Curved structures in recurrence plots: The role of the sampling time

A. Facchini[1,2,*] and H. Kantz[2,†]

[1]*Center for the Study of Complex Systems, University of Siena, Via Tommaso Pendola 37, I-53100 Siena, Italy*

[2]*Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Strasse 38, D-01187 Dresden, Germany*

(Received 16 November 2006; published 26 March 2007)

We give a theoretical explanation of the formation of the curved macropatterns in the recurrence plots of sinusoidal signals, with nonstationarity in the phase or in the frequency. We show that the large time scales observed and the curved structures are the artificial product of the discretization of the signal. Recurrence plots are highly sensitive to the phase error introduced by the sampling, and we show that this characteristic can be used to detect very small ($\sim 0.5\%$) phase or frequency shifts of the carrier frequency.

DOI: 10.1103/PhysRevE.75.036215 PACS number(s): 05.45.Tp, 43.80.+p

I. INTRODUCTION

The recurrence plot (RP) is a visual tool for the investigation of temporal recurrences in phase space, and was initially designed to display recurring patterns and nonstationarity in time series [1]. Recurrence is the most important characteristic of chaotic systems, while nonstationarity may arise from other reasons such as parameter drifting, time varying driving forces, sudden changes in dynamics, etc. In the recent years, RPs found a wide range of applications in the time series analysis of nonstationary phenomena. For instance, RPs have been used in the analysis of biological systems including neuronal spike trains [2], electromyographic data [3], intercranial EEG recordings [4], electrocadiograms recording [5], protein folding [6], DNA sequences [7], and nonlinear phenomena in voice production [8]. The popularity of RPs lies in the fact that the structures are visually appealing, allowing the investigation of high dimensional dynamics by means of a simple two-dimensional plot.

For a better understanding and quantification of the recurrences, Webber and Zbilut have proposed a set of quantification measures, which are mainly based on the statistical distribution of the line structures among the RP. By means of the recurrence quantification analysis [9], the RP can be used as a tool for the exploration of bifurcation phenomena and dynamics changes in even nonstationary and short time series. Some authors have also related the quantification measures to the invariants of the phase space, like the correlation dimension and the largest Lyapunov exponent [10,11].

Despite this, there is a lack in the literature regarding the origin and the formation of the various patterns that can be observed in a RP. In a recent work, Gao [12] identified the main patterns visible in a RP, while Marwan focused his attention on line structures [12]. A first attempt to characterize curved structures can be found in [13], which shows how complex patterns arise from simple deterministic signals both of artificial and natural origin.

While this first paper provides the existence and the reproducibility of the curved patterns, in this paper we investigate the nature of these structures, showing how the sampling plays a fundamental role in the building of the patterns. In this sense we will speak about *interference patterns*. We will focus our attention on the understanding of the formation of the macrostructures and we will discuss how the interference effect may be used to detect very slight nonstationarity such as frequency and phase shifts in periodic signals, which, as already shown in [13], are too small to be detected by linear techniques like spectrograms and power spectra.

II. THE RECURRENCE PLOT

Starting from the time series $s(t)=\{s_1,\ldots,s_n\}$, the attractor of the underlying dynamics is reconstructed in a phase space by applying the time delay vector method [14]. The reconstructed trajectory can be expressed as a matrix $\mathbf{X}=[\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_m]^T$ where each row is a phase space vector $\mathbf{x}_i=[s_i,s_{i+T},\ldots,s_{i+(D_E-1)T}]$, with $m=n-(D_E-1)T$. D_E and T are called *embedding dimension* and *delay time*.

The *recurrence plot* (RP) is a two-dimensional plot defined by $\mathbf{R}_{i,j}^{D_E,\epsilon}=\Theta(\epsilon-\|\mathbf{x}_i-\mathbf{x}_j\|)$, where $\mathbf{x}_{i,j}\in\mathbb{R}^{D_E}$ are the embedded vectors, $i,j\in[1,\ldots,m]$, $\Theta(\cdot)$ is the Heaviside step function, and ϵ is an arbitrary threshold. In the graphical representation, each nonzero entry of $R_{i,j}$ is marked by a black dot in the position (i,j). Since any state is recurrent with itself, the RP matrix fulfills $R_{i,i}=1$ which hence contains the diagonal *line of identity* (LOI).

To compute a RP, a norm must be defined. We use the L_∞ norm because it is independent of the phase space dimension and no rescaling of ϵ is required. Furthermore, special attention must be given to the choice of the threshold ϵ. There is not a specific guideline for this estimation, but the noise level of the time series must be taken into account. Values suggested are some percentage of the diameter of the attractor (in any case, not more than 10%) [15].

III. INTERFERENCE PATTERNS AND MACROSTRUCTURES

For *macropattern* or *macrostructure* we mean an aggregate of short lines parallel to the LOI forming structures having some "familiar" shape. A simple example is given in Fig. 1(a), in which the recurrence plot of the signal $s(t)=\sin(2\pi 1000t)$ (sampled with a frequency of 11 025 Hz)

