P H Y S I C A L   R E V I E W   L E T T E R S

Temporal Scaling at Feigenbaum Points and Nonextensive Thermodynamics

Peter Grassberger

John-von-Neumann Institute for Computing, Forschungszentrum Jülich, D-52425 Jülich, Germany
Perimeter Institute for Theoretical Physics, Waterloo, Canada, N2L 2Y5

We show that recent claims for the nonstationary behavior of the logistic map at the Feigenbaum point based on nonextensive thermodynamics are either incorrect or can be easily deduced from well-known properties of the Feigenbaum attractor. In particular, there is no generalized Pesin identity for this system, the existing attempts at proofs being based on misconceptions about basic notions of ergodic theory. In deriving several new scaling laws of the Feigenbaum attractor, thorough use is made of its detailed structure, but there is no obvious connection to nonextensive thermodynamics.

During the last decade a vast body of literature has appeared on a new "nonextensive thermodynamics" (NET), which uses a maximum entropy principle with the Shannon entropy replaced by the Havrda-Charvat [1] (Tsallis) entropy [2–5]. Several versions of NET were proposed by Tsallis and others [3], in order to avoid inconsistencies. But as shown most forcefully by Nauenberg [6], none of these versions is consistent with thermodynamics, at least for equilibrium systems.

For nonequilibrium systems the situation is not quite so simple, as the standard maximum entropy principle is usually not applicable. As shown by Jaynes [7], the formalism of statistical physics can be obtained by Occam's razor: If all knowledge can be formulated in terms of constraints, then the only rationally justifiable ansatz for the probability distribution is the one which maximizes Shannon entropy (which has to be replaced by Kullback-Leibler information, if some knowledge existed prior to these constraints). The main reason why standard thermodynamics cannot be applied to most nonequilibrium systems is that prior knowledge cannot be cast into the form of a few constraints. But there is no reason for abandoning an information theoretic interpretation of entropy, and Shannon (Kullback-Leibler) entropy is the only consistent probabilistic measure of (relative) information.

The main reason why NET is still vigorously pursued is, it seems, the claim that it is able to make striking predictions that could not be made within a more conventional framework. Typically, these are for nonequilibrium phenomena with distributions showing power laws with heavy tails. In deriving these distributions, not only is Shannon entropy replaced arbitrarily by Tsallis entropy. Also the constraints are modified in a way which has no clear rational motivation—except that one arrives thereby at expressions more easily handled.

A careful study of most examples where NET was supposedly successful shows that the success is much less clear than claimed. In a later publication, we will substantiate this further by discussing several such examples. In the present Letter, we discuss in depth just one single example, which has been treated in several papers [8–23], and which was claimed to show the success of NET in a particularly clear way.

This example is the nonstationary behavior of the Feigenbaum attractor [24]. More precisely, one can study: (1) sensitive dependence on initial conditions (both for finite and infinitesimal perturbations, both on the attractor and in its basin of attraction); (2) scaling of different dynamical (Shannon, Renyi, Tsallis) entropies of various ensembles of trajectories with their time length T; (3) convergence of a typical trajectory (with random initial condition in its vicinity) to the attractor; (4) scaling of "static" (i.e., Boltzmann-Gibbs) entropies of an ensemble of points with time.

Most of these problems have been discussed since the early 1980's [25–30]. Also, it had been realized from the very beginning that some of these questions are subtle due to large fluctuations ("multifractality" of the Feigenbaum attractor). In the first papers based on NET [8–10] this was missed, leading to wrong claims that a single nonextensivity parameter could describe scaling at the onset of chaos. Although this was corrected recently, we shall see that the effect of fluctuations is still not fully appreciated in some of these papers [20,23].

In the following we give theoretical arguments only for the Feigenbaum map [24] $g(x)$ defined by $g[g(x)] = -\alpha^{-1}g(\alpha x)$ and $g(x) = 1 - cx^2 + O(x^4)$ for $x \to 0$, but we use the logistic map $a - x^2$ with $a = 1.401\,155\,189\ldots$ for numerical calculations. Problems of universality have been discussed, e.g., in [24,26].

