Physica A 533 (2019) 121967

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Oscillations in the Tsallis income distribution

Everton M.C. Abreu [a,b,c,*], Newton J. Moura Jr. [d], Abner D. Soares [e], Marcelo B. Ribeiro [f,g]

[a] *Departamento de Física, Universidade Federal Rural do Rio de Janeiro–UFRRJ, Seropédica, RJ 23890-971, Brazil*
[b] *Departamento de Física, Universidade Federal de Juiz de Fora–UFJF, Juiz de Fora, Brazil*
[c] *Programa de Pós-Graduação Interdisciplinar em Física Aplicada, Instituto de Física, Universidade Federal do Rio de Janeiro-UFRJ, 21941-972, Rio de Janeiro, RJ, Brazil*
[d] *Instituto Brasileiro de Geografia e Estatística–IBGE, Rio de Janeiro, Brazil*
[e] *Comissão Nacional de Energia Nuclear–CNEN, Rio de Janeiro, Brazil*
[f] *Instituto de Física, Universidade Federal do Rio de Janeiro–UFRJ, Rio de Janeiro, Brazil*
[g] *Observatório do Valongo, Universidade Federal do Rio de Janeiro–UFRJ, Rio de Janeiro, Brazil*

HIGHLIGHTS

- Income distribution analysis through Tsallis thermostatistical formalism.
- Complexification of the q-parameter shows the oscillatory behavior.
- The numerical example demonstrates that the q-complexification represents the empirical data.

ARTICLE INFO

Article history:
Received 20 June 2017
Received in revised form 2 July 2019
Available online 10 July 2019

Keywords:
Income distribution
Tsallis statistics
Complex q-parameter

ABSTRACT

Oscillations in the complementary cumulative distribution function (CCDF) of individual income data have been found in the data of various countries studied by different authors at different time periods, but the dynamical origins of this behavior are currently unknown. Although these datasets can be fitted by different functions at different income ranges, the Tsallis distribution has recently been found capable of fitting the whole distribution by means of only two parameters. This procedure showed clearly such oscillatory feature in the entire income range feature, but made it particularly visible at the tail of the distribution. Although log-periodic functions fitted to the data are capable of describing this behavior, a different approach to naturally disclose such oscillatory characteristics is to allow the Tsallis q-parameter to become complex. In this paper we use this idea in order to describe the behavior of the CCDF of the Brazilian personal income recently studied empirically by Soares et al. (2016). Typical elements of periodic motion, such as amplitude and angular frequency coupled to this income analysis, were obtained by means of this approach. A highly non-linear function for the CCDF was obtained through this methodology and a numerical test showed it capable of recovering the main oscillatory feature of the original CCDF of the personal income data of Brazil.

1. Introduction

The study of the individual income distribution of populations has a long history. Vilfredo Pareto (1848–1927), the pioneer of this type of analysis, studied the distribution of personal income at the end of 19th century for some regions and

* Corresponding author at: Departamento de Física, Universidade Federal Rural do Rio de Janeiro–UFRRJ, Seropédica, RJ 23890-971, Brazil.
E-mail address: evertonabreu@ufrrj.br (E.M.C. Abreu).

https://doi.org/10.1016/j.physa.2019.121967
0378-4371/

countries in specific years and sets of years. Pareto looked at the problem systematically and concluded that, individually speaking, the richest people in a society have the *complementary cumulative distribution function* (CCDF) of income obeying a power law function [1]. Consequently, the *probability density function* (PDF) $p(x)$ of the personal income x of the richest persons may be given by,

$$p(x) = \beta x^{-(1+\alpha)} \, , \tag{1}$$

where β is a normalization constant. Through the years, this power law behavior has become known as the *Pareto power-law* and, consequently, the exponent α is now known as the *Pareto index.* This law has been interpreted later as being a classic example of a fractal distributions, where the Pareto index plays the role of the single fractal dimension of the distribution [2]. Higher values of the Pareto index imply in less uneven distribution of the personal income. In other words, a rise in the Pareto index corresponds to a fall in income inequality. The interesting detail is that this result, that is, the power-law nature of the income distribution of the richest persons in a society, has not been disputed by different investigations carried out since then, which considered several different samples obtained at different times for different populations in distinct countries or groups of countries [3–13, and references therein].

Despite its empirical success, the Pareto power-law does not work for the overwhelmingly majority less rich part of the population. Namely, it only describes well the income data of those belonging to the narrow "slice" of the richest population. To consider the income data of the group composed by the less rich people the method that has been used since shortly after Pareto's time is to fit the less rich data segment by various other functions, like the exponential, the log-normal, the gamma function, the Gompertz curve, as well as other ones [3–5,7,12–14]. There are also successful approaches for analyzing the whole data range using less than simple functions with many parameters, but usually such approaches require four or more parameters in order to fit the entire distribution. In addition, using the two-fitting-functions methodology means assuming that societies are divided in two classes only: on one side the very rich, formed by about 1% of the whole population, and on the other side the remaining 99%. One basic problem of this methodology is the absence of a middle class, which certainly exists in between these two groups. In addition, the question remains of whether or not this class division is a real feature of societies, or basically a mathematical artifact convenient for data fitting. Considering these objections, a simple function able to fit the entire income distribution which, at the same time, allows for various features to emerge at different income range is certainly of high interest.