*Electronic address: a.facchini@unisi.it

†Electronic address: kantz@mpipks-dresden.mpg.de

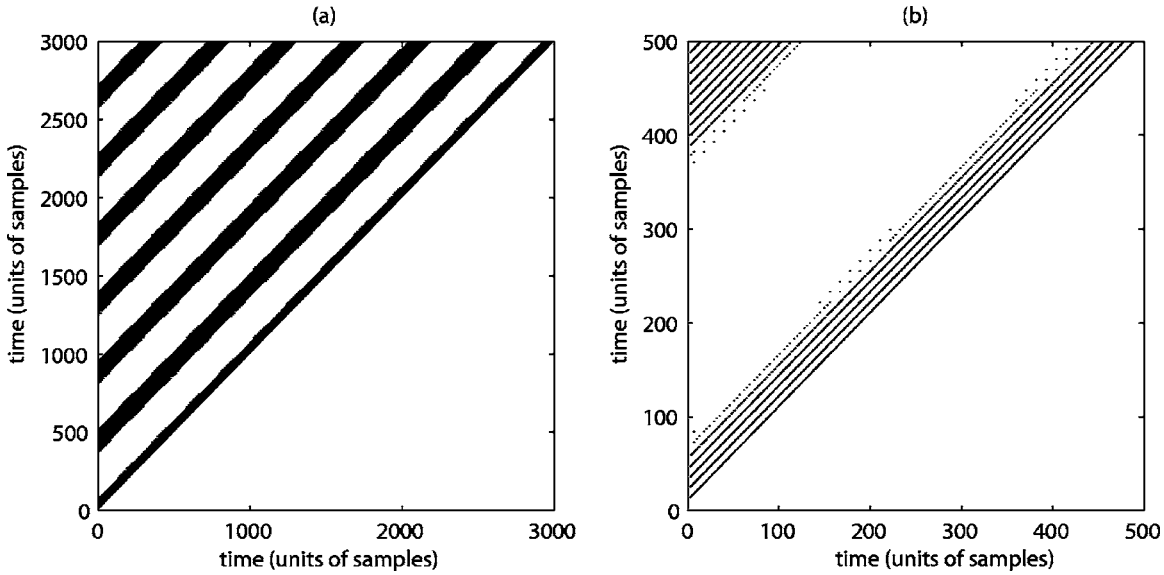

FIG. 1. The RP of the signal $\sin(2\pi 1000t)$, with $f_s=11\,025$ Hz. Here two time scales are clearly visible. The smaller is related to the period of the signal, the larger is due to the interference effect.

is shown. Figure 1(b) is an enlargement, and shows clearly that the *macrolines* are formed by a distribution of lines spaced by the period of the signal, which is 11 025 sample times. Therefore even if the RP has a simple origin, one observes the coexistence of two time scales, the first one related to the period of the signal, and the second one, much larger, related to a frequency that cannot be retraced in the signal itself.

Other macropatterns can be produced by adding different types of modulations to a periodic signal. By adding a slight linear increase of the frequency, the macrolines begin to bend, producing macrohyperbolas (see Fig. 2). Signals with periodic modulation of the carrier frequency or of the phase show circular and curved patterns. Figure 3(a) shows the pattern formed by a periodically phase modulated signal, while Fig. 3(b) shows the pattern formed by a periodically frequency modulated signal.

All the recurrence plots of this section and of the successive section are computed using $D_E=2$, $T=3$, $\epsilon=0.05$. With different embedding dimensions and delay times the structure of the patterns is conserved. The LOI is removed.

IV. ORIGIN OF THE PATTERNS: THE ROLE OF THE SAMPLING

Only a signal continuous in time has perfect recurrences. The same signal recorded with a discrete sampling time generally will have imperfect recurrences, since the exact ones may be omitted by the sampling. For a sinusoidal signal, ideally, the RP should contain all lines parallel to the LOI whose distances to the diagonal are an integer multiple of the period. Due to the error introduced by the sampling, several of these lines are broken or absent. In this sense, the thresh-

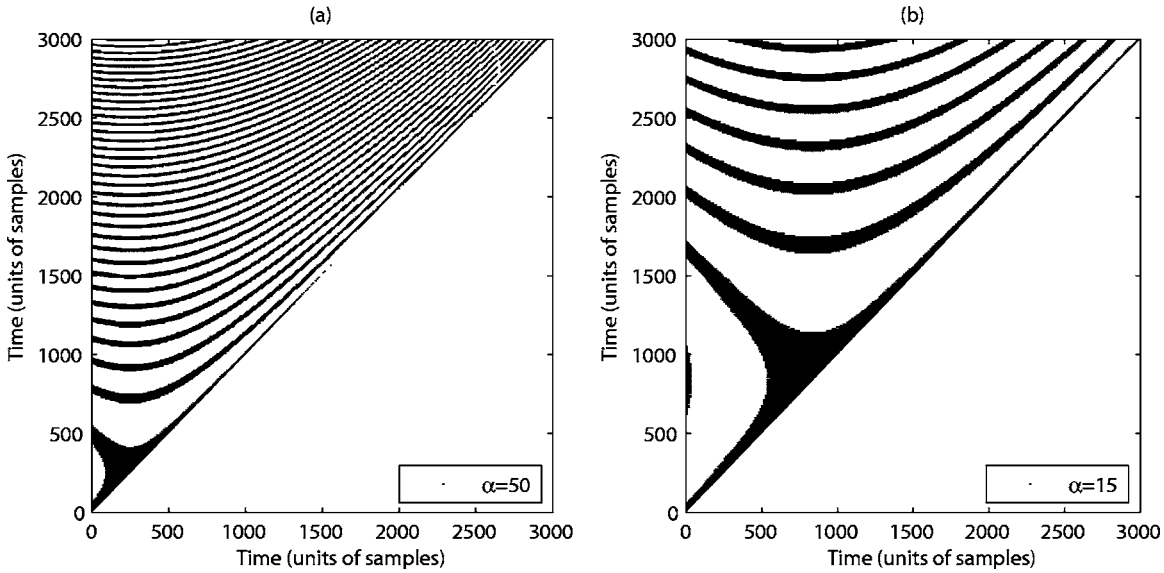

FIG. 2. (a) Recurrence plot of the signal $\sin(2\pi 1000t+2\pi 50t^2)$. (b) $\sin(2\pi 1000t+2\pi 15t^2)$, $f_s=11\,025$ Hz.