Let us first discuss the dependence on infinitesimal changes in the initial conditions, i.e., the behavior of $\Lambda_n(x_0) = \ln|dx_n/dx_0|$, where $x_{i+1} = g(x_i)$. As pointed out in [26], $\Lambda_n(x_0)$ fluctuates very strongly with n and x_0. Different ways of averaging over x_0 give, therefore, rise to different scalings with n. If we take arithmetic averages over Λ_n (i.e., geometric averages over $|dx_n/dx_0|$), we get

$$\int dx_0 w_0(x_0) \ln|dx_n/dx_0| \sim \gamma \ln n \qquad (1)$$

with $\gamma = 0.599 \pm 0.003$ for all smooth initial distributions $w_0(x)$, at least when we also do an additional averaging

140601-1

over n to damp out remaining oscillations [31]. On the other hand, as proven rigorously in [26], arithmetic averages over $|dx_n/dx_0|$ give

$$\int dx_0 w_0(x_0)|dx_n/dx_0| \sim n^{\text{const} \ln n}. \quad (2)$$

Without averaging over x_0 one can of course obtain completely different behavior. For $x_0 = 1$, e.g., one finds $|dx_n/dx_0| \to \text{const}$ for $n = 2^k$, $k \to \infty$, while

$$|dx_n/dx_0| = \alpha^k = (n+1)^{\log_2 \alpha} \quad (3)$$

(exactly) for $n = 2^k - 1$ [32]. Tsallis *et al.* prefer to write Eq. (3) as $|dx_n/dx_0| = [1 + \lambda(1-q)n]^{1/(1-q)}$ with $\lambda = 1/(1-q) = \log_2 \alpha$, and call the right-hand side a q exponential.

Notice that Eq. (3) is *not* a scaling law, since it holds only for special values of n. But one obtains a scaling law by taking averages over n (see Fig. 1),

$$\bar{\Lambda}_n(x_0) \equiv n^{-1} \sum_{i=1}^{n} \Lambda_i(x_0) \sim \beta \ln n \quad \text{for } x_0 = 1 \quad (4)$$

with $\beta = 0.713\,393\,80\dots$. Notice that the constants in Eqs. (1), (3), and (4), are not directly related.

Defining $\bar{\Lambda}_n(x)$ for any x on the attractor as in Eq. (4) gives the natural generalization of the Lyapunov exponent. It fluctuates strongly with x [28,31] and for many applications like Pesin's identity [33] one needs averages over x with respect to the natural measure $\mu(x)$. In the present case, one can easily see that the average

$$\langle \bar{\Lambda}_n \rangle \equiv \int d\mu(x) \bar{\Lambda}_n(x) \quad (5)$$

is identically zero for all n [35], which indicates that there cannot exist a close analogy to Pesin's identity. The identities suggested and verified in [8,20,36] are trivial as the authors started with an initial distribution $w_0(x)$ narrowly localized around x_0 and followed the evolution only up to values of n for which $w_n(x)$ is still smooth and described by the flow linearized around the trajectory starting at x_0 [37]. In addition, they considered, instead of the Kolmogorov-

Sinai (KS) entropy, just the difference $S_n - S_0$ between two static "Boltzmann-Gibbs" entropies

$$S_n = -\int dx\, w_n(x) \ln w_n(x). \quad (6)$$

In this case one has of course $S_n - S_0 = \Lambda_n(x_0)$, but this has no connection to any (generalized) Pesin identity.

Eqs. (3) and (4) apply to trajectories starting *on the attractor*. In order to understand the origin of Eqs. (1) and (2) one has to study how trajectories starting in its vicinity approach the attractor. For this one has to use the detailed triadic Cantor structure of the Feigenbaum attractor [24]. Associated to this structure is a set of open disjoint intervals $I_{k,i}$, the closure of which covers the interval $[-1/\alpha, 1]$ which contains the attractor [26]. $I(0, 1)$ is the hole cut out from the middle of the Cantor set in the first stage of construction; $I(1, 1)$ and $I(1, 2)$ are the second generation holes, $I(2, 1)$ to $I(2, 4)$ are the holes cut out in the third step, etc. Each $I_{k,i}$ contains exactly one point on the instable periodic orbit of period 2^k. Let us define $I_k = \bigcup_i I_{k,i}$; i.e., I_k is all that is cut out during the kth step. The evolution of a typical trajectory can then be viewed as a tumbling through the I_k's, with k never decreasing with n. The average increase of k is for large n exactly given by

$$\langle k \rangle = \text{const} + \log_2 n. \quad (7)$$

As shown in [26], Eq. (1) follows (up to the precise value of γ) from Eq. (7) and from the scaling behavior proven by Feigenbaum.

