Journal of Statistcal Physics, Vol. 69, Nos. 1/2, 1992

On the Phase Space Approach to Complexity

Hans C. Fogedby[1]

Received October 1, 1991; final April 28, 1992

We elaborate in some detail on a new phase space approach to complexity, due to Y.-C. Zhang. We show in particular that the connection between maximal complexity and power law noise or correlations can be derived from a simple variational principle. For a 1D signal we find $1/f$ noise, in accordance with Zhang.

KEY WORDS: Complexity; phase space; entropy; signals; probability distribution; coarse graining; reduction formula; characteristic function; Gaussian distribution; power law noise; variational principle; $1/f$ noise.

1. INTRODUCTION

The emergent properties in many natural and artificial structures, that is, properties not present in or deducible from the fundamental laws of interaction between the individual constituents, is a fascinating aspect of the behavior of interacting systems composed of many parts. In physics the classical example is that of second-order phase transitions in equilibrium systems,[2,3] where the system in question develops long-range spatiotemporal correlations. More recently, the case of fractal behavior in diffusion-limited aggregation (DLA)[4] and spatiotemporal long-range correlations in driven self-organized systems[5] are also examples of emergent properties in irreversible and driven systems. In computer science and physics modeling we also mention the asymptotic behavior of cellular automata as a case of emergent properties.[11]

In recent years there have been several attempts to introduce a quantitative measure for the emergent complex behavior of collective systems.[6-8,11,13,14] A measure of complexity K providing a global characterization distinguishing complex behavior, much like the pressure P or

[1] Institute of Physics, University of Aarhus, 8000 Aarhus C, Denmark.

0022-4715/92/1000-0411$06.50/0

temperature T, is a global characteristic for thermodynamic equilibrium systems. In analogy with the entropy, which takes its largest value for the equilibrium system subject to constraints, the measure K should thus assume its extremal, say largest, value for systems which are in an intermediate state between perfect order and complete disorder and which also exhibit long-range correlations in space and time. We thus expect an appropriate complexity measure to be maximal at the critical point of a second-order phase transition and in the case of self-organized critical behavior,[5] to mention two cases.

For equilibrium systems the concept of entropy is of central importance[3] and characterizes the "degree of order" in the system. Subjectively, the "degree of order" is, however, not the same as "degree of complexity" and the entropy does not distinguish the most complex state of the system and as such is not appropriate as a complexity measure. In fact, as alluded to, it seems that the somewhat elusive notion of complexity is partly subjective and related to the concept of meaning.[11] Also, as we shall discuss below, the perceived complexity of a collective system must be related to the measuring procedure and the degree of resolution.

More specifically, let us briefly comment on two recent approaches to complexity: Computational or algebraic complexity is a measure defined for a specific system or realization as the length of the minimal algorithm or code which leads to an exact copy of the structure when implemented on a computer.[11] Computational complexity is associated with a specific "string" or realization and is therefore not very suitable in a physics context, where we usually deal with an ensemble of systems and wish to introduce the notion of complexity as a probabilistic measure.

A complexity measure has also been introduced for general hierarchical discrete structures[6,7] with the important property of having its largest value for systems intermediate between the completely ordered and completely disordered states. The limitation of "hierarchical" complexity is, however, that it is a rather formal construct based on the generic tree structure of a hierarchical system and does not refer in any direct way to the spatiotemporal characteristics of the system under consideration.

In a recent paper Zhang[1] has proposed a new definition of complexity, here called phase space complexity, which in our view has certain advantages as compared with previous definitions. The complexity measure is based on a phase space approach and is derived from a scale-dependent entropy or information content, drawing from Shannon's definition.[3]

Unlike computational or hierarchical complexity, the new definition of phase space complexity applies to an ensemble of systems. Since it is defined in terms of the probability distribution, it is thus more adaptable to a physical situation.

As mentioned above, the entropy is not an appropriate measure of the complexity of a system. A completely ordered system with a small entropy or a completely disordered system with maximum entropy is the least complex. Like the entropy, the phase space complexity is an extensive quantity; on the other hand, it is defined as a definite weighted sum of scale-dependent entropies.

The perceived complexity of a system is intimately related to the degree of resolution by a measuring procedure of the spatiotemporal correlations. Clearly, on a sufficiently coarse-grained level the correlations are absent and the system "loses" its complexity. The phase space complexity being defined in terms of resolution or scale-dependent entropies or information contents attempts to capture precisely this feature.

