Origin of the Usefulness of the Natural-Time Representation of Complex Time Series

Sumiyoshi Abe,[1,*] N. V. Sarlis,[2] E. S. Skordas,[2,3] H. K. Tanaka,[2,4] and P. A. Varotsos[2,3,†]

[1]*Institute of Physics, University of Tsukuba, Ibaraki 305-8571, Japan*
[2]*Solid State Section, Physics Department, University of Athens, Panepistimiopolis, Zografos 157 84, Athens, Greece*
[3]*Solid Earth Physics Institute, Physics Department, University of Athens, Panepistimiopolis, Zografos 157 84, Athens, Greece*
[4]*Earthquake Prediction Research Center, Tokai University, Shizuoka 424-8610, Japan*

(Received 3 December 2004; published 2 May 2005)

The concept of natural time turned out to be useful in revealing dynamical features behind complex time series including electrocardiograms, ionic current fluctuations of membrane channels, seismic electric signals, and seismic event correlation. However, the origin of this empirical usefulness is yet to be clarified. Here, it is shown that this time domain is in fact optimal for enhancing the signals in time-frequency space by employing the Wigner function and measuring its localization property.

DOI: 10.1103/PhysRevLett.94.170601 PACS numbers: 05.40.−a, 05.45.Tp, 05.90.+m, 87.10.+e

Characterization of complex time series $\{x(t)\}$ and prediction of catastrophic events have always been of general common interest in biology, earth science, and physics (e.g., [1–5]). In the analysis of such time series, the role of time itself and the possibility of introducing its reparametrization is very important. The concept of time reparametrization is, of course, not novel in the theory of stochastic processes and applications, e.g., [6–11]. In general [11], let $\{x(t)\}$ be a stochastic process, and $\theta(t)$ an increasing function of t. The process $X(t) \equiv x[\theta(t)]$ is called a compound process. The index t denotes the clock (conventional) time, and $\theta(t)$ is called trading time or time deformation process. A special form of compound process is subordination (e.g., [8]), as developed by Bochner [6], and applied to financial markets by Mandelbrot and Taylor [7] and later refined by Clark [9] to explain the speculative prices. Note that the range of values of the new process $\{X(t)\}$ is a subset of the range of values of $\{x(t)\}$.

In recent works [12–19], it has been shown that novel dynamical features hidden behind the time series can emerge if we (abandon $\{x(t)\}$ or $\{X(t)\}$ and) represent the time series in terms of the "natural time". Natural time, labeled [20] χ, is defined [12,13] by ascribing to the kth pulse (once the initial pulse is identified) the value $\chi_k = k/N$, where N is the total number of pulses considered, and then representing [see Fig. 1(b)] each pulse by the energy emitted in that pulse, which for dichotomous signals are proportional to its duration Q_k. This way, the whole complex time series $\{x(t)\}$ is transformed to the pairs (χ_k, Q_k); see Figs. 1(a) and 1(b). The usefulness of this representation in time series analysis, has already been demonstrated when distinguishing similar looking signals that are emitted from systems of different dynamics. Examples are the discrimination of sudden cardiac death individuals from healthy humans through analysis of their electrocardiograms [16,19], or of seismic electric signals (SES) activities (i.e., a series of electrical pulses detected before earthquakes [21–23]) from irrelevant background noise [14,15]. Another application of natural time refers to the manifestation of aging and scaling properties in seismic event correlation [17,18]. We emphasize that these results could not be obtained if the analyses were carried out in the conventional time domain. The most important point regarding natural time may be that it enables us to follow the dynamical evolution of a system and identify when it enters into a critical stage. Therefore, it can play a major role in predicting impending catastrophic events such as a strong earthquake occurrence [23] and sudden cardiac death [16,19,24]. However, the question remains to be solved why natural time exhibits more advantages than conventional time.

In this Letter, we address ourselves to the problem of optimality of the natural-time representation of time series resulting from complex systems that may contain catastrophic events. For this purpose, first we study the structures of the time-frequency representations [25] of the signals by employing the Wigner function [26] to compare the natural-time representation with the ones, either in conventional time or other possible reparametrizations. We shall see that significant enhancement of the signal is observed in the time-frequency space if natural time is used, in marked contrast to other time domains. To quantify this localization property, we examine the generalized entropic measure proposed by Tsallis [27], which has been widely discussed in the studies of complex dynamical systems. In time series analysis, it is desired to reduce uncertainty and extract signal information as much as possible. Consequently, the most useful time domain should maximize the information measure, and hence minimize the entropy. We find that this can statistically be ascertained in natural time, by investigating a multitude of different time domains.