A recent approach for representing the whole income data is to fit the data using the Tsallis functions instead of a combination of two functions as depicted above. In this approach the individual income distribution is classified in terms of the well known Tsallis parameter, *i.e.*, the q-parameter, and another normalization constant. Borges [15] used two q-parameter, where one controls the slope of the intermediate income range and the other describes the tail of the distribution. He was then able to analyze the income distribution concerning some *counties* of the USA from 1970 to 2000, Brazil from 1970 to 1996, Germany from 1992 to 1998 and the United Kingdom from 1993 to 1998. The conclusion was that an increase in q with time points to growing inequality. Greater values of q indicate greater probability to find counties much richer than others. Ferrero [16,17] used the Tsallis function to fit the entire income data of several *countries*, but only at single years, not being able though to indicate an evolution in the Tsallis parameters.

Recently [18,19] have used the Tsallis formalism to empirically study the income distribution of Brazil during a relatively large yearly time window, from 1978 to 2014. The results showed that the two fitted parameters of the Tsallis function follow a cycling behavior over time. Moreover, a linear fit of the distribution in each year showed that the data oscillate periodically around the fitted straight line with an amplitude that grows with the income values. A closer look at the fitted data made by other authors using different methods applied to different samples collected at different time periods showed a similar oscillatory pattern, which means that there seems to be indeed a second order dynamical effect not previously identified in the income data [20, p. 164]. This kind of oscillatory behavior has not been noted before in the income distribution data, although it has been observed in financial markets and other systems [20].

In this work we have analyzed this periodic oscillation through the alternative approach of allowing the Tsallis parameter to become complex, as suggested by Ref. [21]. Under this methodology the q-parameter discloses such periodic behavior in the income distribution curves, allowing us to define periodic elements that are ordinary in physics, like the amplitude and angular frequency. Such methodology is exemplified by a numerical example where the original empirically obtained CCDF of Brazil for the year 2011 is recovered once our approach is applied to the results.

This paper is organized as follows. Section 2 presents the Tsallis functions and some of their properties required in our analysis. Section 3 presents and discusses the complexification process and its influence in the analysis of the individual income distribution. Section 4 presents a numerical example of the procedure developed here, and Section 5 ends the paper with our conclusions.

2. Tsallis functions

It is well known that the Tsallis thermostatistics [22,23] is based on both the q-logarithm and q-exponential functions, given by,

$$\ln_q x \equiv \frac{x^{(1-q)} - 1}{1-q}, \tag{2}$$

$$e_q{}^x \equiv \Big[1 + (1-q)x\Big]^{1/(1-q)} . \tag{3}$$

These functions are defined such that for $q = 1$ both expressions become the standard logarithm and exponential functions, namely, $e_1{}^x = e^x$ and $\ln_1 x = \ln x$. Hence, the Tsallis q-functions are in fact the usual exponential and logarithmic expressions twisted in such a way as to be used in Tsallis' theory of non-extensive statistical mechanics [23].

At this point it should be noted that there are other ways to deform these two common functions viewing other application, such as the personal income distribution. This is the case of the κ-generalized exponential, introduced by Ref. [24], which can be used to fit the entire income data range in similar manner as the Tsallis q-functions. This is especially significant because both of them have the power-law and exponential as their limiting cases. The interested reader can find more applications of the κ-generalized function in Refs. [25–27].

From the definitions above it is clear that,

$$e_q{}^{(\ln_q x)} = \ln_q\left(e_q{}^x\right) = x. \tag{4}$$

Moreover, $\ln_q 1 = 0$ for any q. Hence, if we have a value x_0 such that $x/x_0 = 1$, as a result $\ln_q(x/x_0) = 0$. Two other properties of the q-exponential useful for our purposes here are as follows [28],

$$\left[e_q{}^{f(x)}\right]^a = e_{1-(1-q)/a}{}^{af(x)} , \tag{5}$$

$$\frac{d}{dx}\left[e_q{}^{f(x)}\right] = \frac{\left[e_q{}^{f(x)}\right]^q}{f'(x)}. \tag{6}$$

These results will be useful when we explore the fact that the Tsallis parameter q can be represented in the complex plane, as discussed in Ref. [21]. It is worth noting that from Eqs. (3), (5) and (6) it is not obvious that a complex q can help us disclose some income distribution details that are hidden in these functions, such as a periodic behavior, as we shall show below.