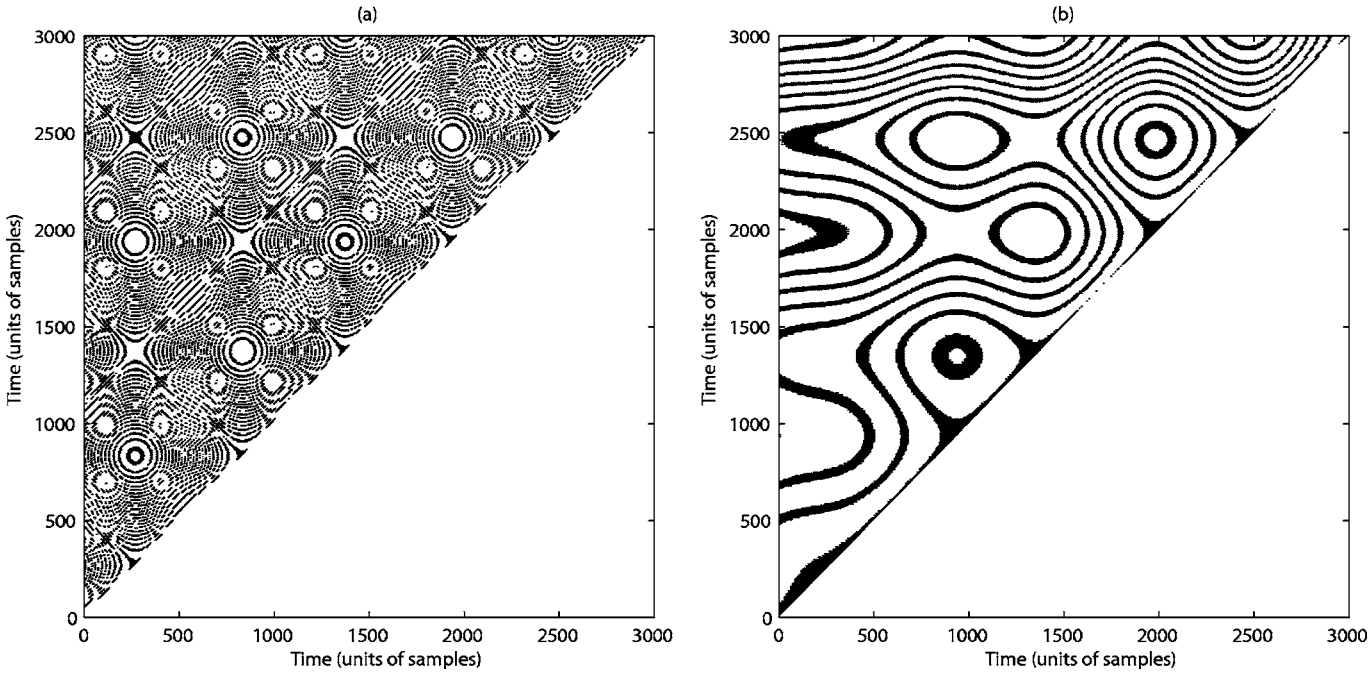

FIG. 3. (a) Recurrence plot of the signal $\sin[2\pi 1000t+2\pi\sin(2\pi 10t)]$. (b) Recurrence plot of the signal $\sin[2\pi 1000t+2\pi\sin(2\pi 10t)t]$. $f_s=11\,025$ Hz.

old ϵ represents the tolerable phase mismatch between two delay vectors in order to be still neighbors in the sense of $|\mathbf{x}_i-\mathbf{x}_j|<\epsilon$. A recurrence happens when an integer multiple of the sampling interval matches another integer multiple of the period so that their difference modulo 2π is less than the phase tolerance. It should be noticed that an increase of the threshold joins the broken lines, but the others tend to thicken and even if large values are used (60% or more of the width of the phase space) the macrolines continue to exist.

The role of the error can be easily shown considering a periodic signal $s(t)=\sin(2\pi f_c t)$. Denoting the sampling frequency f_s, the discrete representation of the signal is $s_j=\sin(2\pi f_c/f_s j)$. In the following we will indicate with $p=f_s/f_c$ the number of samples necessary for the representation of an oscillation. Normally p is not integer, therefore we will write

$$p=\alpha+\beta,\quad \alpha\in\mathbb{N},\quad \beta\in\mathbb{R},\quad |\beta|<\frac{1}{2}. \tag{1}$$

In this framework an oscillation contains α samples, while in every cycle a phase shift of β is introduced. The phase shift is fundamental for the formation and the structures of the patterns, and plays the most important role, since the sampling introduces an intrinsic phase error, which is not visible when the sampled signal is reconstructed in the time domain.

We will first focus on a nonmodulated sinusoidal signal, which is helpful to set up the vocabulary and the framework that will be useful when we will analyze modulated sinusoids. In the next sections, the carrier frequency is fixed to $f_c=1000$ Hz, and the sampling frequency is fixed to 11 025 Hz, which is one of the standard sampling frequencies for audio recording.

V. NONMODULATED SIGNALS

We consider the signal $s(t)=\sin(2\pi f_c t)$ and its numerical representation:

$$s_j=\sin\left(2\pi\frac{f_c}{f_s}j\right). \tag{2}$$

If the signal is sampled according to the Shannon theorem, its time representation is faithful. This is not true looking at the position of the samples after one cycle. There is a shift between a generic sample s_j and its corresponding $s_{j+\alpha}$ after one oscillation. Following our formalism we will compute through a numerical procedure the terms α and β:

$$p=\alpha+\beta,\quad \alpha=\text{round}(p),\quad \beta=p-\alpha,\quad p,\beta\in\mathbb{R},\quad \alpha\in\mathbb{N}, \tag{3}$$

round$(\cdot)$ returns the nearest integer to p.