Consider a signal $\{x(t)\}$ represented in conventional time, t. The normalized time-frequency Wigner function associated with it is defined by

$$W(t, \omega) = A \int d\tau e^{-i\omega\tau} x(t - \tau/2) x(t + \tau/2), \quad (1)$$

where $A = [\pi \int dt x^2(t)]^{-1}$ is the normalization constant and ω is the frequency. Numerically, it is necessary to discretize and make finite both time and frequency, and the integral has to be replaced by a sum. To make a comparison of the natural-time analysis with Eq. (1), it is convenient to rescale χ_k by $N\chi_k$, which is precisely the pulse number, $k \equiv t_k$. The quantity Q_k has a clear meaning for dichotomous time series (Fig. 1), whereas for non-dichotomous time series, the threshold should be appropriately put (e.g., the mean value plus half of the standard deviation) to transform it to a dichotomous one. The normalized Wigner function associated with Q_k is now given as follows:

$$W(k, \tilde{\omega}) = B \sum_{i=0}^{N-1} Q_{k-i} Q_{k+i} \cos[\tilde{\omega}(t_{k+i} - t_{k-i})], \qquad (2)$$

where $B = [\pi \sum_{k=1}^{N} Q_k^2]^{-1}$ stands for the normalization constant and $\tilde{\omega}$ is the dimensionless "frequency" (see the later comment). In the sum, Q_k with $k \leq 0$ and $k > N$ should be set equal to zero. It is noted that Eq. (2) is a discrete version of the continuous Wigner function in Eq. (1) and unlike the ordinary definition, the transformation in Eq. (2) is not orthogonal, in general.

In Fig. 2, we present the plots of the Wigner functions in the time-frequency spaces with conventional time and natural time. Remarkably, significant enhancement of the signal is observed in the latter case, with the scale of enhancement being about 10 times. In contrast to a moderate profile in the conventional time representation, a localized structure emerges in natural time.

In the natural-time domain, the time difference between two consecutive pulses (i.e., interoccurrence time) is equally spaced and dimensionless, and is here taken to be unity: $t_{k+1} - t_k = 1$. However, for comparison, later we will consider various time domains, in which the occur-

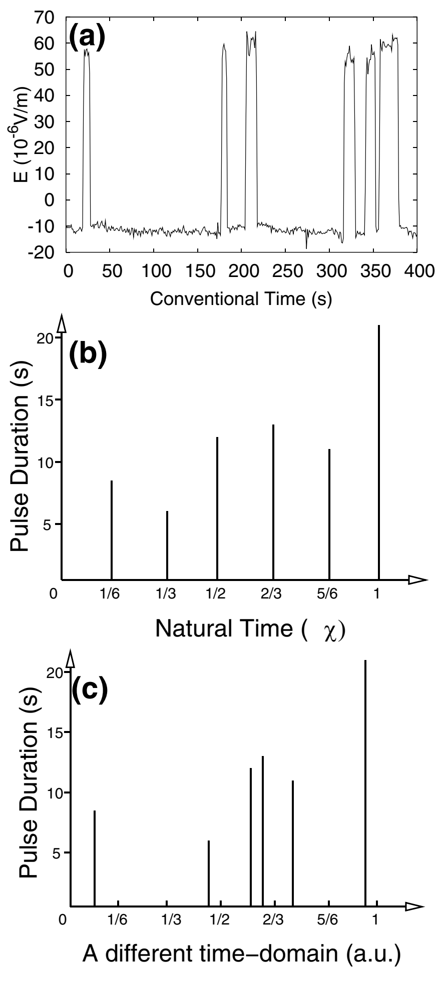

FIG. 1. An example of observed time series of SES activity represented in (a) conventional time, (b) natural time, and (c) a randomly generated time. In (b), the natural time serves as an index of the occurrence of each pulse (reduced by the total number of pulses), while the amplitude is proportional to the duration of each electric pulse [12–15].