In the present paper we elaborate in some detail on this new approach to complexity. In Section 2 we discuss systems or rather signals and their associated probability distributions. In Section 3 we consider the concepts of entropy and define the new phase space complexity. We remark that given the choice of a complexity measure based on the scale-dependent entropies, the requirement of an extensive character essentially dictates its form. In Section 4 we carry out the coarse-graining procedure and introduce a reduction formula in order to evaluate conveniently the complexity for general probability distributions. In Section 5 we limit the discussion to signals or systems with a Gaussian probability distribution and compute explicitly the phase space complexity. In Section 6 we show that the extremal value of the complexity for Gaussian probability distributions can be inferred from a simple variational principle sibject to the constraint that the power of the signal is kept constant. A maximal complexity generally corresponds to long-range correlations in the system or signal indicating a complex behavior. For a one-dimensional signal we find in particular, in agreement with the findings in ref. 1, that the correlations exhibit a $1/f$ behavior for maximal complexity. In Section 7 we present a summary and a conclusion.

Before we embark on the formal analysis we wish to remark that the presence of long-range correlations and of $1/f$ behavior in particular following from a maximal complexity cannot be considered as an independent explanation or derivation of the ubiquitous $1/f$ noise in natural systems which has puzzled researchers for a long time,[5,16] but presumably is a feature of the way we have defined the phase space complexity, i.e., as a weighted sum of scale-dependent entropies. That there should be a connection between the somewhat ambiguous notion of complexity and a natural phenomenon such as $1/f$ noise in interacting systems is, of course, highly speculative. In the present paper our aim is simply to explain and derive some of the properties of the complexity measure introduced in ref. 1.

2. SIGNALS AND PROBABILITY DISTRIBUTIONS

The properties of a physical many-body system manifest themselves by the "signals" emitted by the system. For instance, the "signature" of a magnetic spin system in thermal equilibrium is shown in the fluctuation spectrum of the magnetization. We shall here understand a signal in the broadest sense as either a time series $s(t)$ or a spatial distribution $s(\mathbf{x})$, where $\mathbf{x}$ is a d-dimensional vector. In this extended sense a two-dimensional spatial morphology also represents a physical "signal." Generally, we may also consider a signal with several components s^i, $i = 1,..., N$; in the present context we shall, however, limit ourselves to a one-component signal.

The signal s emitted by a system is characterized by a probability distribution $P(s)$. More specifically, considering a discrete 1D signal s_i, $1 < i < N$, with continuous amplitude, the distribution is

$$P(\{s_i\},\ 1 < i < N) \tag{2.1}$$

properly normalized, i.e.,

$$\int \prod_{i=1}^{N} ds_i\, P(\{s_i\}) = 1 \tag{2.2}$$

In Fig. 1 we show a 1D signal representing, for instance, a time series. The dashed line is just a guide to the eye indicating the random erratic character of the signal. In Fig. 2 we show an example of a 2D signal which could, for instance, represent a percolating cluster.

For equilibrium systems the probability distribution is given by the canonical ensemble,[3]

$$P(\{s_i\}) = \frac{1}{Z} \exp[-\beta E(\{s_i\})] \tag{2.3}$$

where β is the inverse temperature of the heat bath and Z is the partition function

$$Z = \int \prod_{i=1}^{N} \exp[-\beta E(\{s_i\})] \tag{2.4}$$

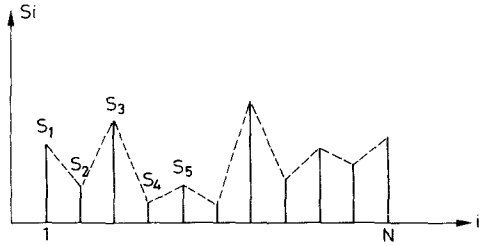

Fig. 1. A time series of an arbitrary random signal. We assume that the time variable is discrete. The dashed line is a guide to the eye.

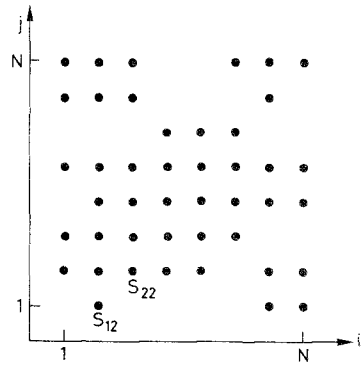

Fig. 2. A 2D spatial morphology or "signal" on a two-dimensional square lattice. The signal could represent a percolating cluster.