For this specific signal, $\alpha=11$ samples and $\beta=0.025$. Since $\beta\neq 0$, there is a constant shift in the position of the samples after every cycle. Figure 4 shows the path of the first sample in time. This is done by downsampling the signal by $\alpha=11$ samples. If the representation of the signal were correct, the path would be a straight line. In this case, we observe that all the samples $\{s_i\}_{i=1}^{11}$ (contained in the first oscillation) go along a sinusoidal path, which has a frequency much smaller than the original signal. Therefore the sampling error introduces a new time scale which is only visible considering the motion of the samples. The position of the phase space vectors is modified too by the interference, as shown in Fig. 5(a).

Since the signal is periodic, one could expect that a generic state vector $\mathbf{x}_0(i)$ $(i=1,\ldots,\alpha)$ is recurrent with all the other vectors $\mathbf{x}_k(i)$, where $\mathbf{x}_k(\cdot)$ indicates the same state vec-

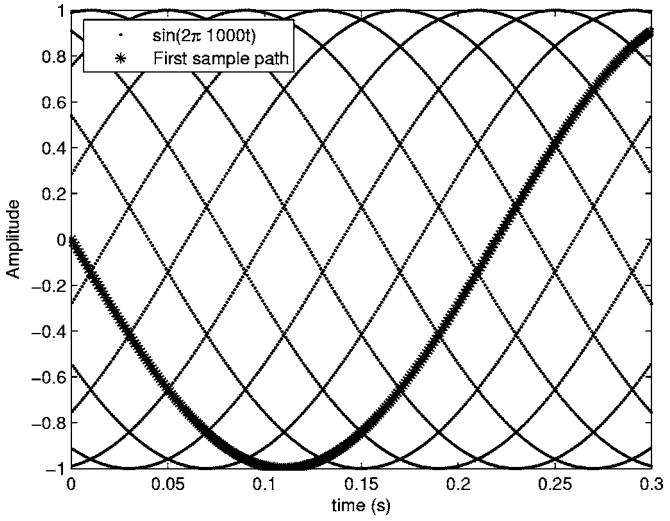

FIG. 4. The plot of the signal $\sin(2\pi f_c t)$, with $f_c=1000$ Hz and $f_s=11\,025$ Hz, every 11 samples. This shows how the position of the zeros changes with time. The other samples follow the same sinusoidal path.

tor after an arbitrary number of oscillations k.

If we represent in a three-dimensional space the time evolution of the phase vectors, one would expect that every vector $\mathbf{x}_k(i)$, lying on the (x,y) plane, moves on a straight line along the z axis. Figure 5(a) shows that instead of a straight line, the path is a helix.

Therefore, in the phase space, one can see that after a certain number of oscillations k' the vector $\mathbf{x}_{k'}(0)$ is no more recurrent with $\mathbf{x}_0(0)$, and that, in the mean time, $\mathbf{x}_0(0)$ is recurrent with $\mathbf{x}_{k'}(1)$. As shown in Fig. 6, the cross-recurrent effect is extended to all the α samples present in an oscillation. This gives an explanation of the macrolines observed in the RP of a sinusoidal signal. When the recurrence between $\mathbf{x}_0(0)$ and $\mathbf{x}_{k'}(0)$ is missed, a new recurrence arises between $\mathbf{x}_0(0)$ and $\mathbf{x}_{k'}(1)$. After $2k'$ oscillations, this new recurrence is missed too, and $\mathbf{x}_0(0)$ will be recurrent with $\mathbf{x}_{2k'}(2)$. In general, a certain sample $\mathbf{x}_0(i)$, contained in the first oscillation, will be recurrent with all the other α samples at time intervals multiples of k'.

The RP shows then three scales, the first, and smallest, is related to the main frequency of the signal, the second and much larger, is an effect of the interference: it is the time necessary to perform k' oscillations. The third scale, not visible, is the time at which $\mathbf{x}_0(i)$ is again recurrent with itself, i.e., with $\mathbf{x}_{\alpha k'}(i)$.

In order to understand the formation of the macrolines and to compute their macrofrequency, we consider that the signal has perfect recurrences only when

$$2\pi\frac{f_c}{f_s}j = 2k\pi \quad \Rightarrow \quad j_k = \alpha k + \beta k \tag{4}$$

for $k=1$, $j_1=p$. After a certain number of oscillations k', the product $\beta k'=1$ and $j_{k'} \in \mathbb{N}$ [16]. Now, we consider $k=1/\beta$ as the number of cycles at which p again becomes an integer. Under this assumption the number of samples at which a