In [20] the authors studied another average over $\Lambda_n(x)$, in between Eqs. (1) and (2),

$$\xi_q(n) \equiv \int dx_0 w_0(x_0)|dx_n/dx_0|^{1-q}. \quad (8)$$

From straightforward simulations they concluded that $\xi_q(n)$ becomes (asymptotically) linear in n, $\xi_q(n) \sim n$, for $q = 0.36$ and called this value $q_{\text{sen}}^{\text{av}}$. But neither analytic nor numerical estimates of $\xi_q(n)$ seem easy. A direct numerical estimate as in [20] is prone to large errors, since the integrand of Eq. (8) [with constant $w_0(x_0)$ and for $q = q_{\text{sen}}^{\text{av}}$] is very sharply peaked near the unstable periodic orbits of small periods. Choosing initial values x_0 at random one will miss these peaks, unless one has extremely high statistics. But it is not clear anyhow why Eq. (8) should be of interest. In contrast to what its name suggests, $q_{\text{sen}}^{\text{av}}$ does not measure *the* average sensitivity to initial conditions but just one particular average.

Another consequence of Eq. (7) is that the (geometrically) average distance from the attractor of a point starting randomly in $[-1/\alpha, 1]$ decreases as [26]

$$d_n \sim n^{-1/D_1} \quad (9)$$

where $D_1 = 0.517\,097\dots$ is the information dimension of the Feigenbaum attractor.

The same argument can also be immediately used to derive the scaling of the Boltzmann-Gibbs entropy with n,
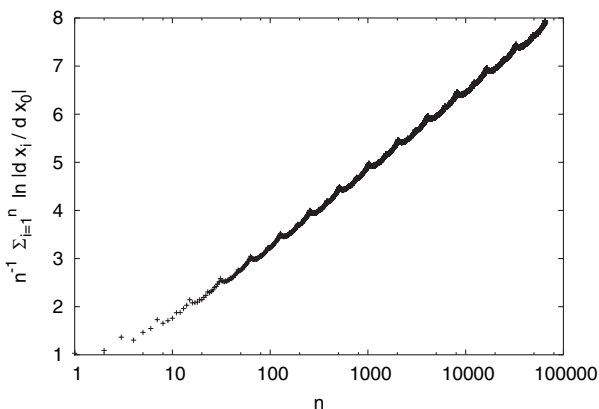

FIG. 1. Scaling of the time averaged local sensitivity exponents, for trajectories starting at $x_0 = 1$.

after starting with a random ensemble [23]. Assume we have N_n points distributed randomly with respect to $w_n(x)$. The distance between point $x_n^{(j)}$ and its nearest neighbor is called r_j. The Kozachenko-Leonenko estimator [38] for differential entropies gives then for large N_n

$$S_n = \text{const} + \ln N_n + \sum_{i=j}^{N_n} \ln r_j. \quad (10)$$

To apply this for the present problem, we choose furthermore $N_n = nN_0$ so that N_n scales as the number of intervals $I_{k,i}$ with $k = \text{const} + \log_2 n$. We notice also that all $I_{k,i}$ with $i = 1, \ldots, 2^k$ have roughly the same statistical weight. Then the distance to the nearest neighbor scales as the size of the interval in which the point is located, $r_j \in |I_{k,i}|$ with $x_j \in I_{k,i}$, and it follows that

$$S_n = \text{const} - \frac{1 - D_1}{D_1} \log_2 n. \quad (11)$$

Data obtained by spraying N points uniformly onto $[a - a^2, a]$, letting them evolve according to the critical logistic map, and estimating S_n by binning this interval into M bins, are shown in Fig. 2. Here we used $N > 2 \times 10^9$, which is largely sufficient for convergence (we used the entropy estimator corrections given in [39]). Convergence with the number of bins is much slower. We see very clear changes as we increase M from $\approx 10^7$ to $\approx 3 \times 10^8$. This explains also the small remaining discrepancy.