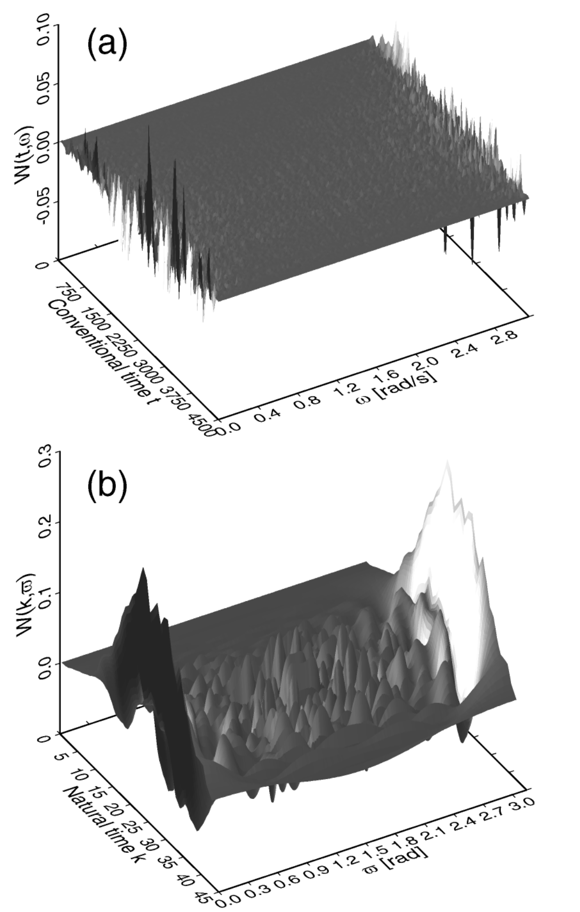

FIG. 2. The plots of the Wigner functions of the SES activity A in Fig. 3 given below in (a) the conventional time domain and (b) the natural-time domain. Significant enhancement of the signal is recognized in the natural-time domain at both edges but mainly in the localized structures in the intermediate region. Note that, instead of χ_k, $N\chi_k = k$ is used (see the text). ω has the unit [rad/ sec], whereas $\tilde{\omega}$ has [rad].

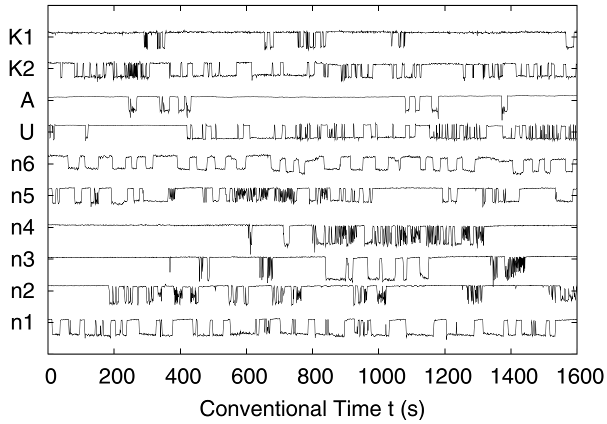

FIG. 3. Excerpts of 4 SES activities, labeled K1, K2, A, U; and 6 artificial noises, labeled n1-n6, in arbitrary scales.

rence time $t_k = Nu_k$ in Eq. (2) is made random. The conventional time representation is characterized by a constant time increment Δt (e.g., 1 sec), and the occurrence of the ith event is at $t_i = i\Delta t$. Differences between the three time domains are shown in Fig. 1. To generate the random time domains artificially, we randomize u_k by making use of the uniform distribution defined in the interval (0,1) so that the average interoccurrence time is again unity. Performing Monte Carlo simulation, we have constructed more than 1000 different time domains and integrated over ω ($\tilde{\omega}$) from 0 to π [rad/ sec] ([rad]), which can cover the regimes of interest [recall that when $t_k = k$, $W(k, \omega + \pi) = W(k, \omega)$].

To quantify the degrees of disorder in the time-frequency spaces with various time domains, we employ, as mentioned, the Tsallis entropy [27] defined by

$$S_q = \frac{1}{1-q}\left(\int d\mu W^q - 1\right), \tag{3}$$

where $\int d\mu$ is the collective notation for integral and sum over the time-frequency space and q is the positive entropic index. In the limit $q \to 1$, this quantity tends to the form of the Boltzmann-Gibbs-Shannon entropy $S = -\int d\mu W \ln W$. This limit cannot however be taken, since the Wigner function is a pseudodistribution and takes negative values, in general. S_q is, however, well defined if q is even. Thus, we propose to use the value