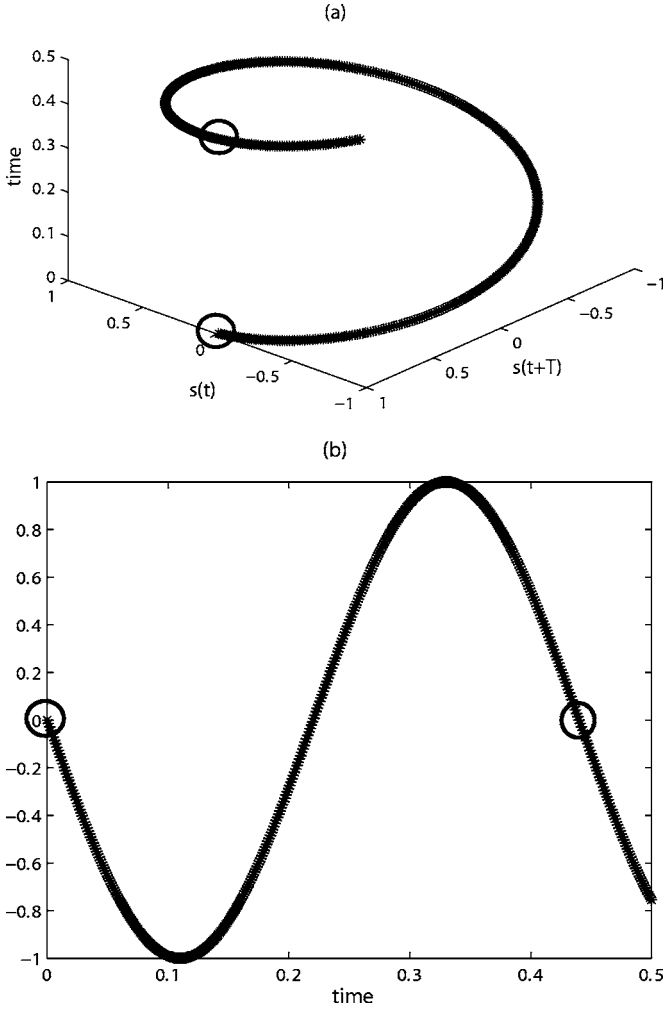

FIG. 5. (a) Motion of the state vector $\mathbf{x}_k(0)$ in the phase space [(x,y) plane] and in time (z axis). Instead of moving on a straight line, the path is a helix. (b) The path of the sample $s_k(0)$ in the time. The circles indicate the time at which a new correct recurrence happens.

new recurrence occurs is $j_{k'}=\alpha/\beta+1$. The macrofrequency (MF) is now straightforwardly determined as

$$f_M = \frac{f_s}{j_{k'}} = \left|\frac{f_s\beta}{\alpha+\beta}\right|. \tag{5}$$

With $f_c=1000$, $f_s=11\,025$, using this formalism, the three time scales are

(1) $j_c=\alpha=11$ samples, i.e., the time scale of the carrier frequency f_c;

(2) $j_{k'}=\alpha/\beta+1=441$ samples, i.e., the macrofrequency $f_M=25$ Hz of the lines; and

(3) $j^*=j_c j_{k'}$ samples, i.e., the recurrence time between the same phase space vectors, whose MF is 2.27 Hz.

This is in agreement and extends our preceding results in [13].

In general, the value of the MF depends on f_c and f_s. Fixing f_s, we can compute the value of the MF for different

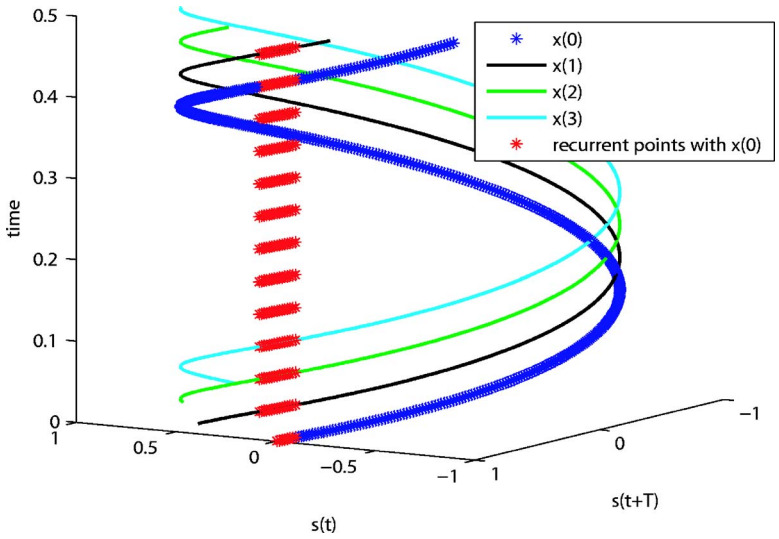

FIG. 6. (Color online). Recurrences of the state vector $\mathbf{x}(0)$ with $\mathbf{x}_{ik'}(i)$ $(i=1,\ldots,\alpha)$. Because of the motion of the vectors, $\mathbf{x}(0)$ recurs with itself and with the other α vectors.

values of the carrier frequency. Figure 7 shows how the MF varies with f_c, while the sampling frequency acts as a scale factor for the picture. The picture shows also how the value of the MF increases and decreases dramatically, and, for some specific values of f_c, for which $\beta=0$, it shows strong discontinuities. From the picture, one can see that the RP representation of even periodic signals varies largely with the frequency. Furthermore, for those frequencies which have a large value of the MF, we observe a second order interference. In this case we speak about macrofrequencies of order n (f_{M_n}). By example, we know that the MF of a 1000 Hz signal is 25 Hz. If we consider a signal at 3062 Hz, we find that its MF is again 1000 Hz, therefore the RP will show again the same macrostripes at 25 Hz, i.e., the value of the MF shown by the RP is not univocal.

In general, a signal whose MF reaches high values cannot be visualized because of the interference. Experimentally we found that a good representation of the f_{M_1} is available only when the ratio $\rho=f_M/f_c \leq 0.18$, for higher values the stripes in the RP should be referred to the f_{M_2} or of higher order. Figure 8 shows the first three orders of $f_{M_{i(f_c)}}$.

In the movie `non_mod.avi` [17] this effect is clearly visible. It shows the evolution of the RP for the frequency range $f_c \in [1000, 1200]$, the green point shows the corresponding $f_{M_{1(f_c)}}$.