Although S_n had been discussed as an interesting quantity in the NET literature [11,17,22,23], we are not aware of a previous estimate. Instead, several authors [11,17,23] have studied a different quantity W_n, which they supposed to have the same scaling. Instead of taking a sum over all nonempty bins with weights $p_i \log(1/p_i)$, W_n is just the number of nonempty bins. Scaling of this quantity is much more subtle—both theoretically and numerically. We
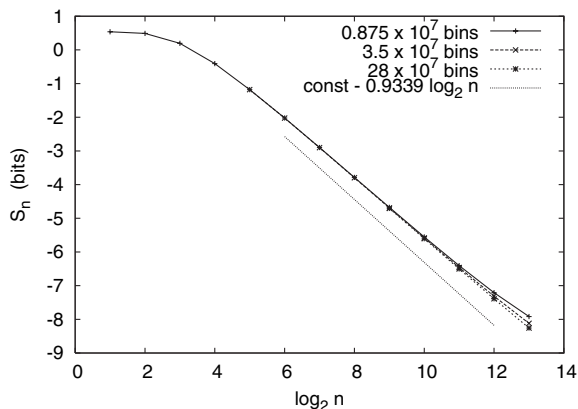

FIG. 2. Time dependence of the Boltzmann-Gibbs entropy S_n, i.e., of the differential Shannon entropy of the x distribution, starting with a uniform distribution over the interval $[a - a^2, a]$. The three curves correspond to different binnings. The lines connecting the data points are just to guide the eye, and omit the log-normal oscillations. The straight line indicates the theoretically predicted slope.

should expect larger finite n, finite M, and finite N corrections for W_n than for S_n, but when n, M, and N are sufficiently large [40], we can apply basically the same reasoning to W_n as we did for S_n: It scales like the total length of all intervals $I_{k,i}$ with $k = \log_2 n$. Calculating the latter is easy and gives

$$W_n \sim n^{-0.800\,138\,194\ldots}. \quad (12)$$

The exponent measured in [11] was quite different (0.71 ± 0.01 instead of 0.80), but this is easily explained by the expected large corrections to scaling, which have even led to claims of nonuniversality in [23].

The KS entropy for a one-humped map with maximum at $x = 0$ is obtained from the sequences $s_n = [\text{sgn}(x_1), \ldots, \text{sgn}(x_n)]$, as

$$H_{\text{KS}} = \lim_{n \to \infty} n^{-1} H_n \quad (13)$$

with $H_n = -\sum_{s_n} p(s_n) \ln p(s_n)$. As shown in [27,30], for the Feigenbaum attractor one has $H_n = \ln n + O(1)$. Thus the KS entropy is zero, but there is a logarithmic increase of H_n. With the same methods one can prove *exactly* that all Renyi entropies $H^q = \lim_{n \to \infty} H_n^q$ with $H_n^q = (1 - q)^{-1} \ln \sum_{s_n} [p(s_n)]^q$ are equal for this problem, $\lim_{n \to \infty} H_n^q / H_n = 1$ for all q [41]. For Tsallis-type generalized KS entropies, this gives $K_n^q \equiv (1 - q)^{-1}(\exp[(1 - q)H_n^q] - 1) \sim n^{1-q}$, at variance with [14].

The q independence of the Renyi entropies is surprising in view of the multifractality of the attractor. For a chaotic attractor, the H_n^q are closely related, via generalizations of Pesin's identity, to qth moments of $\Lambda(x)$ [42]. This shows again that the onset of chaos is more subtle than expected in the NET literature.

