Universal patterns in sound amplitudes of songs and music genres

R. S. Mendes,[1,2] H. V. Ribeiro,[1,2,*] F. C. M. Freire,[1] A. A. Tateishi,[1,2] and E. K. Lenzi[1,2]

[1]*Departamento de Física, Universidade Estadual de Maringá, Av. Colombo 5790, 87020-900, Maringá, Paraná, Brazil*

[2]*National Institute of Science and Technology for Complex Systems, CNPq, Rua Xavier Sigaud 150, 22290-180, Rio de Janeiro, Rio de Janeiro, Brazil*

(Received 20 August 2010; revised manuscript received 10 November 2010; published 12 January 2011)

We report a statistical analysis of more than eight thousand songs. Specifically, we investigated the probability distribution of the normalized sound amplitudes. Our findings suggest a universal form of distribution that agrees well with a one-parameter stretched Gaussian. We also argue that this parameter can give information on music complexity, and consequently it helps classify songs as well as music genres. Additionally, we present statistical evidence that correlation aspects of the songs are directly related to the non-Gaussian nature of their sound amplitude distributions.

DOI: 10.1103/PhysRevE.83.017101 PACS number(s): 89.90.+n, 89.20.−a, 43.75.St, 05.45.Tp

In recent years, studies of complex systems have become widespread among the scientific community, especially in statistical physics [1–5]. Many of these investigations deal with data records ordered in time or space (i.e., time series), trying to extract some features, patterns, or laws that may be present in the systems studied. This approach has been successfully applied to a variety of fields, from physics and astronomy [6] to genetics [7] and economics [8]. Moreover, this framework has been used to investigate and model interdisciplinary fields, such as religion [9], elections [10], vehicular traffic [11], tournaments [12], and many others. These few examples and social phenomena in general [13] illustrate that physicists have gone far from their traditional domain of investigations.

Music is a well-known worldwide social phenomenon linked to the human cognitive habits, modes of consciousness, as well as historical developments [14]. In exploring music's social role, some authors investigated collective listening habits. For instance, Lambiotte and Ausloos [15] analyzed data from people's music libraries, finding that audience groups with size distribution follow a power law. They also investigated correlations among these music groups, reporting nontrivial relations [16]. In another work, Silva *et al.* [17] studied the network structure of composers and singers of Brazilian popular music (mpb). There is also interest in studying the patterns of music sales [18] and the success of particular musicians [19–21].

Despite these cultural aspects, songs form a highly organized system presenting very complex structures and long-range correlations. All these features have attracted the attention of statistical physicists. In a seminal paper, Voss and Clarke [22] analyzed the power spectrum of radio stations and observed a $1/f$ noise-like pattern. They also showed that the correlation can extend to longer or shorter time scales, depending on the music genre. Hsü and Hsü [23] investigated the changes of acoustic frequency in Bach's and Mozart's compositions, finding self-similarity and fractal structures. In contrast, they report no resemblance to fractal geometry [24] for modern music. Fractal structures have also been reported in the study of sequences of music notes [25], and Su and Wu [26] suggested that the multifractal spectrum could be used to distinguish different styles of music. By using sound amplitudes of songs, Bigerelle and Iost [27] achieved a classification based on fractal dimension using the entire frequency range. However, as discussed by Ro and Kwon [28], the $1/f$ analysis in the region below 20 Hz might not classify music genres. Gündüz and Gündüz [29] reported analysis of several Turkish songs using many techniques. Beltrán del Río *et al.* [30] evaluated the rank distribution of music notes of a large selection, finding good agreement with a two-parameter β distribution. Dagdug *et al.* [31] investigated a specific piece of Mozart with detrended fluctuation analysis (DFA) [32]. Applying DFA in a volatility-like series, Jennings *et al.* [33] found quantitative differences in the Hurst exponent depending on the music genre.

In this brief literature review, we see that special attention was paid to the fractal structures of music, correlations, and power spectrum analysis. However, much less attention has been paid to understanding amplitude distribution. This last point has been noted by Diodati and Piazza [34]. In their work, they investigated the distribution of times and sound amplitudes larger than a fixed value. By using this kind of return interval analysis [35], they found Gaussian distributions in the amplitudes for jazz, pop, and rock music, while non-Gaussians emerge for classical pieces. Here, we directly investigate the amplitude distributions of songs of several genres without employing a threshold value as considered by Diodati and Piazza. Moreover, our analysis goes toward finding patterns in the amplitude sound distribution by using a suitable one-parameter probability distribution function (PDF). In the following, we present the dataset used in our investigation, the analysis of the shape of the resulting distributions, and our conclusions.