3. q-parameter complexification

3.1. Complex heat capacity

The fact that the nonextensivity parameter q can be represented by a complex formulation is not new. In Ref. [29] the q-parameter can be seen as a measure of the thermal bath heat capacity C, where

$$C = \frac{1}{q-1} . \tag{7}$$

Moreover, such complex C is well known in the literature [30,31] and can be written as follows,

$$C = C_\infty + \frac{C_0 - C_\infty}{1+(\omega\tau)^2}(1 - i\omega\tau) , \tag{8}$$

where C_∞ is the heat capacity for infinitely fast degrees of freedom (DOF), ω is the frequency, C_0 is the heat capacity at equilibrium of DOF where the frequency is set to zero and τ, the time constant, is the kinetic relaxation time constant of a certain DOF. The form of Eq. (8) suggests us that we can write it as $C = C' + i\,C''$, where

$$C' = C_\infty + \frac{C_0 - C_\infty}{1+(\omega\tau)^2} \tag{9}$$

and

$$C'' = \frac{(C_0 - C_\infty)\omega\tau}{1+(\omega\tau)^2} . \tag{10}$$

Hence, let us write a complex form of q as $q = q_r + iq_i$. Substituting in Eq. (7) and associating with Eq. (8), we have that

$$\frac{q_r - 1}{q_i} = \frac{C_0 + C_\infty(\omega\tau)^2}{(C_0 - C_\infty)\omega\tau} , \tag{11}$$

and for $C_0 \ll C_\infty$ we have $(1-q_r)/q_i \simeq \omega\tau$, which shows that the relationship between the real and imaginary parts of q is proportional to the frequency. More details of this discussion can be found in Ref. [29].

3.2. Complex income distribution means periodic behavior

Let us start with the suggestion of Wilk and Włodarczyk [21] for the complexification of the q-parameter. A three-parameters Tsallis distribution (TD) may be written as follows,

$$f(x) = C\,\mathrm{e}_q^{-x/T} = C\left(1 - \frac{x}{mT}\right)^{-m}, \tag{12}$$

where

$$m = \frac{1}{(q-1)}, \tag{13}$$

is a real power index, T is a scale parameter identified in thermodynamic applications, in general the standard temperature, and C is a normalization constant. The proposal is to consider m, or q, complex. Hence, the TD keeps its main quasi-power like form, however, that brings about some log-periodic oscillations. As examples, one can mention that such behavior has been encountered in many subjects, such as earthquakes [32,33], chaos [34], tracers on random systems [35–37], random quenched and fractals [38–42], specific heat [43], clusters [44], growth models [45], stock markets [46–50] and, finally, non-extensive statistical mechanics log-periodic oscillations [51]. When $m \to \infty$, namely, $q \to 1$, we have that this power-like distribution is analogous to the standard exponential distribution $f(x) = C\mathrm{e}^{-x/T}$.

The complexification proposal means turning m complex in Eq. (12), yielding,

$$m = m' + \mathrm{i}m'' \tag{14}$$

which means that we can also have a complex non-extensive q-parameter written as below,

$$q = 1 + \frac{1}{m} = q' + \mathrm{i}q''. \tag{15}$$

It is simple to see that

$$q' = 1 + \frac{m'}{|m|^2} \qquad \text{and} \qquad q'' = -\frac{m''}{|m|^2} \tag{16}$$

where

$$|m|^2 = m'^2 + m''^2. \tag{17}$$

Hence, the goal here is to analyze the results obtained in Ref. [20], where the personal income distribution of Brazil shows a periodic behavior as a function of the income variable for each yearly sample, in the light of the complexification of the q-parameter. Therefore, we wish to describe mathematically such an oscillatory behavior.

3.3. Non-extensive analysis of the income distribution of Brazil

Let us now turn our attention to the main issue of this article. As discussed in Ref. [20], if the entire income distribution range can be fitted by one function with only two parameters, a well-defined two-classes-base income structure implicitly assumed when the income range is described by two distinct functions may be open question. Therefore, such income–class division could possibly be only a result of fitting choices and not of an intrinsic feature of societies. The TD is known to become a pure power-law for large values of its independent variable x, and an exponential when x tends to zero. However, this behavior is not equivalent to assuming from the start a two-classes approach to the income distribution problem because the TD will only have power-law and exponential like behaviors as limiting cases. Thus, a possible different behavior at the intermediate level might not be described by neither of these functions. This means that the TD does not necessarily imply in two very distinct classes based on well-defined income domain ranges, but possibly having an intermediate income range of unknown size which might behave as neither of them. Bearing these points in mind, let us now proceed with the description of the income distribution in terms of the TD and its subsequent complexification.

Let $\mathcal{F}(x)$ be the *cumulative distribution function* (CDF) of the personal income, representing the proportionality, or probability, that a person receives an income less than or equal to x. Let us now denote its complementary version, the CCDF, by $F(x)$, which then describes the probability that a person receives an income greater or equal to x. It is then clear that,

$$\mathcal{F}(x) + F(x) = 100\ . \tag{18}$$

Here the maximum probability is normalized to 100% instead of the standard unity value. The boundary conditions involved in both functions are $\mathcal{F}(x) = F(\infty) \cong 0$ and $\mathcal{F}(\infty) = F(0) \cong 100$. In addition, the following properties apply to these income functions,

$$\frac{\mathrm{d}\mathcal{F}(x)}{\mathrm{d}x} = -\frac{\mathrm{d}F(x)}{\mathrm{d}x} = f(x), \tag{19}$$

$$\int_0^\infty f(x)\,\mathrm{d}x = 100\ , \tag{20}$$

where $f(x)$ is the PDF [5,12].