In the first papers [8–11], it was supposed but never substantiated that the parameter q of NET can be obtained, also for the Feigenbaum map, by some maximum entropy principle. Although this was never withdrawn, q is now fixed such that one obtains linear time dependencies, when logarithms are replaced by q-deformed logarithms (see, e.g., [20]). The reason for this is not clear, since there is no need for time dependencies to be linear. But even worse, with the proliferation of different (although closely related) scaling laws, one obtains—for the single case of the Feigenbaum map—a rich zoo of different q values [4,20]. With the additional scaling laws found in the present Letter, there would be even more q's—unless one accepts at last that NET fails at least at the onset of chaos.

One reason for preferring "q exponentials" (i.e., generalized Pareto distributions [43]) over power laws could be that the former give correct deviations from pure power laws at small arguments. But in all cases studied in the present Letter [except Eq. (3)] the small-n limits are neither given by pure powers nor by q exponentials.

Finally, as a last remark: In [3,15,20] the Feigenbaum map was chosen as the prototype of a supposedly *weakly mixing* system. But it is not mixing at all.

———————

[1] J. Havrda and F. Charvat, Kybernetica **3**, 30 (1967).

[2] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).

[3] C. Tsallis, Chaos Solitons Fractals **13**, 371 (2002).

[4] C. Tsallis, cond-mat/0403012.

[5] For an up to date bibliography containing more than 1500 references, see http://tsallis.cat.cbpf.br/biblio.htm.

[6] M. Nauenberg, Phys. Rev. E **67**, 036114 (2003); **69**, 038102 (2004).

[7] E. T. Jaynes and G. L. Bretthorst, *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, England, 2003).

[8] C. Tsallis, A. R. Plastino, and W.-M. Zheng, Chaos Solitons Fractals **8**, 885 (1997).

[9] U. M. S. Costa, M. L. Lyra, A. R. Plastino, and C. Tsallis, Phys. Rev. E **56**, 245 (1997).

[10] M. L. Lyra and C. Tsallis, Phys. Rev. Lett. **80**, 53 (1998).

[11] F. A. B. F. de Moura, U. Tirnakli, and M. L. Lyra, Phys. Rev. E **62**, 6361 (2000).

[12] J. Yang and P. Grigolini, Phys. Lett. A **263**, 323 (1999).

[13] M. Buiatti, P. Grigolini, and L. Palatella, Physica A (Amsterdam) **268**, 214 (1999).

[14] S. Montangero, L. Franzoni, and P. Grigolini, Phys. Lett. A **285**, 81 (2001).

[15] V. Latora, M. Baranger, A. Rapisarda, and C. Tsallis, Phys. Lett. A **273**, 97 (2000).

[16] U. Tirnakli, G. F. J. Añaños, and C. Tsallis, Phys. Lett. A **289**, 51 (2001).

[17] E. P. Borges, C. Tsallis, G. F. J. Añaños, and P. M. C. de Oliveira, Phys. Rev. Lett. **89**, 254103 (2002).

[18] C. Tsallis, A. Rapisarda, V. Latora, and F. Baldovin, *Nonextensivity: From Low-Dimensional Maps to Hamiltonian Systems*, Springer Lecture Notes in Physics Vol. 602 (Springer, Berlin, 2002), p. 140.

[19] F. Baldovin and A. Robledo, Phys. Rev. E **66**, 045104(R) (2002); Europhys. Lett. **60**, 518 (2002).

[20] G. F. J. Añaños and C. Tsallis, Phys. Rev. Lett. **93**, 020601 (2004).

[21] R. Tonelli, G. Mezzorani, F. Meloni, M. Lissia, and M. Coraddu, cond-mat/0403360.

[22] M. Coraddu, F. Meloni, G. Mezzorani, and R. Tonelli, Physica A (Amsterdam) **340**, 234 (2004).

[23] R. Tonelli, report, 2005.

[24] M. J. Feigenbaum, J. Stat. Phys. **19**, 25 (1978); **21**, 669 (1979).

[25] B. A. Huberman and J. Rudnick, Phys. Rev. Lett. **45**, 154 (1980).

[26] P. Grassberger and M. Scheunert, J. Stat. Phys. **26**, 697 (1981).

[27] P. Grassberger, Int. J. Theor. Phys. **25**, 907 (1986).

[28] G. Anania and A. Politi, Europhys. Lett. **7**, 119 (1988).