Not all sound is music, but certainly music is made by sounds. The sounds that we hear are a consequence of pressure fluctuations traveling in the air and hitting our ears. These audible pressure fluctuations can be converted into a voltage signal u_t using a record system and stored, for instance, in a compact disk (CD). Our analysis is focused on this time series u_t, which we call sound amplitude. In the case of songs stored on CDs, u_t has a standard sampling rate of 44.1 kHz and encompasses the full audible human range (approximately between 20 and 20 kHz).

As a database, we have 8115 songs of nine different music genres: classical (907), tango (992), jazz (700), hip-hop (876),

*hvr@dfi.uem.br

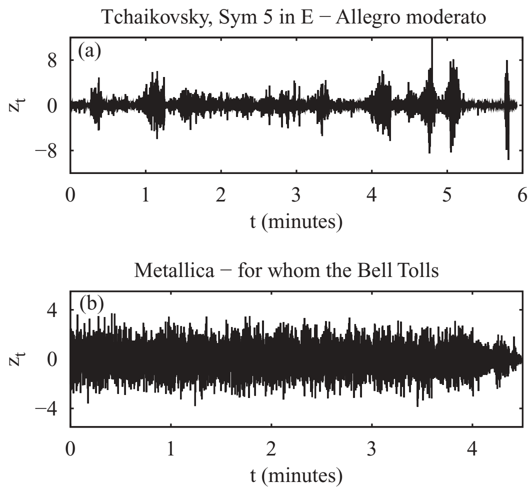

FIG. 1. The normalized sound amplitude of (a) a classical piece and (b) a heavy-metal song (labeled in the figure). Note that the signals are quite different; the first one presents a more complex structure characterized by "bursts" while the second resembles Gaussian noise.

mpb (580), flamenco (524), pop (998), techno (900), and heavy metal (1638). The songs were chosen as to cover a large number of composers and singers. For instance, for classical music, we have taken pieces from Bartók, Beethoven, Berlioz, Brahms, Bruch, Chopin, Dvorak, Fauré, Grieg, Malher, Marcello, Mozart, Rachmaninov, Strauss, Schuber, Schumann, Scriabin, Shostakovich, Sibelius, Stravinsky, Tchaikovsky, Verdi, Vivaldi, and others.

When a time series is analyzed, a way to view its variability (complexity) is at least in part by investigating its PDF. In the case of music, the mean amplitude is approximately zero since a vibration essentially occurs around this value. In addition, the mean (global) intensity is not relevant to the variability (complexity) of a song. Motivated by these facts, our research is based on the PDF of recorded data regardless of their mean value and their real amplitudes. In other words, we consider that the complexity of a song is not related to its mean intensity but to the relative variability of the amplitudes. Thus, instead of employing the amplitude u_t in different time instants t, we focus attention on u_t subtracted from its mean value μ and divided by its standard deviation σ. This corresponds to using $z_t = (u_t - \mu)/\sigma$ instead of u_t. Figure 1 illustrates the behavior of z_t for two songs, a classical piece and a heavy-metal song. This figure is enough to reveal qualitative differences between these two songs. In the classical piece, we can observe some kind of bursts giving rise to a non-Gaussian distribution. However, for the heavy-metal song, the signal is very similar to Gaussian noise—no complex structure is perceptible.

Motivated by these distinct behaviors, we investigated the distribution of z_t for all the songs in our data set. In Fig. 2, we show the PDF for some representative songs. As we can verify from this figure, the shape of distributions goes from a long tail to Laplace to Gaussian distribution. A family of functions that has the Gaussian and the Laplace distributions as a particular case is given by the stretched Gaussian [36,37] $p(z) = N \exp(-b|z|^c)$, where N is the normalization constant, b is directly related to the standard deviation, and c is a positive parameter. Since the distribution $p(z)$ is normalized to unity and the variable z is defined in such a way that its standard deviation is equal to 1, the parameters N and b become functions exclusively of c, leading to

$$p(z) = \frac{c}{2}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)^3}\right)^{1/2} \exp\left(-\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{c/2} |z|^c\right), \quad (1)$$

with $\Gamma[w]$ being the Euler γ function. Also in Fig. 2, the least square fits to the data of the function are shown. Observe that we find a good agreement between the data and the model for the songs represented in this figure, and a similar agreement has been found for the others (at least in the central part of the distribution).