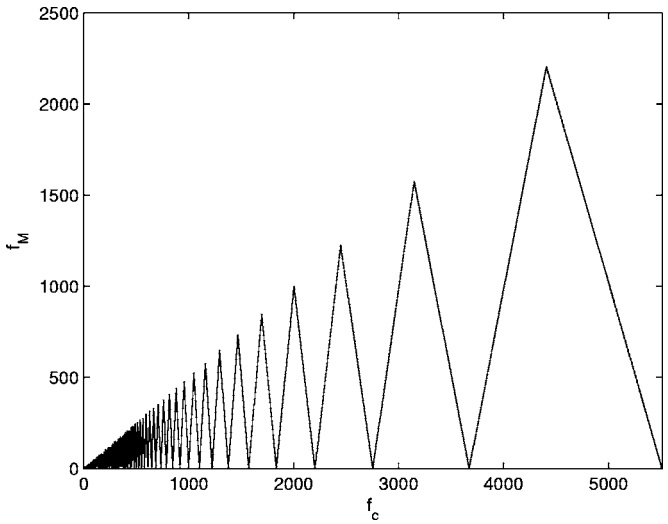

FIG. 7. The MF in function of f_c for a nonmodulated sinusoid. Despite the simplicity of the signal, the MF shows with the increase of the frequency a nontrivial behavior. The frequencies corresponding to $\beta=0$ correspond theoretically to $f_M=0$.

VI. MODULATED SIGNALS

The computation of the MF given in Eq. (5) can be extended when the frequency or the phase of the signal varies with time. We will then consider cases:

(1) linear increase of the frequency;
(2) periodic modulation of the phase; and
(3) frequency modulation.

We will consider f_c as a function of time (or j), and for the

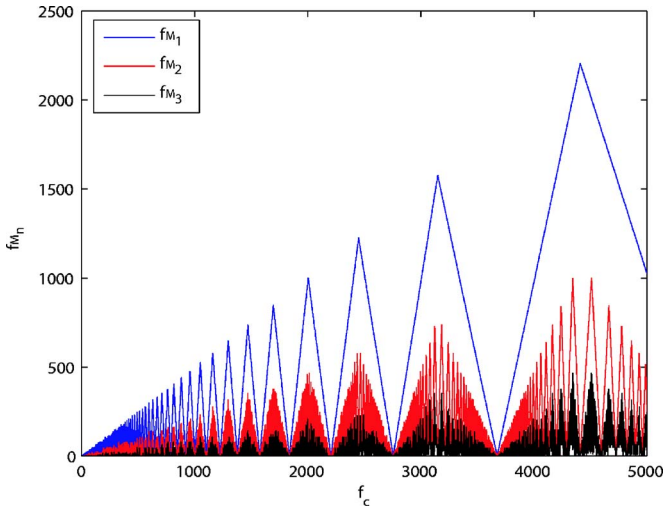

FIG. 8. (Color online). The orders of the MF (f_{M_n}) for the unmodulated signals.

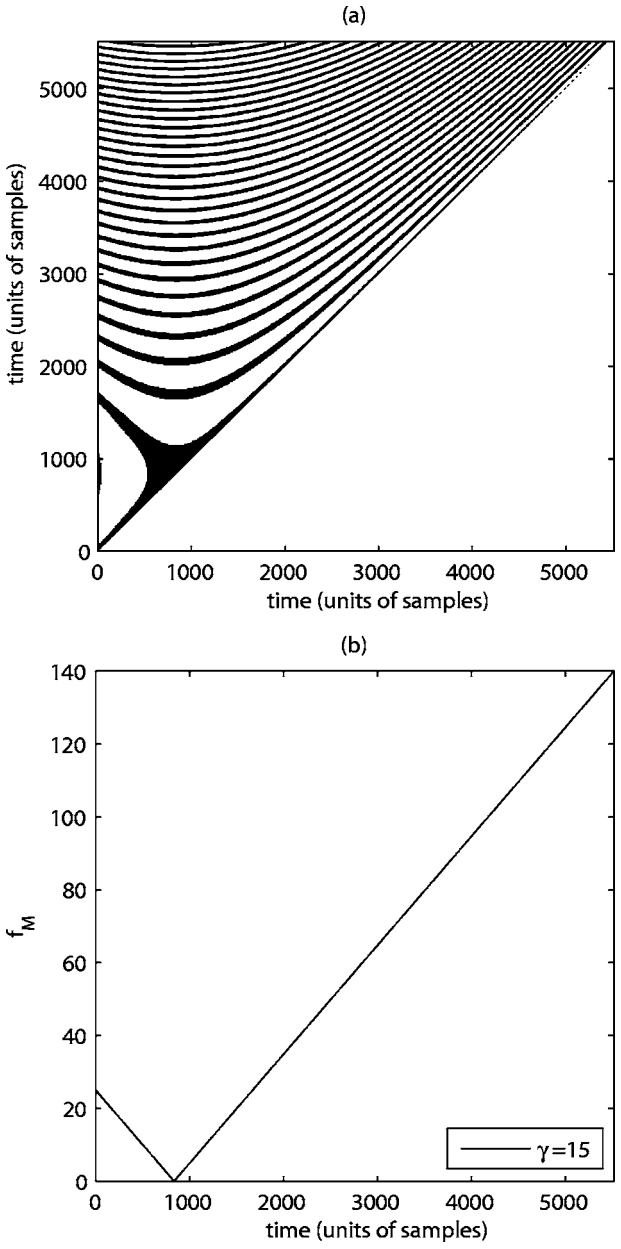

FIG. 9. The MF in function of the time, $f_c=1000$, $\gamma=15$.