[29] H. Hata *et al.*, Prog. Theor. Phys. **82**, 897 (1989).

[30] J. Freund *et al.*, Phys. Rev. E **54**, 5561 (1996).

[31] These oscillations and the ones seen in Fig. 1 are due to the Cantor structure of the Feigenbaum attractor and will not be discussed further.

[32] Equation (3) follows immediately from the Feigenbaum-Cvitanovic relation $\alpha g[g(x)] = -g(\alpha x)$ and from the well-known facts [24] that $g(1) = -1/\alpha$, $g'(1) = -\alpha$, and $g'(-1/\alpha) = 1$. The equation for $n = 2^k$ uses in addition the fact that $x_{(2^k-1)} = g^{(2^k-1)}(1)$ is close to zero, where $g(x)$ has its quadratic maximum. See also [19]. For Feigenbaum type maps $g_z(x)$ with maxima of order z, the same easy derivation gives $|dx_n/dx_0| = \alpha^{(z-1)k}$. For no obvious reasons, this relation was claimed to show the importance of NET in [10].

[33] For an invariant measure which is absolutely continuous in all expanding dimensions, Pesin's identity (see, e.g., [34]) relates the Kolmogorov-Sinai entropy H_{KS} to the Lyapunov spectrum via $H_{KS} = \sum_{\lambda_i > 0} \lambda_i$. For a measure with partial dimensions $D_i < 1$ in the expanding dimensions; this generalizes to $H_{KS} = \sum_{\lambda_i > 0} D_i \lambda_i$. The *latter* should be further generalized if one looks for a Pesin-type identity for the Feigenbaum map.

[34] E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge, England, 1993).

[35] The natural measure can be approximated as $d\mu(x)/dx \approx N^{-1} \sum_{i=N+1}^{2N} \delta[x - g^{(i)}(1)]$ for $N \gg 1$. Thus $\langle \bar{\Lambda}_n \rangle$ is given by the increase of the curve shown in Fig. 1 when the x coordinate increases by n, averaged over $[N+1, 2N]$ with $N \gg n$, and then divided by n. In the limit $n/N \to 0$ this obviously converges to zero.

[36] F. Baldovin and A. Robledo, Phys. Rev. E **69**, 045202(R) (2004).

[37] In [4,15,16], the authors first make a partition of the attractor basin into W equally long bins and take N starting points x_0^j, $j = 1 \ldots N$, in one of them. The spatial order-q Tsallis entropy after n iterations, estimated from the points x_n^j, is denoted as $S_q(n)$. Then the "generalized KS entropy" of order q is defined as

$$K_q = \lim_{n \to \infty} \lim_{W \to \infty} \lim_{N \to \infty} n^{-1} S_q(n). \qquad (14)$$

For $q \to 1$ this is claimed in [4,15,16] to coincide with the standard KS entropy, but in the latter the limits $n \to \infty$ and $W \to \infty$ are exchanged, the points are chosen randomly according to the natural invariant measure, and $S_1(n)$ should be replaced by the Shannon entropy for *trajectories* of length n.

[38] L. F. Kozachenko and N. N. Leonenko, Problems of Information Transmission (Engl Trans) **23**, 95 (1987).

[39] P. Grassberger, physics/0307138.

[40] In contrast to Eq. (11), where one has to take first $N \to \infty$, then $M \to \infty$, and at last $n \to \infty$, these limits have to be taken simultaneously in Eq. (12). Taking, e.g., $N \to \infty$ first, all W_n would be equal to W_0. Equation (12) is obtained with $M > n^{\log_2 \alpha}$ and taking $N/M \to \infty$ after $n \to \infty$. It relies on the fact that the distribution of k values for fixed n is extremely sharply cut off at small k [26].

[41] If we start on the attractor, and if $n = 2^k$ ($k = 1, 2, \ldots,$), then there are $3k$ equally probable s_n, giving $H_n = \ln(3n/2)$. For other values of n there two classes of allowed sequences s_n, with the sequences in the first class twice as probable as those in the second class.

[42] P. Grassberger *et al.*, J. Stat. Phys. **51**, 135 (1988).

[43] J. Pickands, Ann. Stat. **3**, 119 (1975).