$E(\{s_i\})$ is the energy or Hamiltonian of the signal or configuration $\{s_i\}$. Strictly speaking, the distribution in Eqs. (2.3) and (2.4) applies only to time-independent signals or configurations. For time-dependent signals $\{s_i(t)\}$ we must invoke an appropriate dynamical description of, for example, the Langevin type:

$$\frac{ds_i(t)}{dt} = \Gamma \frac{\delta E(\{s_i\})}{\delta s_i(t)} + \eta_i(t) \tag{2.5a}$$

$$\langle \eta_i(t)\, \eta_j(t') \rangle = \frac{2\Gamma}{\beta} \delta_{ij} \delta(t - t') \tag{2.5b}$$

where the form of the white noise $\eta_i(t)$ ensures the coupling to a thermal heat bath.

Generally, the form of $P(\{s_i\})$ is not known for driven systems. For later purposes we shall consider a simple Gaussian form for $P(\{s_i\})$ applied to a 1D signal:

$$P_{\text{Gauss}}(\{s_i\}) = \frac{1}{Z_{\text{Gauss}}} \exp\left(-\frac{1}{2} \sum_{ij} s_i\, g_{ij}^{-1} s_j \right) \tag{2.6a}$$

$$Z_{\text{Gauss}} = \int \prod_{i=1}^{N} ds_i \exp\left(-\frac{1}{2} \sum_{ij} s_i\, g_{ij}^{-1} s_j \right) \tag{2.6b}$$

The signal is here alone characterized by the correlation function g_{ij}.

3. ENTROPY AND COMPLEXITY

The traditional measure of "the amount of uncertainty" represented by the probability distribution $P(\{s_i\})$ for the signal s_i is given by Shannon's information-theoretic entropy (in units of k_B)[3]:

$$S = -\int \prod_{i=1}^{N} ds_i \, P(\{s_i\}) \log P(\{s_i\}) \tag{3.1}$$

For a completely disordered signal $P(\{s_i\}) = \text{const}$ the amount of uncertainty is maximal and the entropy S attains its largest value; the only constraint is the normalization condition for P in Eq. (2.2). Correspondingly, for a well-defined signal at a given time instant or space point $s = s_i$ we have $P(s = s_i) = 1$, and $P = 0$ otherwise; the amount of uncertainty is minimal and the entropy S assumes its smallest value, $S = 0$.

From our previous remarks it follows that the entropy is not a good measure of the complexity of a signal. In a certain sense the completely deterministic signal with a small entropy or the completely disordered signal with maximum entropy is the least complex. An appropriate definition of complexity must be in accordance with our subjective understanding of complex behavior and is clearly associated with the presence of correlations in the signal. For a system in thermal equilibrium with P given by Eq. (2.3) the complexity depends on the parameter β (the inverse temperature) and we anticipate that an appropriate definition of complexity attains its largest value at the critical point in cases where the system described by the Hamiltonian $E(\{s_i\})$ undergoes a second-order phase transition, the critical point with its long-range spatial and temporal correlations giving rise to a maximally complex signal.

As pointed out by Zhang,[1] the information content and thereby the ensuing complexity of a signal must be related to the scale of resolution be it in time or space, and is thus intimitely related to the measuring procedure. This is clearly how the "subjective" element enters in an appropriate definition of complexity.

In order to take into account the way the entropy or information content depends on the scale of resolution, it is natural to introduce a coarse-grained signal[1] depending on a scale parameter τ. For a 1D discrete signal s_i, $1 < i < N$, the scale parameter τ extends between 1 and N; $\tau = 1$ corresponds to the original signal measured with maximum resolution and $\tau = N$ represents the maximally coarse-grained signal. At each level of coarse graining we can define a corresponding entropy $S(\tau)$, where by construction $S(1) = S$ and $S(N) \approx 0$. Since the dimensionless scale factor τ applies to the linear extent of the signal, it follows from the extensive

character of the entropy that $S(\tau) \sim (N/\tau)^d$ for a general d-dimensional signal.

As a measure of the complexity K of a d-dimensional signal the following form is proposed[1]:

$$K = \int_1^N d\tau\, \tau^{d-1} S(\tau) \tag{3.2}$$

which samples in a uniform way over the scale-dependent entropies. This definition of K has the virtue of having the same dimension as the entropy and of being extensive; K is, however, not additive, in contrast to S.