The only model parameter is c, and it may give useful information about music complexity. First note that for values of c smaller than 1, heavy tail distributions emerge. In some sense, these heavy tails reflect the complex structures that we see in Fig. 1(a), that is, larger fluctuations. Increasing c makes the tails shorter and recovers some known distributions (Laplace for $c = 1$ and Gaussian for $c = 2$). In this context, a shorter tail indicates that larger fluctuations become rare, leading to music signals very similar to Gaussian noise [see Fig. 1(b)]. From the musical point of view, the word *complexity* may be related to several aspects of the song or even to music taste. In the present context, it should be viewed a comparative measurement, that is, a measure of how the empirical distributions differs from the Gaussian one.

Based on this discussion, we may use c to sort the songs and music genres in a kind of complexity order (smaller c is related to a large complexity). In order to construct this rank for music genres, we evaluate the mean value of c over all songs of each music genre considered here as shown in Fig. 3(a). Our findings agree with other works in the sense that there is a quantitative difference between classic and light or dancing music [33,34]. However, it is interesting to note that music genres are not well defined [40]. Thus, any taxonomy may be controversial, representing an open problem of automatic classification like other problems of pattern recognition. To take a glance at this complicated problem, we also evaluated the probability distribution of c for each music genre as shown in Fig. 3(b). We can see that there are overlapping regions for all genres, reflecting the fuzzy boundaries existent in the music genre definitions.

Despite the complex situation that emerges in the problem of automatic genre classification [41–44], our model is very simple. From the qualitative point of view, the characteristics of songs and music genres are related with multidimensional aspects such as timber, melody, harmony, and rhythm, among others. Thus, as a minimalist model, the classification presented here must be viewed as a kind of global measure for these qualitative aspects. In addition, we have to note that correlation aspects are lost when we consider only the histogram as presented in Fig. 2. In the same way, information is also lost when someone considers only some correlations. However, we remark that the results concerning the genre classification, here obtained by using only the PDF of sound amplitude, are in statistical agreement