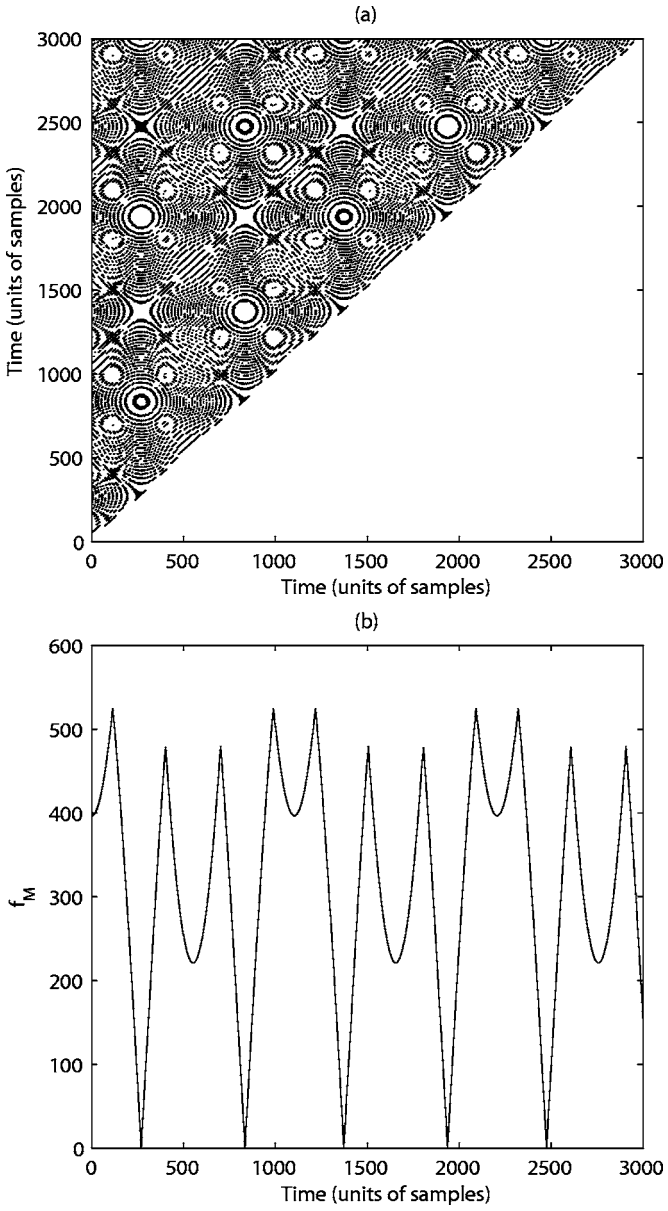

FIG. 10. The MF in function of the time, $f_c=1000$, $f_m=10$.

computation of the $p(j)$ we will consider the instantaneous frequency of the signal.

A. Linear increase of the frequency

The signal is $\sin(2\pi f_c t+2\pi\gamma t^2)$ while the instantaneous frequency is $\frac{1}{2\pi}\frac{d\omega}{dt}=f_c+2\gamma t=f_c(t)$. Therefore the number of samples becomes

$$p(j)=\frac{f_s}{f_c(j)}=\frac{f_s}{f_c+\frac{2\gamma j}{f_s}}=\frac{f_s^2}{f_c f_s+2\gamma j}. \tag{6}$$

Now $\alpha(j)$ and $\beta(j)$ can now be computed as functions of time following Eq. (5).

Figure 9 shows the variation of the MF for a signal with $f_c=1000$ and $\gamma=15$. The value of the MF increases with time, and this is in agreement with the RP shown in the above plot.

The movie `lin_mod.avi` [17] shows the evolution of the recurrence plot for increasing values of $\gamma\in[0,100]$: after a fast change in the patterns, the shape of the RP becomes more stable and assumes the characteristic hyperbolic shape. It should be also noticed that the RP starts with a decreasing MF and with a *singular* pattern, corresponding to very low values of the MF. After that, there is a constant increase. As a final remark, the movie shows some gapped patterns in the structure (especially for high γ values). This may be related to higher macrofrequencies orders.

B. Periodic phase shift

Considering $s(t)=\sin[2\pi f_c t+2\pi\sin(2\pi f_m t)]$, following the same procedure we obtain