4. COARSE GRAINING AND REDUCTION FORMULA

More specifically, let us consider a one-dimensional discrete signal $\{s_i^{(1)}\}$, $1 < i < N$, characterized by the normalized probability distribution $P^{(1)}(\{S^{(1)}\}, 1 < i < N)$. A coarse graining of the signal determined by the scale factor τ is now achieved by defining the "block" variables, i.e., a mean value of $s_i^{(1)}$ over a scale τ,

$$s_l^{(\tau)} = \frac{1}{\tau} \sum_{i=(l-1)\tau+1}^{l\tau} s_i^{(1)}, \qquad 1 < l < \frac{N}{\tau} \tag{4.1}$$

In Fig. 3 we indicate how the coarse graining is performed.

The original signal is characterized by the distribution $P^{(1)}$ and our first aim is now to find the probability distribution for the coarse-grained signal $\{s_l^{(\tau)}\}$, $1 < l < N/\tau$, i.e., $P^{(\tau)}(\{s_l^{(\tau)}\}, 1 < l < N/\tau)$. This task is most easily carried out by deriving a simple reduction formula expressing $P^{(\tau)}$ in terms of $P^{(1)}$.

Integrating over the variables to be coarse grained, we obtain in a straightforward manner

$$P^{(\tau)}(\{s_l^{(\tau)}\}) = \int \prod_{i=1}^{N} ds_i^{(1)}\, P^{(1)}(\{s_i^{(1)}\}) \prod_{l=1}^{N/\tau} \delta\left(s_l^{(\tau)} - \frac{1}{\tau} \sum_{i=(l-1)\tau+1}^{l\tau} s_i^{(1)}\right) \tag{4.2}$$

Fig. 3. The coarse-graining procedure. Here $s_i^{(1)}$ is the original signal for $\tau = 1$ and $s_i^{(4)}$ the coarse-grained signal corresponding to a scale factor $\tau = 4$.

The coarse-grained or scale-dependent entropy $S(\tau)$ is now evaluated according to Eq. (3.1), i.e.,

$$S(\tau) = -\int \prod_{i=1}^{N/\tau} ds_i^{(\tau)} P^{(\tau)}(\{s_i^{(\tau)}\}) \log P^{(\tau)}(\{s_i^{(\tau)}\}) \tag{4.3}$$

and the complexity K from Eq. (3.2) for $d=1$. We notice that the complexity K depends on (a) the initial distribution $P(\{s_i^{(1)}\})$ and (b) the coarse-graining transformation.

It is useful to express the reduction formula in Eq. (4.2) in terms of the characteristic functions or generators for the probability distributions $P^{(1)}$ and $P^{(\tau)}$. We have[15]

$$Q^{(\tau)}(\{k_i\}, 1 < i < N/\tau) = \int \prod_{i=1}^{N/\tau} ds_i \exp(ik_i s_i^{(\tau)}) P^{(\tau)}(\{s_i^{(\tau)}\}, 1 < i < N/\tau) \tag{4.4}$$

By inserting Eq. (4.4) in Eq. (4.2), it is easy to verify that the reduction formula now takes the more transparent form

$$Q^{(\tau)}(\{k_i\}, 1 < i < N/\tau) = Q^{(1)}(\{\{k_i/\tau\}_\tau\}, 1 < i < N/\tau) \tag{4.5}$$

where $\{k_i/\tau\}_\tau$ indicates τ consecutive variables in $Q^{(1)}$ taking the same value k_i/τ. For $\tau = 2$ and $N = 4$ we have, for instance,

$$Q^{(2)}(k_1, k_2) = Q^{(1)}(k_1/2, k_1/2, k_2/2, k_2/2) \tag{4.6}$$

The reduction formula in Eq. (4.5) is quite general and can easily be generalized to the case of a d-dimensional signal. Provided we can find the characteristic function $Q^{(1)}$ for a given distribution $P^{(1)}$ and, furthermore, recover $P^{(\tau)}$ from $Q^{(\tau)}$, the procedure for evaluating $S(\tau)$ and finally the complexity K is straightforward.

5. GAUSSIAN CASE

In the Gaussian case the characteristic function $Q^{(1)}(\{k_i\})$ takes a particularly simple form. Using Eqs. (2.6a)–(2.6b) and Eq. (4.4) and "completing the square" in the exponent, we obtain a simple functional Gaussian integral yielding

$$Q^{(1)}_{\text{Gauss}}(\{k_i\}, 1 < i < N) = \exp\left(-\tfrac{1}{2} \sum_{ij}^{N} k_i g_{ij}^{(1)} k_j\right) \tag{5.1}$$