The empirical suggestion that the income distribution can be modeled by the TD comes from the fact that when $F(x)$ is obtained from income data and plotted in a log–log scale, its functional curve decreases as the income x increases. In addition, the general shape of the empirical CCDF function, particularly its "belly", is analogous to the behavior of e_q^{-x} for $q > 1$ when plotted in a log–log scale [see Ref. [23], p. 40, Fig. 3.4]. Besides, as mentioned above the Tsallis functions have power-law like behavior for high income values, agreeing then with the Pareto power-law. Taking together these observations into account, Ref. [20] advanced the following description for the individual income distribution,

$$F(x) = A\mathrm{e}_q^{-Bx}\ , \tag{21}$$

where A and B are positive parameters. Since we have a boundary condition in the form of $F(0) = 100$, the expressions above implies that $A = 100$. Hence, Eq. (21) may be rewritten as follows,

$$F(x) = 100\,\mathrm{e}_q^{-Bx}\ . \tag{22}$$

Note that a cursory examination of Eq. (22) does not present any obvious evidence of a periodic behavior, or that the complexification of q will disclose any oscillatory feature, since the q-parameter is just an index. We shall show below that allowing q to become complex will expose such features.

3.4. Periodic behavior of income distribution function

Let us start with the definition (3) in order to rewrite Eq. (22) as below,

$$F(x) \;=\; 100\left[1 \;+\; (1-q)(-Bx)\right]^{1/(1-q)}\ . \tag{23}$$

This equation can also be expressed as a function of m using Eq. (13),

$$F(x) \;=\; 100\left[1 \;+\; \frac{B}{m}x\right]^{-m}. \tag{24}$$

Following the complexification suggested in Eqs. (14) to (17), Eq. (24) may be written as below,

$$F(x) \;=\; 100\left[a(x) \;+\; \mathrm{i}b(x)\right]^{-(m'+\mathrm{i}m'')}, \tag{25}$$

where

$$a(x) \;=\; 1 \;+\; \frac{B\,m'}{|m|^2}\,x\ , \tag{26}$$

$$b(x) \;=\; -\,\frac{B\,m''}{|m|^2}\,x\ . \tag{27}$$

Expanding the terms in Eq. (25) results in the following expression,

$$F(x) \;=\; 100\,\mathcal{A}(x)\,\mathrm{e}^{-\mathrm{i}\omega(x)} \;=\; 100\,\mathcal{A}(x)\left[\cos\omega(x) \;-\; \mathrm{i}\,\sin\omega(x)\right], \tag{28}$$

where

$$\begin{cases} \mathcal{A}(x) \;=\; \exp\left[m''\varphi(x) - m'\theta(x)\right],\ \text{(a)} \\ \omega(x) \;=\; m'\varphi(x) + m''\theta(x), \qquad \text{(b)} \end{cases} \tag{29}$$

and

$$\begin{cases} \varphi(x) \;=\; \tan^{-1}\left[\dfrac{-B\,m''x}{\left(|m|^2 + B\,m'x\right)}\right],\ \text{(a)} \\ \theta(x) \;=\; \ln\sqrt{1 + \dfrac{Bx}{|m|^2}\,(Bx + 2m')}\ .\text{(b)} \end{cases} \tag{30}$$

We have to remember that B is always positive. Now, if we consider only the real part of Eq. (28), the original description of the income data in terms of the TD, as given by Eq. (22), turns out to be written as follows,

$$F(x) = 100\,\mathrm{e}_q^{-Bx} = 100\,\mathcal{A}(x)\left|\,\cos\omega(x)\right|\ , \tag{31}$$

where the absolute value of $\cos\omega(x)$ guarantees that we will always have the empirically obtained positive values for $F(x)$.

This result shows clearly that the periodic behavior empirically observed by Ref. [20] in the income data seems to be described by the expression above. Or that the observed oscillatory behavior can at least be expected, since it is built in the TD. Note that the original q-parameter is present in both $\mathcal{A}(x)$ and $\omega(x)$ through m' and m'', which are the respective real and imaginary parts of q.

The fact that the individual income distribution function can be written in the standard exponential form given by Eq. (28) shows us that we can define a periodic function where $\mathcal{A}(x)$ can be seen as the amplitude of this oscillatory motion and $\omega(x)$ its angular frequency. Considering the q-parameter only by a real term hides this periodic behavior. The complexification adds new components such as the kind derived above so that it is able of revealing a periodic behavior that does appear in the empirical curves obtained from the income distribution data.

The issue of always considering the q parameter as an entirely complex number, namely with both real and imaginary parts, is an open question although we believe that it depends on the problem we are dealing with. We can imagine an analogy of this feature with the dual characteristic of the electron in quantum physics, since it has a wave-particle duality behavioral feature that depends on the experience that we are analyzing. In this way we could talk about a q-duality, which would deserves alternative interpretations and could certainly be a target of further research.

4. Numerical application

Our problem now is to analyze the q-complex version of Eq. (31) concerning income distribution real data through a numerical example. The aim is to calculate the values of m' and m'' from Eq. (14), that is, to obtain for each empirical income value x_k given in the tables of Ref. [20] a pair $m_k{}'$ and $m_k{}''$.

