Error threshold in optimal coding, numerical criteria, and classes of universalities for complexity

David B. Saakian*
Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan
and Yerevan Physics Institute, Alikhanian Brothers Street 2, Yerevan 375036, Armenia
(Received 9 March 2004; revised manuscript received 17 June 2004; published 19 January 2005)

The free energy of the random energy model at the transition point between the ferromagnetic and spin glass phases is calculated. At this point, equivalent to the decoding error threshold in optimal codes, the free energy has finite size corrections proportional to the square root of the number of degrees. The response of the magnetization to an external ferromagnetic phase is maximal at values of magnetization equal to one-half. We give several criteria of complexity and define different universality classes. According to our classification, at the lowest class of complexity are random graphs, Markov models, and hidden Markov models. At the next level is the Sherrington-Kirkpatrick spin glass, connected to neuron-network models. On a higher level are critical theories, the spin glass phase of the random energy model, percolation, and self-organized criticality. The top level class involves highly optimized tolerance design, error thresholds in optimal coding, language, and, maybe, financial markets. Living systems are also related to the last class. The concept of antiresonance is suggested for complex systems.

DOI: 10.1103/PhysRevE.71.016126 PACS number(s): 89.75.−k, 02.70.−c, 89.70.+c, 05.50.+q

I. INTRODUCTION

A. Complexity

The definition of statistical complexity is an entirely open problem in statistical mechanics (see [1,2] for an introduction to the problem and [3] for a recent discussion). There are a lot of different definitions having sometimes a common context. A certain success was the discovery of the idea of the "schema," highly compressed information, introduced by Gell-Mann for complex adaptive (we assume, for all complex) systems. Some attempts, based mainly on entropy concepts, have been undertaken to define the concept of complexity. The approach [4–6] relevant for our investigation is of special interest. A very interesting aspect of complex phenomena is related to the edge of chaos (the border between chaotic and deterministic motion), the phase of complex adaptive systems (CASs) [2,7]. The concept of the edge of chaos, independently suggested by Bak *et al.* [8], Kauffman [7], and Langton [9], is not well defined quantitatively. However, it is widely accepted that this concept is connected with the sandpile [10]. This concept is of special importance due to its possible relation to the creation of life and evolution [7]. This paper is devoted to relations between this phenomenon and some aspects of information theory and optimal coding [11]. We assume that the definition of a single (or best) complexity measure is a subjective one, even with the reasonable constraint that complexity should vanish for totally ordered or disordered motion (see the dispute [12,13]). More strict is the definition of different universality classes of complexity, which is presented in this paper. We suggest several numerical criteria for complex adaptive properties. In practice we suggest identifying the universality class of complexity from the experimental data and choosing a model from the same class to describe the phenomenon.

We assume the following picture of complex phenomena. The following hierarchy is presented: instead of the microscopic motion of molecules or spins we deal with the macroscopic thermodynamic variables. In addition, some further structures arose, sometimes proportional to the number of fractional degrees of the particles. One can understand qualitatively the complexity as a measure of the additional structures. We are not going to scrutinize the concrete features of those structures. We will just evaluate the total measure of the structures on the basis of the free energy expression, including finite size corrections. The suggested complexity measures could be applied to pure systems, as well as to those defined via the disorder ensemble (as in spin glasses). For interesting cases of complex adaptive systems, a hierarchy in the definition of the model, either a disorder ensemble (as in spin glasses) or the scale of the system (spatial or temporal) should be represented. The structures themselves are derived from microscopic motions of spins via order parameters. When those order parameters fluctuate, they can be handled like microscopic spins or molecules. Therefore, in such cases, including the optimal coding of the article, we can identify complex phenomena as a situation with changing reality or creation of new reality (the thermodynamic reality is a mapping of molecular motions into a few thermodynamic variables). A good analogy is the weak interaction in high energy physics on different scales. On the level of low energies it can be described through a simple picture by Fermi. However, on the level of 100 GeV, when broken symmetry is restored, there is an absolutely different reality.

Our approach (to investigate just the ensemble averaged free energy instead of the logarithm of the number of different ground states in spin glasses [49]) is coherent with the idea of Jayenes [14] saying that there is a single probability in physics, measurable during observations, and there is no need to fracture it.