The coarse-graining procedure is now carried out by using the reduction formula in Eq. (4.5). We obtain for $Q^{(\tau)}(\{k_i\})$,

$$Q^{(\tau)}_{\text{Gauss}}(\{k_i\},\ 1<i<N) = \exp\left(-\tfrac{1}{2}\sum_{ij}^{N/\tau} k_i\, g^{(\tau)}_{ij} k_j\right) \tag{5.2}$$

where the coarse-grained correlation function $g^{(\tau)}_{ij}$ is given by

$$g^{(\tau)}_{ij} = \frac{1}{\tau^2} \sum_{m=(i-1)\tau+1}^{i\tau} \sum_{n=(j-1)\tau+1}^{j\tau} g^{(1)}_{mn}, \qquad 1<i,j<\frac{N}{\tau} \tag{5.3}$$

The matrix $g^{(\tau)}_{ij}$ is thus constructed by dividing the matrix $g^{(1)}_{ij}$ into blocks of size τ by τ and defining $g^{(\tau)}_{ij}$ as $(1/\tau)^2$ times the sum of matrix elements in the block with entries $(i-1)\tau+1$ to $i\tau$ and $(j-1)\tau+1$ to $j\tau$. Since the characteristic function in the Gaussian case is covariant under coarse graining, it is an easy task to derive the probability distribution $P^{(\tau)}$ for the coarse-grained signal,

$$P^{(\tau)}(\{s^{(\tau)}_i\}) = \frac{1}{Z^{(\tau)}} \exp\left[-\frac{1}{2}\sum_{ij}^{N/\tau} s^{(\tau)}_i (g^{(\tau)}_{ij})^{-1} s_j\right] \tag{5.4a}$$

$$Z^{(\tau)} = \int \prod_{i=1}^{N/\tau} ds_i \exp\left[-\frac{1}{2}\sum_{ij}^{N/\tau} s^{(\tau)}_i (g^{(\tau)}_{ij})^{-1} s^{(\tau)}_j\right] \tag{5.4b}$$

We note that the probability distribution $P^{(\tau)}$ is covariant under coarse graining; this is, of course, a special feature of the Gaussian distribution.

For stationary time signals or spatially homogeneous morphologies the correlation function $g^{(\tau)}_{ij}$ only depends on $i-j$ and a further analysis is facilitated by a Fourier spectral decomposition.[1] Introducing the Fourier transforms

$$s^{(\tau)} = \left(\frac{\tau}{N}\right)^{1/2} \sum_k \exp(ikl)\, s^{(\tau)}(k) \tag{5.5a}$$

$$g^{(\tau)}_{mn} = \frac{\tau}{N} \sum_k \exp[ik(m-n)]\, g^{(\tau)}(k) \tag{5.5b}$$

and

$$(g^{(\tau)}_{mn})^{-1} = \frac{\tau}{N} \sum_k \exp[ik(m-n)]\, g^{(\tau)}(k)^{-1} \tag{5.5c}$$

where $k = 2\pi n\tau/N$ and $n = 1,\ldots, N/\tau$, we obtain, noting that $s^{(\tau)}(k)^* = s^{(\tau)}(-k)$,

$$P^{(\tau)}=\frac{1}{Z^{(\tau)}}\exp\left[-\frac{1}{2}\sum_k |s^{(\tau)}(p)|^2\, g^{(\tau)}(p)^{-1}\right] \tag{5.6a}$$

$$Z^{(\tau)}=\int \prod_p ds^{(\tau)}(p)\exp\left[-\frac{1}{2}\sum_k |s^{(\tau)}(p)|^2\, g^{(\tau)}(p)^{-1}\right] \tag{5.6b}$$

The relationship between the correlation functions $g^{(\tau)}$ and $g^{(1)}$ at different levels of coarse graining or resolution is now given by

$$g^{(\tau)}(k)=\frac{1}{\tau^3}\left[\frac{\sin(k/2)}{\sin(k/2\tau)}\right]^2 g^{(1)}\left(\frac{k}{\tau}\right) \tag{5.7}$$

For the Gaussian distribution in Eqs. (5.6a)–(5.6b) the entropy $S(\tau)$ takes a simple form. Using Eq. (4.3), we find

$$S(\tau)=\tfrac{1}{2}\sum_k \{1+\log[2\pi g^{(\tau)}(k)]\} \tag{5.8}$$

which in conjunction with Eq. (5.7) for $g^{(\tau)}(k)$ defines the entropy or information content of a Gaussian signal with correlation function $g^{(1)}(k)$ on a resolution level characterized by τ. A scale factor $\tau=1$ corresponds to a complete resolution of the original signal; $\tau=N$ to no resolution at all.