Let us begin by rewriting Eq. (30)(a) and (b) to conform with discrete real numerical data, as follows,

$$\varphi_k = \varphi(x_k) = \arctan\left(\frac{-Bm_k{}''x_k}{m_k{}'^2 + m_k{}''^2 + Bm_k{}'x_k}\right), \tag{32}$$

$$\theta_k = \theta(x_k) = \ln\sqrt{1 + \frac{Bx_k\left(Bx_k + 2m_k{}'\right)}{m_k{}'^2 + m_k{}''^2}}\,, \tag{33}$$

where x_k is the kth income data point as produced by Ref. [20]. From Eq. (29)(a) we can write the following expression for the kth tabled income data,

$$\mathcal{A}_k = \mathcal{A}(x_k) = \exp\left[m_k{}'' \arctan\left(\frac{-Bm_k{}''x_k}{m_k{}'^2 + m_k{}''^2 + Bm_k{}'x_k}\right) - m_k{}' \ln\sqrt{1 + \frac{Bx_k\left(Bx_k + 2m_k{}'\right)}{m_k{}'^2 + m_k{}''^2}}\right]. \tag{34}$$

Similarly, Eq. (29)(b) allows us to write the expression below,

$$\omega_k = \omega(x_k) = m_k{}' \arctan\left[\frac{-Bm_k{}''x_k}{m_k{}'^2 + m_k{}''^2 + Bm_k{}'x_k}\right] + m_k{}'' \ln\sqrt{1 + \frac{Bx_k\left(Bx_k + 2m_k{}'\right)}{m_k{}'^2 + m_k{}''^2}}\,. \tag{35}$$

The aim is to solve Eq. (31) numerically. To accomplish this task, let us write it in the following numerically discrete form,

$$F(x_k) = 100\,\mathcal{A}(x_k)\left|\cos\omega(x_k)\right|, \qquad \Longrightarrow \qquad F_k = 100\,\mathcal{A}_k\left|\cos\omega_k\right|. \tag{36}$$

For the kth income one needs to calculate both $m_k{}'$ and $m_k{}''$. From Eqs. (15) to (17) we have that

$$|m|^2 = \frac{1}{|q-1|^2} = m_k{}'^2 + m_k{}''^2, \tag{37}$$

which allows us to write the expression below,

$$m_k{}'' = \pm\frac{\sqrt{1 - m_k{}'^2\,|q-1|^2}}{q-1}\,. \tag{38}$$

Considering that there are two solutions for $m_k{}''$ due to the square root, substituting Eq. (38) into Eqs. (34) and (35) we respectively obtain the results below,

$$\mathcal{A}_k = \exp\left\{-\frac{\sqrt{1 - m_k{}'^2|q-1|^2}}{q-1}\arctan\left[\frac{Bx_k(q-1)\sqrt{1 - m_k{}'^2|q-1|^2}}{1 + Bm_k{}'x_k(q-1)^2}\right]\right.$$

$$- m_k{}' \ln \sqrt{1 + Bx_k(q-1)^2\Big(Bx_k + 2m_k{}'\Big)} \Bigg\}, \tag{39}$$

and

$$\omega_k = - m_k{}' \arctan\left[\frac{Bx_k(q-1)\sqrt{1-m_k{}'^2|q-1|^2}}{1+Bm_k{}'x_k(q-1)^2}\right] + \frac{\sqrt{1-m_k{}'^2|q-1|^2}}{q-1} \ln\sqrt{1 + Bx_k(q-1)^2\Big(Bx_k + 2m_k{}'\Big)}\, . \tag{40}$$

Finally, substituting both expressions above into Eq. (36) the result may be written as below,

$$\begin{aligned} F_k = {} & 100 \exp\Bigg\{ - \frac{\sqrt{1-m_k{}'^2|q-1|^2}}{q-1} \arctan\left[\frac{Bx_k|q-1|\sqrt{1-m_k{}'^2|q-1|^2}}{1+Bm_k{}'x_k|q-1|^2}\right] \\ & - m_k{}' \ln\sqrt{1+Bx_k|q-1|^2\Big(Bx_k+2m_k{}'\Big)}\Bigg\} \\ & \times \Bigg|\cos\Bigg\{ - m_k{}' \arctan\left[\frac{Bx_k|q-1|\sqrt{1-m_k{}'^2|q-1|^2}}{1+Bm_k{}'x_k|q-1|^2}\right] \\ & + \frac{\sqrt{1-m_k{}'^2|q-1|^2}}{q-1} \ln\sqrt{1+Bx_k|q-1|^2\Big(Bx_k+2m_k{}'\Big)}\Bigg\}\Bigg|\, . \end{aligned} \tag{41}$$

Ref. [20] used income data of Brazil to obtain the distribution F_k in terms of income values x_k for each year in the observed time window and then fitted the TD to the empirical distribution in order to find both parameters q and B for all k-values in a given year. Once this was done, knowing the values of $[q, B]$ in that given year one can solve Eq. (41) numerically for each data pair $[F_k, x_k]$ to finally obtain the unknown quantity $m_k{}'$.