$$q = 2, \tag{4}$$

which, using Eqs. (2) and (3) , results in

$$S_2 = 1 - \frac{1}{2\pi} \times \left\{ \frac{\sum_{k=1}^{N}\sum_{l=0}^{N-1}\sum_{l'=0}^{N-1} Q_{k-l}Q_{k+l}Q_{k-l'}Q_{k+l'}\left(\delta_{l+l',0} + \frac{\sin[\pi(t_{k+l}-t_{k-l}+t_{k+l'}-t_{k-l'})]}{\pi(t_{k+l}-t_{k-l}+t_{k+l'}-t_{k-l'})} + \delta_{l,l'} + \frac{\sin[\pi(t_{k+l}-t_{k-l}-t_{k+l'}+t_{k-l'})]}{\pi(t_{k+l}-t_{k-l}-t_{k+l'}+t_{k-l'})}\right)}{\left[\sum_{k=1}^{N}(Q_k^2 + \sum_{l=1}^{N-1} Q_{k-l}Q_{k+l}\frac{\sin[\pi(t_{k+l}-t_{k-l})]}{\pi(t_{k+l}-t_{k-l})})\right]^2} \right\}. \tag{5}$$

To examine how the natural-time representation is superior to other ones, we have made a comparison of the values of S_2 for 10 different time series [15] of electric signals (see Fig. 3): 4 SES activities and 6 "artificial" noises (i.e., noises emitted from nearby electrical sources). The results are shown in Table I in which we give the values of $p(S_2 < S_2^{\rm nat})$, i.e., the probability that S_2 calculated with a time domain different than the natural-time domain is smaller than the value $S_2^{\rm nat}$ calculated with natural time. This probability $p(S_2 < S_2^{\rm nat})$ was estimated as follows: For each time domain produced by Monte Carlo simulations, the corresponding S_2 value was calculated through Eq. (5) and compared to $S_2^{\rm nat}$. An inspection of Table I shows that among the signals investigated only two, i.e., A and n6, have a considerable probability $p(S_2 < S_2^{\rm nat})$, i.e., $\approx 28.5\%$ and 26%, respectively. This can be attributed to the small number of pulses ($N \approx 40$) of these signals for the following reason: In Fig. 4 we present the dependence of $p(S_2 < S_2^{\rm nat})$ versus the number of pulses for the simplified example of all $Q_k = 1$; this figure shows that $p(S_2 < S_2^{\rm nat})$ decreases upon increasing N, starting from $\approx 36\%$ at $N \approx 50$. In other words, Table I reveals that, for signals with a reasonable number of pulses, e.g., larger than 2×10^2, the quantity $S_2^{\rm nat}$, in fact, tends to be minimum compared to those of other representations attempted. In addition, it is

TABLE I. The number of N pulses and the values of $p(S_2 < S_2^{\rm nat})$ for the 10 electric signals analyzed. The estimation error is at the most 1.6%.

Signal	N	$p(S_2 < S_2^{\rm nat})(\%)$
K1	312	3.7
K2	141	6.9
A	43	28.5
U	80	8.1
n6	42	26.0
n5	432	2.8
n4	396	1.6
n3	259	2.7
n2	1080	<0.1
n1	216	5.7

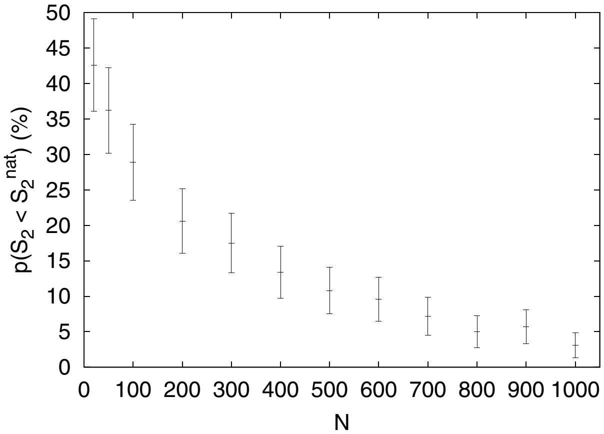

FIG. 4. The values of $p(S_2 < S_2^{\rm nat})$ versus the number of pulses for the simple example of a time series consisting of pulses with all $Q_k = 1$.

mentioned that $S_2^{\rm nat}$ is also appreciably smaller than S_2 in conventional time (see Fig. 2).