In the long-wavelength or low-frequency limit $k\to 0$, depending on the nature of the signal, we can approximate the expression for $g^{(\tau)}(k)$ in Eq. (5.7),

$$g^{(\tau)}(k)<(1/\tau)\, g^{(1)}(k/\tau) \tag{5.9}$$

in accordance with ref. 1. We note that the coarse graining scales both the amplitude of g and its range. In this approximation the entropy takes the simple form

$$S(\tau)=\frac{1}{2}\frac{N}{\tau}(1+\log 2\pi-\log\tau)+\frac{1}{2}\sum_k \log g^{(1)}\left(\frac{k}{\tau}\right) \tag{5.10}$$

and we find for the complexity, using Eq. (3.2) in the case $d=1$,

$$K=\frac{1}{2}N\log N(1+\log 2\pi-\log N)+\int_1^N d\tau \sum_k \log g^{(1)}\left(\frac{k}{\tau}\right) \tag{5.11}$$

The generalization of the above expressions to the d-dimensional case is straightforward. For a d-dimensional signal $s_{\mathbf{l}}^{(\tau)}$, $\mathbf{l}=(l_1, l_2,..., l_d)$, we find

$$g^{(\tau)}(\mathbf{k})=\frac{1}{\tau^{3d}}\prod_{i=1}^{d}\left[\frac{\sin(k_i/2)}{\sin(k_i/2\tau)}\right]^2 g^{(1)}\left(\frac{\mathbf{k}}{\tau}\right) \tag{5.12}$$

where $\mathbf{k} = (k_1, k_2, ..., k_d)$ is the wavevector of the Fourier resolutions of $s_{\mathbf{l}}^{(\tau)}$, $g_{\mathbf{mn}}^{(\tau)}$, and $g_{\mathbf{mn}}^{(\tau)^{-1}}$, i.e.,

$$s_{\mathbf{l}}^{(\tau)} = (\tau/N)^{d/2} \sum_{\mathbf{k}} \exp(i\mathbf{kl})\, s^{(\tau)}(\mathbf{k})$$

etc. Correspondingly, we have for the probability distribution

$$P^{(\tau)} = \frac{1}{Z^{(\tau)}} \exp\left[-\frac{1}{2} \sum_{\mathbf{k}} k\, |s^{(\tau)}(\mathbf{k})|^2\, g^{(\tau)}(\mathbf{k})^{-1} \right] \tag{5.13a}$$

$$Z^{(\tau)} = \int \prod_{\mathbf{k}} ds^{(\tau)}(\mathbf{k}) \exp\left[-\frac{1}{2} \sum_{\mathbf{k}} |s^{(\tau)}(\mathbf{k})|^2\, g^{(\tau)}(\mathbf{k})^{-1} \right] \tag{5.13b}$$

for the entropy

$$S(\tau) = \frac{1}{2} \left(\frac{N}{\tau} \right)^d (1 + \log 2\pi - d \log \tau) + \frac{1}{2} \sum_{\mathbf{k}} \log g^{(1)}\left(\frac{\mathbf{k}}{\tau} \right) \tag{5.14}$$

and for the complexity

$$K = \frac{1}{2} N^d \log N (1 + \log 2\pi - d \log N) + \int_1^N d\tau\, \tau^{d-1} \sum_{\mathbf{k}} \log g^{(1)}\left(\frac{\mathbf{k}}{\tau} \right) \tag{5.15}$$

6. VARIATIONAL PRINCIPLE AND 1/*f* NOISE

In the Gaussian case the complexity depends on the correlations in the original signal as characterized by $g^{(1)}$. In order to compare Gaussian signals with different correlations, it is convenient to consider the mean power of the signal as a "common measuring stick"[1]; in other words, we attempt to compare the complexity of Gaussian signals with different correlations but with the same "energy content."[1]

The average power of the signal $s_{\mathbf{l}}^{(1)}$ is given by

$$W = \frac{1}{N^d} \sum_{\mathbf{k}} \langle |s^{(1)}(\mathbf{k})|^2 \rangle \tag{6.1}$$