The methodology described above can then be used to test the whole complexification procedure. One starts by choosing the dataset of a single year as provided by Ref. [20], calculate $m_k{}'$ for each data point using the empirical values $[F_k, x_k, q, B]$ in the chosen year to find the roots of the expression (41), substitute the results $[m_k{}', x_k, q, B]$ back into Eq. (41) to recover the distribution F_k and then compare the original distribution with the recovered one. If the recovered points follow closely the original empirical distribution, including the observed oscillation, that would demonstrate that the complexification procedure above really discloses the oscillations present in the distribution.

Fig. 1 shows the results obtained with this approach using the empirical CCDF for the income data of Brazil in the year 2011. One can clearly see that the recovered distribution does follow closely the empirical one, including the more pronounced oscillation at the tail due to increasing values of the amplitude $\mathcal{A}_k$. These results could not bettered due to high non-linearity of Eq. (41), since calculating the root of this expression resulted in strong numerical fluctuations due to catastrophic loss of significant digits. Reducing these fluctuations required the use of no less than 20 digits in the numerical evaluation. Using more than 20 digits did not improve the results because the parameters $[q, B]$ were fitted to the data with up to 3 digits only.

The results shown in Fig. 1 could, perhaps, be improved if the numerical fluctuations due to the loss of significant digits were to be somehow reduced. One possibility for doing that would be by means of recalculating both parameters q and B with at least 20 digits, since they were both originally obtained with only 3 digits. But, the task of reducing these fluctuations is beyond the scope of this paper because our aim here is just to show that allowing the parameter q of the TD to become complex results in revealing the oscillatory nature of the distribution, as demonstrated by the graphs of Fig. 1.

5. Conclusions

Since the Pareto's work the study of the income distribution of the whole population has been a target of economic experts, and for some time now, to the econophysics literature. Despite the relative great number of parameters used in several functions to fit the income data, from three to five in some cases, some mathematical approaches have thrived in the description of the entire data range. In this work we have used Tsallis' non-extensive point of view with a complexification mode where its q-parameter is represented by a complex number. The objective here was to use this complexification in order to justify analytically the results obtained in Ref. [20], which show a periodic behavior in the income data.

As shown above, in doing this we, however, increased the number of parameters required to fit the data, from the original two to three. Although increasing the number of unknown parameters in a problem is a bad procedure, in our case the complexification has disclosed an extra behavior of the income distribution in the form of a periodic motion in the income distribution, motion which was already present in several studies of income distribution, but was only explicitly

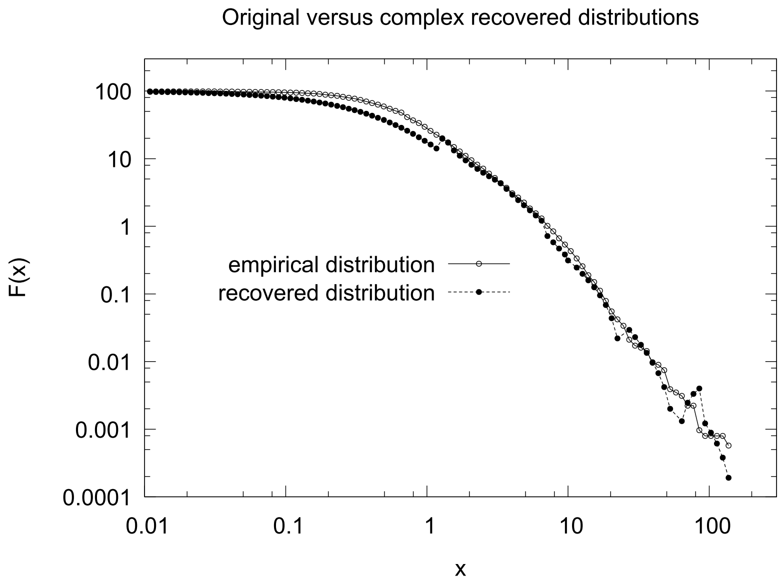

Fig. 1. This plot presents the CCDF of Brazil for the year 2011 with $q = 1.265$ obtained with the empirical data (open circles) and the recovered ones (filled circles) obtained by means of finding the roots of Eq. (41). Eqs. (12) and (13) are the parameter transformation from q to m, and although q has a single value for the whole distribution, the complexification process produces real and imaginary parts *for each point of the distribution*. So, as the distribution is made of k points there will be k values for complex m, as shown by Eqs. (37), (38) and (41). Therefore, the k points produced by the solution of Eq. (41) are shown in this plot. Since Eq. (41) is actually an expanded form of Eq. (36), and that it has an oscillatory term coupled with an amplitude term, then, once the k values of F_k are found by Eq. (41), the oscillatory nature of the distribution is reproduced by the fitting. However, due to its high non-linearity and the fact that it has several subtractions of two values close to unit, the numerical problem suffered from strong numerical instability due to the catastrophic loss of significant digits. This instability was significantly ameliorated by the use of 20 digits during the numerical evaluations of Eq. (41). Despite the instability, the oscillatory nature of the distribution is clearly visible in the recovered points, meaning that allowing the parameter q of the TD to become complex results in revealing the oscillatory nature of the distribution, especially at its tail.

acknowledged by Soares et al. [20]. By interpreting this oscillatory motion of the income distribution as a typical periodic motion allowed us to define commonly used oscillatory parameters such as amplitude and angular frequency.