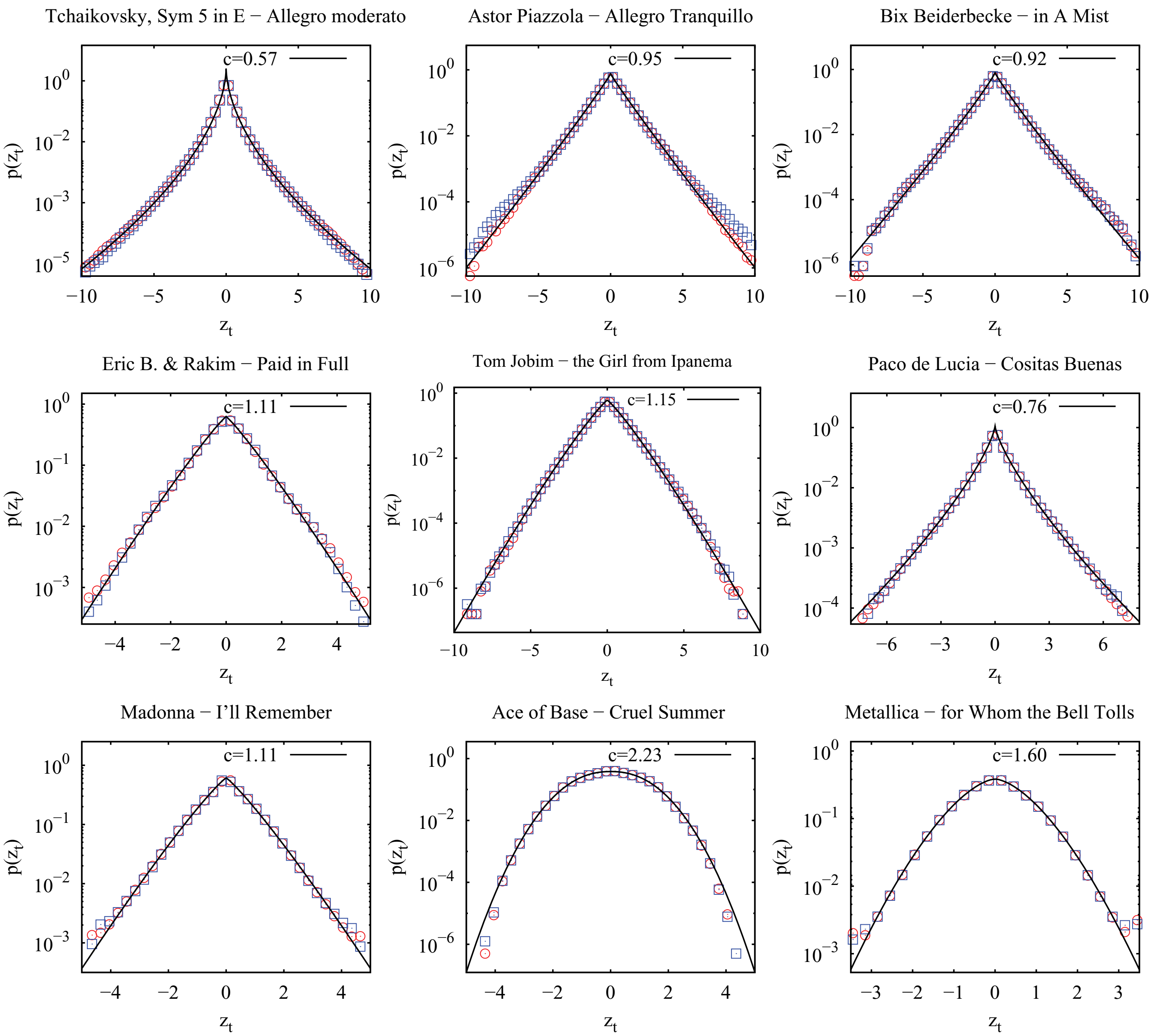

FIG. 2. (Color online) Histograms of some representative songs (labeled in the figure) in comparison with the stretched Gaussian, Eq. (1). The squares (circles) are in the right (left) channel of the stereo audio. As we see, the two channels are quite similar in the sense that the statistical results do not dramatically change when considering the right or left channel.

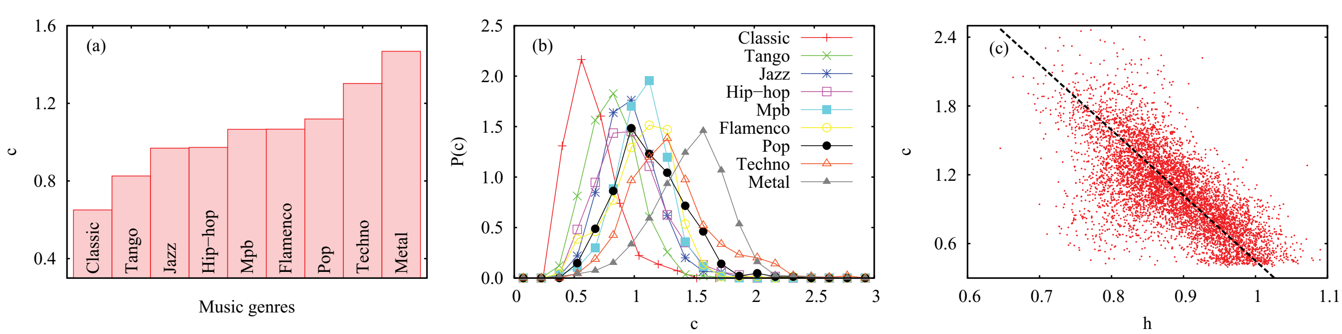

FIG. 3. (Color online) (a) In ascending order, the mean values of the parameter c corresponding to the stretched Gaussians employed for each music genre considered here. (b) The distribution of the parameter c for each genre. (c) Scatter plot of the parameter c vs the Hurst exponent, h, obtained via detrended fluctuation analysis (DFA) [38,39] of sound intensity z_t^2. The dashed line is a guide for our eyes.

with the results of other methods based on correlation analysis. This fact suggests a kind of coupling between the correlation aspects and the non-Gaussian PDFs. Aiming to highlight this feature, we evaluated the Hurst exponent (h) of the time series z_i^2 and plotted it versus the PDF parameter c in Fig. 3(c). The data presented in this figure suggest an approximate linear relation between c and h (Pearson correlation about -0.7), providing statistical evidence that the non-Gaussian nature of the PDFs are directly related to the correlations in songs. Therefore, these two complementary aspects and others compose the multidimensional nature of music quantification and classification.

To sum up, we investigated the probability distribution of the normalized sound amplitudes for more than 8000 musical pieces. The empirical findings suggest a universal form of distribution, which agreed well with a stretched Gaussian. Because of the normalization and the standard deviation fixed as 1, our distribution has only one parameter, c. We argue that this parameter goes toward quantifying the complexity of songs as well as music genres. In addition to this universal feature, we presented empirical evidences that non-Gaussian nature of sound amplitude PDF is related to the correlation aspects. As an application, we also hope that the distribution of sound amplitudes presented here may have implications for stochastic music compositions.

We thank CNPq/CAPES for the financial support and the local university radio for kindly providing the songs.

