the length of these laminar phases. LAM will decrease if the RP consists of more single recurrence points than vertical structures.

We define the average length of vertical structures (cp. (2.20))

$$TT = \frac{\sum_{v=v_{min}}^{N} v P^{\varepsilon}(v)}{\sum_{v=v_{min}}^{N} P^{\varepsilon}(v)}, \qquad (2.28)$$

which we call *trapping time* TT. The computation also uses the minimal length v_{min} as in LAM (2.27). The measure TT contains information about the amount and the length of the vertical structures in the RP. It measures the mean time that the system will abide at a specific state (how long the state will be trapped).

Finally, we use the *maximal length of the vertical structures* in the RP

$$V_{max} = \max(\{v_l ; l = 1 \ldots L\}) \qquad (2.29)$$

as a measure which is the analogue to the standard measure L_{max} (2.21).

In contrast to the known RQA measures, these new measures are able to find chaos-chaos transitions (Marwan et al., 2002b). Hence, these measures make the investigation of intermittency possible, even if they are only represented by rather short and nonstationary data series. Since for periodic dynamics these measures are zero, chaos-order transitions can also be identified.

An application to the logistic map $x_{n+1} = a\, x_n (1 - x_n)$ illustrates the potentials of LAM, TT and V_{max}. We generate for each control parameter $a \in [3.5, 4]$, $\Delta a = 0.0005$ a separate time series (Fig. 2.8). In the analyzed range