Using the distribution in Eqs. (5.13a)–(5.13b), we obtain in the Gaussian case

$$W = \frac{1}{N^d} \sum_{\mathbf{k}} g^{(1)}(\mathbf{k}) \tag{6.2}$$

In the one-dimensional case it has been shown,[1] using the Kullback inequality,[12] that the complexity K, subject to a constraint of fixed power W, takes the maximal value $K_{\max}$ for the case of $1/f$ noise correlations in the signal, i.e., $g^{(1)}(k) \simeq 1/k$ for $K = K_{\max}$. Here we wish to investigate the extremal properties of K using a standard variational principle.[3] First we note that by successively exchanging the sum over wavenumbers $\mathbf{k}$ and the integration over the scale parameter τ in Eq. (5.15) we can perform the τ integration explicitly. Using the approximation

$$\int_1^{N/m} d\tau \sum_{n=1}^{N/\tau} \simeq \sum_{n=m}^{N} \int_1^{N/n} d\tau + \sum_{n=1}^{m} \int_1^{N/m} d\tau \tag{6.3}$$

we obtain in $d = 1$

$$K = \frac{1}{2} N \log N(1 + \log 2\pi - \log N) + \frac{1}{2} \sum_k \left(\frac{2\pi}{k} - 1 \right) \log g^{(1)}(k) \tag{6.4}$$

in accordance with ref. 1. For $d = 2$, applying Eq. (6.3) twice, we have

$$\begin{aligned} K = {} & \frac{1}{2} N^2 \log N(1 + \log 2\pi - 2 \log N) \\ & + \frac{1}{4} \sum_{k_1} \sum_{k_2 > k_1} \left[\left(\frac{2\pi}{k_2} \right)^2 - 1 \right] \log g^{(1)}(\mathbf{k}) \\ & + \frac{1}{4} \sum_{k_1} \sum_{k_2 < k_1} \left[\left(\frac{2\pi}{k_1} \right)^2 - 1 \right] \log g^{(1)}(\mathbf{k}) \end{aligned} \tag{6.5}$$

In the same way we can derive expressions for K for $d > 2$. However, owing to the branching in exchanging the k-sums and the τ-integration, we have not derived a closed expression for K valid for general d.

We are now in a position to apply a variational principle in order to determine the form of $g^{(1)}(\mathbf{k})$ corresponding to an extremal value for K. Introducing the constraint $W = \text{const}$ by means of a Lagrange multiplier λ, we vary the form

$$F = K - \lambda W \tag{6.6}$$

and require $\delta F = 0$ subject to variations of $g^{(1)}$.

For a 1D signal we obtain

$$\frac{1}{2} \sum_k \left(\frac{2\pi}{k} - 1 \right) \frac{\delta g^{(1)}(k)}{g^{(1)}(k)} - \lambda \frac{1}{N} \sum_k \delta g^{(1)}(k) = 0 \tag{6.7}$$

Solving for $g^{(1)}$ and eliminating λ using Eq. (6.2), we obtain in the long-wavelength limit $k \to 0$,

$$g^{(1)}(k) = \frac{2\pi WN}{\sum_p (2\pi/p - 1)} \frac{1}{k} \tag{6.8}$$

in accordance with ref. 1. In other words, for a 1D spatial signal, i.e., a 1D configuration, the complexity K takes its maximum value in the case of correlations behaving as $1/k$ in wavenumber space. In configuration space the correlations fall off as $\log r$, i.e., $g_{ij} \sim \log |i-j|$. For a temporal signal we find, correspondingly, $g^{(1)}(k) \equiv g^{(1)}(\omega) \sim 1/\omega$, i.e., $1/f$ noise.

For a 2D signal we obtain in a similar manner

$$g^{(1)}(k_1, k_2) = \frac{WN^2[(2\pi/k_1)^2\,\theta(k_1-k_2) + (2\pi/k_2)^2\,\theta(k_2-k_1) - 1]}{\sum_{p_1 p_2}[(2\pi/p_1)^2\,\theta(p_1-p_2) + (2\pi/p_2)^2\,\theta(p_2-p_1) - 1]} \tag{6.9}$$

The anisotropic form of $g^{(1)}(k_1, k_2)$ is due to the choice of coarse-graining procedure. In the long-wavelength limit $k_1, k_2 \to 0$ we find $g^{(1)} \sim 1/k_1^2$, $1/k_2^2$. For a 2D spatial morphology this behavior also corresponds to correlations falling off like $\log r$.

7. SUMMARY AND CONCLUSION

In this paper we have discussed a novel definition of complexity based on a phase space approach.[1] The complexity measure K can in principle be evaluated for any kind of signal be it a time series or a higher-dimensional spatial morphology. The complexity is derived from the probability distribution for the signal by expressing K in terms of a graded sum of entropies evaluated at different levels of resolution of the signal.