The analytical procedure developed here was tested against real data, in this case the income distribution data of Brazil in 2011, and the numerical results showed that allowing for a complex q parameter results in revealing the oscillatory nature of the distribution, especially at its tail. So, from the results obtained here we can say that "the highs and lows" of the yearly income distribution samples are an expected behavior, confirmed using the complex form of the Tsallis q-parameter. However, it not all clear that this periodic oscillation could lead us to the understanding of the real nature of the q-parameter, namely, if it is complex or not depends on the features of the problem we are dealing with.

Finally, this work shows us a glimpse of the task ahead as far as the empirical studies of income distribution are concerned. The data give us $F(x)$, x, B, q and m, so the empirical task is to determine both m' and m'' from the data in order to end up with only three parameters required to characterize the whole yearly distribution, oscillatory feature included, namely B, m' and m''.

Acknowledgments

E.M.C.A. thanks the Brazilian Federal Agency CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), a federal scientific support agency, for partial financial support, Grants numbers 302155/2015-5 and 406894/2018-3. M.B.R. acknowledges partial financial support from the Rio de Janeiro State Scientific Funding Agency (FAPERJ).

References

[1] V. Pareto, Cours d'Économie Politique, Lausanne, 1897.
[2] B.B. Mandelbrot, The Fractal Geometry of Nature, Freeman, San Francisco, 1982.
[3] N.C. Kakwani, Income Inequality and Poverty, Oxford University Press, 1980.
[4] N.J. Moura Jr., M.B. Ribeiro, Evidence for the Gompertz curve in the income distribution of Brazil 1978-2005, Eur. Phys. J. B 67 (2009) 101–120, arXiv:08122664.
[5] F. Chami Figueira, N.J. Moura Jr., M.B. Ribeiro, The Gompertz-Pareto income distribution, Physica A 390 (2011) 689–698, arXiv:10101994.
[6] A. Drăgulescu, V.M. Yakovenko, Evidence for the exponential distribution of income in the USA, Eur. Phys. J. B 20 (2001) 585, arXiv:cond-mat/0008305.
[7] J.C. Ferrero, The statistical distribution of money and the rate of money transference, Physica A 341 (2004) 575–585.