We assume that it is proper to define the criteria of complexity via the free energy rather than via the entropy (see the discussion in Sec. III). In [15] a criterion of complexity,

*Electronic address: saakian@jerewan1.yerphi.am

functioning from galaxies and stars to brains and society, has been suggested: the rate of free energy density change. It is an interesting approach, but we prefer to deal with dimensionless criteria. The point that the subdominant free energy describes the degree of system complexity has already been recognized, at least, for two-dimensional (2D) critical systems. According to [16], the subdominant term is proportional to the conformal charge (effective number of bosonic degrees of freedom). We just suggest applying this criterion for any system, as one of the complexity criteria. For the situation when there is no explicit free energy (optimal coding, sandpiles, etc.) one should try to find some equivalent statistical mechanical formulation of the theory and investigate the free energy. In this work we give a derivation of the subdominant free energy for a case related to optimal coding and identify the universality class of the error threshold. We provide other criteria of complexity and a list of universality classes.

B. Optimal coding

Information processing should certainly be a property of complex adaptive systems. What can be clarified by statistical physics? The connection of statistical physics to information theory is known from the work of Jayenes [14]. In 1989 Sourlas [17] found a connection of the random energy model (REM) of a spin glass by Derrida [18] with an important branch of Shannon information theory, i.e., optimal coding theory [11]. In [19] I proved Sourlas's idea for the principal case of finite velocity codes. An important result has been derived by Rujan [20] regarding coding by statistical mechanics models at finite temperatures. In a series of work [21–25] we derived the main results of Shannon information theory using the REM. Those results were repeated by alternative methods later (see the review [26]). Because we are going to consider models at the junction of statistical mechanics and optimal coding, some features of optimal coding should be briefly mentioned [11].

Let us consider the transition of information, a sequence of ± 1, through a noisy channel to a receiver. There exists original information which is a sequence of +1 and $-1:\epsilon_1,\ldots,\epsilon_N$. If we send some letters through the channel, due to the noise, the letters change their correct values, and information is partially lost. Therefore, to recover further the original message in a proper way it is needed to originally send more information using coding. The encoding is a mapping of the initial message of length N onto a sequence which has the length αN $(\alpha > 1)$

$$(\epsilon_1,\ldots,\epsilon_N) \to (f_1(\epsilon_1,\ldots,\epsilon_N),\ldots,f_{\alpha N}(\epsilon_1,\ldots,\epsilon_N)). \qquad (1)$$

A noisy channel is represented as a mapping of the message by random letters η_j:

$$\begin{aligned}&(f_1(\epsilon_1,\ldots,\epsilon_N),\ldots,f_{\alpha N}(\epsilon_1,\ldots,\epsilon_N))\\ &\quad \to (f_1(\epsilon_1,\ldots,\epsilon_N)\eta_1,\ldots,f_{\alpha N}(\epsilon_1,\ldots,\epsilon_N)\eta_{\alpha N}),\end{aligned} \qquad (2)$$

where the noisy η_j, $1 \leqslant j \leqslant \alpha N$, are independent random numbers with probability distribution

$$P(\eta_j) = \frac{1+m_0}{2}\delta(\eta_i - 1) + \frac{1-m_0}{2}\delta(\eta_i + 1). \qquad (3)$$

The transmitter introduces additional information $(\alpha > 1)$, and the receiver must extract useful information. These two operations are called encoding and decoding. In general, coding is a mapping of the initial message of length N onto a sequence which has the length αN, $\alpha > 1$. Thus, encoding is done with αN functions $f_j = \pm 1$. The value $\alpha^{-1} = R$, the "rate" of information transmission, characterizes the degree of redundancy. Decoding in the general case is the procedure of extracting the initial message out of the noisy sequence $(f_1\eta_1,\ldots,f_{\alpha N}\eta_{\alpha N})$.

When is errorless decoding possible? We have a Boltzmann-Gibbs-Shannon measure of information for a discrete distribution p_i:

$$-\sum_i p_i \ln p_i. \qquad (4)$$

In the original message any letter ± 1 carries information ln 2. In case of noise by Eq. (3) any letter carries the information

$$\ln 2 - h, \qquad (5)$$

$$h = -\left(\frac{1+m_0}{2}\ln\frac{1+m_0}{2} + \frac{1-m_0}{2}\ln\frac{1-m_0}{2}\right).$$

In the last expression we extracted from ln 2 the entropy h of the distribution by Eq. (3).