For a Gaussian signal characterized by a correlation function $g(k)$, where k is the wavenumber or frequency, the scale-dependent entropy and hence the complexity can be evaluated explicitly.[1] Introducing a simple variational principle, which supplements the derivation in ref. 1 and puts the problem within the general context of statistical mechanics methods,[3] we show that for signals with constant power, maximal complexity is attained for signals which in one dimension have long-range correlations of the $1/f$ type, i.e., a correlation function $g(k) \sim 1/k$. Note that by maximizing alone the entropy subject to the above constraint we obtain a white noise behavior, i.e., $g(k) \sim$ const. For a two-dimensional signal or structure, maximum complexity corresponds to spatial correlations decaying logarithmically.

The virtue of the phase space approach[1] is that it applies to any signal for which we can define a probability distribution; in this respect the

present method differs markedly from other approaches.[6,8,11] As far as uniqueness is concerned, the present definition has the virtue of yielding a complexity which apart from logarithmic corrections is an extensive quantity. In this respect the complexity resembles the entropy and enables us to define a complexity density.

In recent years there has been a substantial effort directed toward understanding the ubiquitous nature of $1/f$ noise in driven open systems.[5,16] As emphasized by Zhang,[1] the present definition of complexity has the interesting feature of displaying a maximum for a one-dimensional Gaussian signal with $1/f$ noise correlations. Since we still have so little understanding of either $1/f$ noise or the notion of complexity, it is difficult to assess the importance of this observation.

We also wish to comment on the variational principle presented here. Variational principles play an important role in physics. For instance, in statistical mechanics a variational principle based on the information-theoretic definition of entropy[3] allows for the derivation of the statistical ensembles and eventually leads to a derivation of thermodynamics; in mechanics and field theory the principle of least action, i.e., maximizing the action subject to various constraints, implies the equations of motion. Drawing the obvious parallel to the present discussion, a useful definition and understanding of complexity should allow us to formulate a variational principle using K which in a generic sense would give us some insight into the mechanism giving rise to maximally complex signals. Here we have used a simple variational principle to derive $1/f$ noise, but it seems difficult to go beyond this admittedly rather simple calculation.

Finally, let us summarize some obvious extensions of the present calculation. It clearly would be of interest to extend the derivation in ref. 1 and the present discussion to signals with a non-Gaussian character. Within the scheme of a Ginzburg–Landau formulation,[3] where the Gaussian approximation corresponds to the leading term (with arbitrary spatial correlations), one might include higher-order terms and apply a renormalization-group type calculation in order to evaluate the complexity as a function of temperature. It would then be possible to verify the claim[1] that the complexity has a maximum at the critical temperature. Another line of approach is to consider, for example, the one-dimensional Ising model and evaluate the complexity using the transfer matrix method.[3]

ACKNOWLEDGMENTS

The author wishes to acknowledge useful discussions with Per Bak, Itamar Procaccia, Axel Svane, and Y.-C. Zhang.

REFERENCES

1. Y.-C. Zhang, *J. Phys.* I (Paris) **1**:971 (1991).
2. K. G. Wilson, *Sci. Am.* **241**:140 (1979).
3. M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* (Prentice-Hall, 1989).
4. J. Feder, *Fractals* (Plenum Press, New York, 1988).
5. P. Bak, C. Tang, and K. Wiesenfeld, *Phys. Rev. Lett.* **59**:381 (1987); *Phys. Rev. A* **38**:364 (1988).
6. B. A. Huberman and T. Hogg, *Physica* **22D**:376 (1986).
7. T. Hogg and B. A. Huberman, *Phys. Rev. A* **32**:2338 (1985).
8. C. P. Bachas and B. A. Huberman, *Phys. Rev. Lett.* **57**:1965 (1986).
9. W. Ebeling and G. Nicolis, *Europhys. Lett.* **14**:191 (1991).
10. G. Nicolis and C. Nicolis, *Physica A* **163**:215 (1990).
11. P. Grassberger, Randomness, information, and complexity, Wuppertal, Preprint (1989); Cellular automata, Wuppertal Preprint (1990).
12. S. Kullback, *Information Theory and Statistics* (Wiley, New York, 1959).
13. W. Ebeling and G. Nicolis, *Europhys. Lett.* **14**:191 (1991).
14. G. Nicolis and C. Nicolis, *Physica A* **163**:215 (1990).
15. W. Feller, *An Introduction to Probability Theory and its Applications*, 3rd ed. (Wiley, New York, 1968).
16. W. H. Press, *Comments Astrophys.* **7**:103 (1978).