[8] A. Christian Silva, Applications of Physics to Finance and Economics: Returns, Trading Activity and Income (Ph.D. thesis), University of Maryland, 2005, arXiv:physics/0507022.
[9] V.M. Yakovenko, J.B. Rosser, Colloquium: Statistical mechanics of money, wealth, and income, Rev. Modern Phys. 81 (2009) 1703–1725, arXiv:09051518.
[10] R. Coelho, P. Richmond, J. Barry, S. Hutzler, Double power laws in income and wealth distributions, Physica A 387 (2008) 3847–3851, arXiv:07100917.
[11] A. Banerjee, V.M. Yakovenko, Universal patterns of inequality, New J. Phys. 12 (2010) 075032, arXiv:09124898.
[12] N.J. Moura Jr., M.B. Ribeiro, Testing the goodwin growth-cycle macroeconomic dynamics in Brazil, Physica A 392 (2013) 2088–2103, arXiv:13011090.
[13] Econophys-Kolkata I Workshop, in: A. Chatterjee, S. Yarlagadda, B.K. Chakrabarti (Eds.), Econophysics of Wealth Distributions, Springer, 2005.
[14] B.K. Chakrabarti, A. Chakraborti, S.R. Chakravarty, A. Chatterjee, Econophysics of Income and Wealth Distributions, Cambridge University Press, 2013.
[15] E.P. Borges, Empirical nonextensive laws for the county distribution of total personal income and gross domestic product, Physica A 334 (2004) 255.
[16] J.C. Ferrero, An statistical analysis of stratification and inequity in the income distribution, Eur. Phys. J. B 80 (2011) 255.
[17] J.C. Ferrero, The Monomodal, Polymodal, Equilibrium and Nonequilibrium Distribution of Money, In [13], p. 159, 2005.
[18] C. Tsallis, Possible generalization of Boltzmann–Gibbs statistics, J. Stat. Phys. 52 (1988) 479; Nonadditive entropy: the concept and its use, Eur. Phys. J. A 40 (2009) 257.
[19] V. Schwämmle, C. Tsallis, Two-parameter generalization of the logarithm and exponential functions and Boltzmann–Gibbs-Shannon entropy, J. Math. Phys. 48 (2007) 113301.
[20] A.D. Soares, N.J. Moura Jr., M.B. Ribeiro, Tsallis statistics in the income distribution of Brazil, Chaos Solitons Fractas 88 (2016) 158.
[21] G. Wilk, Z. Włodraczyk, Tsallis distribution with complex nonextensivity parameter q, Physica A 413 (2014) 53.
[22] C. Tsallis, What are the numbers that experiments provide? Química Nova 17 (1994) 468–471.
[23] C. Tsallis, Introduction to Nonextensive Statistical Mechanics, Springer, 2009.
[24] F. Clementi, M. Gallegati, G. Kaniadakis, κ-Generalised statistics in personal income distribution, Eur. Phys. J. B 57 (2007) 187–193, arXiv:physics/0607293.
[25] F. Clementi, T. Di Matteo, M. Gallegati, G. Kaniadakis, The κ-generalised distribution: a new descriptive model for the size distribution of incomes, Physica A 387 (2008) 3201–3208, arXiv:07103645.
[26] F. Clementi, M. Gallegati, G. Kaniadakis, A κ-generalized statistical mechanics approach to income analysis, J. Stat. Mech. (2009) P02037, arXiv:09020075.
[27] F. Clementi, M. Gallegati, G. Kaniadakis, A generalized statistical model for the size distribution of wealth, J. Stat. Mech. (2012) P12006, arXiv:12094787.
[28] T. Yamano, Some properties of q-logarithm and q-exponential functions in Tsallis statistics, Physica A 305 (2002) 486–496.
[29] T.S. Biró, G.G. Barnaföldi, P. Ván, Quark-gluon plasma connected to finite heat bath, Eur. Phys. J. A 49 (2013) 110.
[30] J.E.K. Schawe, A comparison of different evaluation methods in modulated temperature DSC, Thermochim. Acta 260 (1995) 1.
[31] J.-L. Garden, Simple derivation of the frequency dependent complex heat capacity, Thermochim. Acta 460 (1995) 85.
[32] Y. Huang, H. Saleur, C. Sammis, D. Sornette, Precursors aftershocks criticality and self-organized criticality, Europhys. Lett. 41 (1998) 43.
[33] H. Saleur, C.G. Sammis, D. Sornette, Discrete scale invariance, complex fractal dimensions, and log-periodic fluctuations in seismicity, J. Geophys. Res. 101 (1996) 17661.
[34] A. Krawiecki, K. Kacperski, S. Matyjaskiewicz, J.A. Holyst, Log-periodic oscillations and noise-free stochastic multiresonance due to self-similarity of fractals, Chaos Solitons Fractals 18 (2003) 89.
[35] J. Bernasconi, W.R. Schneider, Diffusion in random one-dimensional systems, J. Stat. Phys. 30 (1983) 355.
[36] D. Stauffer, D. Sornette, Log-periodic oscillations for biased diffusion on random lattice, Physica A 252 (1998) 271.
[37] D. Stauffer, New simulations on old biased diffusion, Physica A 266 (1999) 35.
[38] B. Kutnjak-Urbanc, S. Zapperi, S. Milosevic, H.E. Stanley, Sandpile model on the Sierpinski gasket fractal, Phys. Rev. E 54 (1996) 272.
[39] R.F.S. Andrade, Detailed characterization of log-periodic oscillations for an aperiodic Ising model, Phys. Rev. E 61 (2000) 7196.
[40] M.A. Bab, G. Fabricius, E.V. Albano, Critical behavior of an Ising system on the Sierpinski carpet: A short-time dynamics study, Phys. Rev. E 71 (2005) 036139.
[41] H. Saleur, D. Sornette, Complex exponents and log-periodic corrections in frustrated systems, J. Physique I 6 (1996) 327.
[42] C. Tsallis, L.R. da Silva, R.S. Mendes, R.O. Vallejos, A.M. Mariz, Specific heat anomalies associated with Cantor-set energy spectra, Phys. Rev. E 56 (1997) R4922.
[43] R.O. Vallejos, R.S. Mendes, L.R. da Silva, C. Tsallis, Connection between energy spectrum, self-similarity, and specific heat log-periodicity, Phys. Rev. E 58 (1998) 1346.
[44] D. Sornette, A. Johansen, A. Arneodo, J.-F. Muzy, H. Saleur, Complex fractal dimensions describe the hierarchical structure of diffusion-limited-aggregate clusters, Phys. Rev. Lett. 76 (1996) 251.
[45] Y. Huang, G. Ouillon, H. Saleur, D. Sornette, Spontaneous generation of discrete scale invariance in growth models, Phys. Rev. E 55 (1997) 6433.
[46] D. Sornette, A. Johansen, J.-P. Bouchaud, Stock market crashes, precursors and replicas, J. Phys. I 6 (1996) 167.
[47] N. Vanderwalle, P. Boveroux, A. Minguet, M. Ausloos, The crash of October 1987 seen as a phase transition: amplitude and universality, Physica A 255 (1998) 201.
[48] N. Vanderwalle, M. Ausloos, How the financial crash of October 1997 could have been predicted, Eur. J. Phys. B 4 (1998) 139.
[49] N. Vanderwalle, M. Ausloos, P. Boveroux, A. Minguet, Visualizing the log-periodic pattern before crashes, Eur. J. Phys. B 9 (1999) 355.
[50] J.H. Wosnitza, J. Leker, Can log-periodic power law structures arise from random fluctuations? Physica A 401 (2014) 228.
[51] F.A.B.F. de Moura, U. Tirnakli, M.L. Lyra, Convergence to the critical attractor of dissipative maps: Log-periodic oscillations, fractality, and nonextensivity, Phys. Rev. E 62 (2000) 6361.