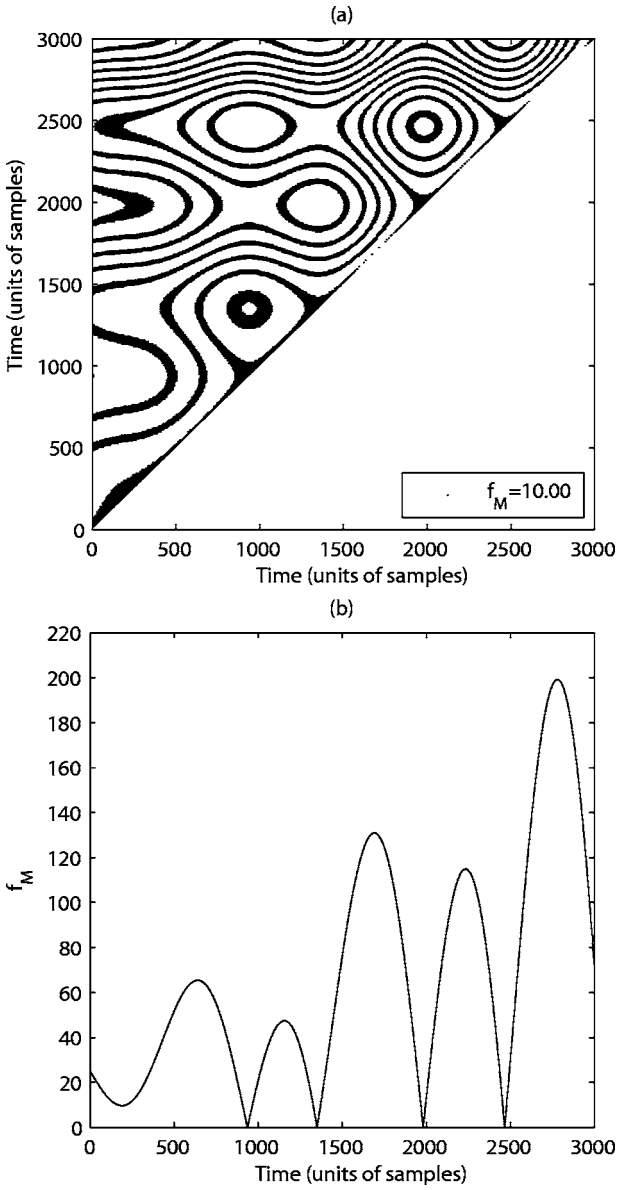

FIG. 11. The MF in function of the time, $f_c=1000$, $f_m=10$.

$$\frac{1}{2\pi}\frac{d\omega}{dt} = f_c + 2\pi f_m \cos(2\pi f_m t), \tag{7}$$

therefore the number of samples in cycle is

$$p(j) = \frac{f_s}{f_c + 2\pi f_m \cos\left(2\pi \frac{f_m}{f_s} j\right)}. \tag{8}$$

Figure 10 shows the strong variations of the MF for this class of signals. The MF oscillates with frequency f_m, this behavior is also found in the spatial structure of the RP. The movie `per_mod.avi` [17] shows how the RP evolves by changing the modulating frequency f_m.

C. Frequency modulation

Considering $s(t)=\sin[2\pi f_c t + 2\pi \sin(2\pi f_m t)t]$, we have obtained

$$f_c(j) = f_c + \frac{2\pi f_m}{f_s} \cos\left(2\pi \frac{f_m}{f_s} j\right) j + \sin\left(2\pi \frac{f_m}{f_s} j\right) \tag{9}$$

therefore

$$p(j) = \frac{f_s}{f_c + \frac{2\pi f_m}{f_s} \cos\left(2\pi \frac{f_m}{f_s} j\right) j + \sin\left(2\pi \frac{f_m}{f_s} j\right)}. \tag{10}$$

Figure 11 shows the evolution of the MF for the frequency modulated signal. In comparison with Fig. 10, the maximum values of the MF changes continuously in time. The different behavior is confirmed if one looks to the movie `freq_mod.avi` [17] that shows the evolution of the patterns with the increase of the modulating frequency.

VII. CONCLUDING REMARKS

In this paper we have interpreted the origin of the curved macrostructures that can be observed by computing the RP of modulated sinusoids. Despite the simplicity of the signals, the patterns were not trivial. By means of a nonmodulated sinusoid, we set up a formal framework based on the fact that the signal discretization introduces an intrinsic phase error β whose effects are evident only in the phase space (or in the time domain). In this sense, the main point regarding the origin of the macrostructures is that the large time scales do not really exist in the signal. They are the artificial product of the interference induced by the sampling of a continuous signal, while the scale of the structures depends on the sampling time.

In [13], we stated experimentally that RPs were able to detect phase or frequency shifts in the order of 5% of f_c. We can now theoretically state that the detection accuracy is one order lower, i.e., 0.5%. In fact, the macrofrequency varies largely even for a shift of 4 or 5 Hz (see, for example, Fig. 7 or the movie `non_mod.avi` [17]). The same conclusions apply for modulated signals, whose pattern structure is complicated by the time dependence of β and α.

ACKNOWLEDGMENTS

A.F. is grateful to Dr. Norbert Marwan and to Professor Nobuo Masataka for interesting discussions and comments.

[1] J. Eckmann, S. Kamphorst, and D. Ruelle, Europhys. Lett. **5**, 973 (1987).
[2] P. Kaluzny and R. Tarnecki, Biol. Cybern. **68**, 527534 (1993).
[3] F. D. Santo, F. Gelli, A. Schmied, J.-P. Vedel, A. Rossi, and R. Mazzocchio, J. Neurosci. Methods **155**, 116121 (2006).
[4] N. Thomasson, T. J. Hoeppner, C. L. Webber, Jr., and J. P.

Zbilut, Phys. Lett. A **279**, 94101 (2001).
[5] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, Phys. Rev. E **66**, 026702 (2002).
[6] J. P. Zbilut, J. C. Mitchell, A. Giuliani, A. Colosimo, N. Marwan, and C. L. Webber, Physica A **343**, 348 (2004).
[7] Z.-B. Wu, Phys. Lett. A **332**, 250 (2004).
[8] L. Matassini, H. Kantz, J. Hołyst, and R. Hegger, Phys. Rev. E **65**, 021102 (2002).
[9] C. Webber and J. Zbilut, J. Appl. Physiol. **76**, 965 (1994).
[10] P. Faure and H. Korn, Physica D **122**, 265 (1998).
[11] C. Letellier, Phys. Rev. Lett. **96**, 254102 (2006).
[12] N. Marwan and J. Kurths, Phys. Lett. A **336**, 349 (2005).
[13] A. Facchini, H. Kantz, and E. Tiezzi, Phys. Rev. E **72**, 021915 (2005).
[14] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, England, 2004).
[15] N. Marwan, Ph.D. thesis, Institut für Physik, Universität Potsdam, 2003 (unpublished).
[16] This strong condition can be relaxed considering $\beta k' \simeq 1$.
[17] See EPAPS Document No. E-PLEEE8-75-092703 for movies showing the evolution of the recurrence plots under different modulating frequency conditions. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.