In conclusion, we have studied if natural time yields an optimal representation for enhancing the signals in the time-frequency space by employing the Wigner function and measuring its localization property by means of the Tsallis entropy. For this purpose, we have compared the values of the entropy for various observed time series represented in a multitude of different time domains. We have found that the entropy is highly likely to be minimum for natural time, implying the least uncertainty in the time-frequency space. This explains why dynamical evolutions of diverse systems can be better described in the natural-time domain, in particular, when systems approach to a critical state. Important examples of the latter are: in cardiology (e.g., an impending sudden cardiac death), seismology (e.g., when a strong earthquake is approached), etc.

S. A. would like to thank the Physics Department of the University of Athens for hospitality extended to him. He was also supported in part by the Grant-in-Aid for Scientific Research of Japan Society for the Promotion of Science.

*Electronic address: suabe@sf6.so-net.ne.jp
†Electronic address: pvaro@otenet.gr

[1] D. Sornette, A. Johansen, A. Arneodo, J. M. Muzy, and H. Saleur, Phys. Rev. Lett. **76**, 251 (1996).
[2] D. Sornette, Phys. Rep. **313**, 237 (1999).
[3] R. Yulmetyev, P. Hänggi, and F. Gafarov, Phys. Rev. E **62**, 6178 (2000).
[4] R. Yulmetyev, F. Gafarov, P. Hänggi, R. Nigmatullin, and S. Kayumov, Phys. Rev. E **64**, 066132 (2001).
[5] R. G. Turcott and M. Teich, Ann. Biomed. Eng. **24**, 269 (1996).
[6] S. Bochner, *Harmonic Analysis and the Theory of Probability* (University of California, Berkeley, 1960).
[7] B. Mandelbrot and H. M. Taylor, Oper. Res. **15**, 1057 (1967).
[8] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. II* (Wiley, New York, 1971).
[9] P. K. Clark, Econometrica: Journal of the Econometric Society **41**, 135 (1973).
[10] E. Ghysels, C. Gourieroux, and J. Jasiak, *Centre for Interuniversity Research and Analysis on Organizations* (Centre for Interuniversity Research and Analysis on Organizations, New York, 1995), Working paper 95s-42.
[11] B. B. Mandelbrot, A. Fisher, and L. Calvet, *Cowles Foundation Discussion* (Yale University, Yale, 1997), Discussion paper 1164.
[12] P. A. Varotsos, N. V. Sarlis, and E. S. Skordas, Practica of Athens Academy **76**, 294 (2001).
[13] P. A. Varotsos, N. V. Sarlis, and E. S. Skordas, Phys. Rev. E **66**, 011902 (2002).
[14] P. A. Varotsos, N. V. Sarlis, and E. S. Skordas, Phys. Rev. E **67**, 021109 (2003).
[15] P. A. Varotsos, N. V. Sarlis, and E. S. Skordas, Phys. Rev. E **68**, 031106 (2003).
[16] P. A. Varotsos, N. V. Sarlis, E. S. Skordas and M. S. Lazaridou, Phys. Rev. E **70**, 011106 (2004).
[17] S. Abe and N. Suzuki, Physica A (Amsterdam) **332**, 533 (2004).
[18] U. Tirnakli and S. Abe, Phys. Rev. E **70**, 056120 (2004).
[19] P. A. Varotsos, N. V. Sarlis, E. S. Skordas, and M. S. Lazaridou, Phys. Rev. E **71**, 011110 (2005).
[20] The symbol χ originates from the ancient Greek word $\chi\rho\acute{o}\nu$os (chronos) which means time (e.g., chronometer, chronology, chronological order come from this word).
[21] P. Varotsos, K. Alexopoulos, K. Nomicos, and M. Lazaridou, Nature (London) **322**, 120 (1986).
[22] S. Uyeda, T. Nagao, Y. Orihara, T. Yamaguchi, and I. Takahashi, Proc. Natl. Acad. Sci. U.S.A. **97**, 4561 (2000).
[23] P. Varotsos, *The Physics of Seismic Electric Signals* (TerraPub, Tokyo, 2005).
[24] D. Graham-Rowe, New Scientist **182**, No. 2441, 10 (2004).
[25] L. Cohen, *Time-Frequency Analysis: Theory and Applications* (Prentice-Hall, Upper Saddle River NJ, 1994).
[26] E. Wigner, Phys. Rev. **40**, 749 (1932).
[27] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).