[1] S. Y. Auyang, *Foundations of Complex Systems* (Cambridge University Press, Cambridge, 1998).
[2] H. J. Jensen, *Self-Organized Criticality* (Cambridge University Press, Cambridge, 1998).
[3] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[4] N. Boccara, *Modeling Complex Systems* (Springer-Verlag, New York, 2004).
[5] D. Sornette, *Critical Phenomena in Natural Sciences* (Springer-Verlag, Berlin, 2006).
[6] S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).
[7] C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons, and H. E. Stanley, Nature **356**, 168 (1992).
[8] R. N. Mantegna and H. E. Stanley, *An Introduction to Econophysics* (Cambridge University Press, Cambridge, 1999).
[9] S. Picoli and R. S. Mendes, Phys. Rev. E **77**, 036105 (2008).
[10] S. Fortunato and C. Castellano, Phys. Rev. Lett. **99**, 138701 (2007).
[11] D. Chowdhury, L. Santen, and A. Schadschneider, Phys. Rep. **329**, 199 (2000).
[12] H. V. Ribeiro, R. S. Mendes, L. C. Malacarne, S. Picoli Jr., and P. A. Santoro, Eur. Phys. J. B **75**, 327 (2010).
[13] C. Castellano, S. Fortunato, and V. Severo, Rev. Mod. Phys. **81**, 591 (2009).
[14] T. DeNora, *The Music of Everyday Life* (Cambridge University Press, Cambridge, 2000).
[15] R. Lambiotte and M. Ausloos, Phys. Rev. E **72**, 066107 (2005).
[16] R. Lambiotte and M. Ausloos, Eur. Phys. J. B **50**, 183 (2006).
[17] D. L. Silva *et al.*, Physica A **332**, 559 (2004).
[18] R. Lambiotte and M. Ausloos, Physica A **362**, 485 (2006).
[19] J. A. Davies, Eur. Phys. J. B **27**, 445 (2002).
[20] E. P. Borges, Eur. Phys. J. B **30**, 593 (2002).
[21] H.-B. Hu and D.-Y. Han, Physica A **387**, 5916 (2008).
[22] R. F. Voss and J. Clarke, Nature **258**, 317 (1975).
[23] K. J. Hsü and A. Hsü, Proc. Natl. Acad. Sci. U.S.A. **88**, 3507 (1991).
[24] K. J. Hsü and A. Hsü, Proc. Natl. Acad. Sci. U.S.A. **87**, 938 (1990).
[25] Z.-Y. Su and T. Wu, Physica A **380**, 418 (2007).
[26] Z.-Y. Su and T. Wu, Physica D **221**, 188 (2006).
[27] M. Bigerelle and A. Iost, Chaos Solitons Fractals **11**, 2179 (2000).
[28] W. Ro and Y. Kwon, Chaos Solitons Fractals **42**, 2305 (2009).
[29] G. Gündüz and U. Gündüs, Physica A **357**, 565 (2005).
[30] M. Beltrán del Río, G. Cocho, and G. G. Naumis, Physica A **387**, 5552 (2008).
[31] L. Dagdug, J. Alvarez-Ramirez, C. Lopez, R. Moreno, and E. Hernandez-Lemus, Physica A **383**, 570 (2007).
[32] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, Phys. Rev. E **49**, 1685 (1994).
[33] H. D. Jennings, P. C. Ivanov, A. M. Martins, P. C. da Silva, and G. M. Viswanathan, Physica A **336**, 585 (2004).
[34] P. Diodati and S. Piazza, Eur. Phys. J. B **17**, 143 (2000).
[35] M. S. Santhanam and Holger Kantz, Phys. Rev. E **78**, 051113 (2008).
[36] The streched Gaussian was introduced in the context of the anomalous diffusion by Richardson [37] by considering a spatial-dependent diffusion coefficient.
[37] L. F. Richardson, Proc. R. Soc. London A **110**, 709 (1926).
[38] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, Phys. Rev. E **49**, 1685 (1994).
[39] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, and A. Bunde, Physica A **295**, 441 (2001).
[40] N. Scaringella, G. Zoia, and D. Mlynek, IEEE Signal Process. Mag. **23**, 133 (2006).
[41] C. N. Silla, C. A. A. Kaestner, and A. L. Koerich, Proc. IEEE Int. Conf. on Systems, Man and Cybernetics, 1687 (2007).
[42] J. Serrà, X. Serra, and R. G. Andrzejak, New J. Phys. **11**, 093017 (2009).
[43] M. M. Mostafa and N. Billor, Expert Systems with Applications **36**, 11378 (2009).
[44] D. C. Correa, J. H. Saito, and L. F. Costa, New J. Phys. **12**, 053030 (2010).