The encoding can work in different ways. To extract the original message without error it is reasonable to put the constraint

$$\alpha N\left(\ln + \frac{1+m_0}{2}\ln\frac{1+m_0}{2} + \frac{1-m_0}{2}\ln\frac{1-m_0}{2}\right) \geqslant N\ln 2. \qquad (6)$$

On the left, we have the information of the received message. On the right, we have the information to be extracted. This is the Shannon fundamental theorem for errorless decoding. Only very special coding schemes correspond to the special case, when the last expression transforms to an equality. Such codes are optimal ones. They are universal mathematical constructions, like the critical Hamiltonian in phase transitions.

C. Statistical mechanics for coding

How could statistical mechanics be applied for optimal coding? To encode the original sequence $\epsilon_1,\ldots,\epsilon_N$, one constructs a Hamiltonian $H(s)$, a function of $(s_1,\ldots,s_N)$:

$$\begin{aligned}-H(s_1,\ldots,s_N) &= f_1(s_1,\ldots,s_N)f_1(\epsilon_1,\ldots,\epsilon_N) + \cdots\\ &\quad + f_{\alpha N}(s_1,\ldots,s_N)f_{\alpha N}(\epsilon_1,\ldots,\epsilon_N)\\ &\equiv h_0(y_1,\ldots,y_{\alpha N}),\end{aligned}$$

$$y_j = f_j(s_1,\ldots,s_N)f_j(\epsilon_1,\ldots,\epsilon_N). \qquad (7)$$

Here the functions f_j, $1 \leqslant j \leqslant \alpha N$, are the products of some p spins, $f_j = s_{j_1},\ldots,s_{j_p}$, and H has a minimal value at $s_i = n_i$. The influence of noise is very simple: every term (word) in Eq. (7) is multiplied by a noise, and instead of the pure Hamiltonian $H(s)$ we have a noisy one,

$$-H(s,\eta) = h_0(y'_1,\ldots,y'_{\alpha N}),$$

$$y'_j = f_j(s_1,\ldots,s_N) f_j(\epsilon_1,\ldots,\epsilon_N)\eta_j. \tag{8}$$

To find the minimum of the Hamiltonian, one could consider the statistical mechanics of the spin system with the Hamiltonian H at very low temperatures,

$$Z = \sum_{s_i=\pm 1} e^{-\beta H(s,\eta)}, \tag{9}$$

where $H(s,\eta) \equiv H(s_1,\ldots,s_N,\ \eta_1,\ldots,\eta_N)$, $\beta \to \infty$. Without noise ($\eta_j=1$) one can calculate the configuration $(s_1,\ldots,s_N)$ giving the main contribution to Z at $\beta\to\infty$. We have the following expression for the mean magnetization:

$$\langle s_i \rangle = \epsilon_i. \tag{10}$$

It has been proved in [19] that Eq. (10) is correct also for nonzero noise below the Shannon error threshold.

In Shannon information theory one considers transmission of a message to a receiver. The influence of the noise corresponds to a simple product of coding words $f_j(\epsilon_1,\ldots,\epsilon_N)$ with a noisy letter η_j (both are accepting the values ±1).

D. Other versions of error threshold in statistical mechanics

What generalizations of the considered scheme are possible to accept? Instead of Eq. (9) one can consider a partition with the quantum noise

$$Z = \mathrm{Tr}\exp\left[-\beta\left(H(\sigma_1^z,\ldots,\sigma_N^z) + \mu\sum_{j=1}^{N}\sigma_i^x\right)\right], \tag{11}$$

and

$$Z = \mathrm{Tr}\exp\left(-\beta\left\{\exp\left[\gamma\left(1-\sum_{j=1}^{N}\sigma_i^x/N\right)\right]H(\sigma_1^z\ldots\sigma_N^z)\right\}\right), \tag{12}$$

where H is a mean-field-like Hamiltonian like

$$H(\sigma_1^z,\ldots,\sigma_N^z) \equiv H_0\left(\sum_i \sigma_i^z\right), \tag{13}$$

having a minimum at the configuration $s_i=1$, $1\leqslant i\leqslant N$. Successful information transmission is connected with the phase where $\langle\sigma_i^z\rangle \equiv m_i > 0$ (there is a nonzero longitudinal magnetization); in Eq. (11) the quantum noise is additive, in Eq. (12) it is a multiplicative one. Equations (11) and (12) are connected to the evolution models [27–30], when genetical information is transmitted to future generations. It is interesting that Eigen derived the correct error threshold in this model [27] just from informational theoretical arguments long before the Sourlas work about the connection of statistical mechanics with information theory. The Eigen model has been exactly solved only recently [29]; Eigen's formula for the error threshold was confirmed.

Our purpose is to connect the complex adaptive phase with the neighborhood of the error threshold (6). We will consider the border between the ferromagnetic and spin glass phases in the random energy model, investigating its statistical mechanics by Eq. (9). We will not consider the informational-theoretical aspects of the problem any more—the subject is well discussed in [25]. In Sec. II we will derive the finite size correction to the free energy and investigate the dependence of the magnetization on the bulk ferromagnetic coupling. In Sec. III we will give a definition of a complex adaptive property and define different universality classes. In Sec. IV we will suggest another concept of complex adaptive systems, i.e., the possibility of antiresonance. In conclusion, we will briefly discuss our results and general aspects of complex adaptive systems.

II. RANDOM ENERGY MODEL

A. Energy configuration formulation

To investigate the complex phenomena we consider an equilibrium statistical physics situation similar to the edge of chaos point, i.e., the border between the ferromagnetic and spin glass (SG) phases in the random energy model. The finite size corrections of free energy will be calculated later on. In the REM N spins $s_i=\pm 1$ interact through $\binom{p}{N} \equiv [N!/p!(N-p)!]$, $p\to\infty$, couplings with the Hamiltonian [18,21]

$$H = -\sum_{1\leqslant i_1,\ldots,i_p\leqslant N}[j^0_{i_1,\ldots,i_p} + j_{i_1,\ldots,i_p}]s_{i_1},\ldots,s_{i_p}. \tag{14}$$

Here $j^0_{i_1,\ldots,i_p}$ are ferromagnetic couplings

$$j^0_{i_1,\ldots,i_p} = \frac{J_0 N}{\binom{p}{N}}, \tag{15}$$

and for quenched disorder $j_{i_1,\ldots,i_p}$ we have a distribution

$$\rho_0(j_{i_1,\ldots,i_p}) = \frac{1}{\sqrt{\pi}}\sqrt{\frac{\binom{p}{N}}{N}}\exp\left(-j^2_{i_1,\ldots,i_p}\frac{\binom{p}{N}}{N}\right). \tag{16}$$

We see that there are ferromagnetic and random couplings, and J_0 defines the ferromagnetic degree.

In our spin model there are 2^N different energy configurations. It has been found by Derrida that for large values of p there is a factorization for the energy level distribution. For $\alpha\neq\beta$ [18]

$$\rho(E_\alpha,E_\beta) = \rho(E_\alpha)\rho(E_\beta). \tag{17}$$

For the first configuration with $s_i=1$ [21],

$$\rho_1(E_1) = \frac{1}{\sqrt{\pi N}}\exp[-(E_1+J_0N)^2/N], \tag{18}$$

and for the other 2^N-1 levels [18],

$$\rho(E) = \frac{1}{\sqrt{\pi N}}\exp(-E^2/N). \tag{19}$$

The REM has two equivalent definitions: via the energy configuration Eqs. (18) and (19) and via the spin Hamiltonian version Eq. (14). It is possible to solve the REM through the ordinary spin glass approach, as well as by using the factorization property Eq. (17). According to the energy configuration approach, we perform averaging via the energy level distribution (instead of random couplings in the usual case of disordered systems)

$$\langle \ln Z\rangle \equiv \left\langle \ln\sum_{\alpha} \exp(-\beta E_{\alpha})\right\rangle_E . \tag{20}$$

Here β is the inverse temperature. It is possible to derive the result [21] that at high enough values of $J_0 > \sqrt{\ln 2}$ [see Eq. (27)] at low temperatures the system is in the ferromagnetic phase with magnetization

$$m_i = 1. \tag{21}$$

Using the trick [18]

$$\langle \ln Z\rangle = \Gamma'(1) + \int_{-\infty}^{\infty} \ln t \frac{d\langle \exp(-tZ)\rangle}{dt}dt, \tag{22}$$

one can factorize the integration via different energy levels E_α. The average is over energy distributions Eqs. (18) and (19). It is enough only to calculate for $\langle \exp(-te^{\beta E_i})\rangle$ for the single level. We consider

$$f(u) \equiv \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} \exp[-y^2 - e^u \exp(-\lambda y)]dy, \tag{23}$$

where $\lambda = \beta\sqrt{N}$ and $\langle \exp(-te^{\beta E_i})\rangle = f(\ln t)$. We can further derive for $\langle e^{-tZ}\rangle \equiv \langle \exp[-t\exp(\Sigma_{\alpha=1}^M E_\alpha)]\rangle$

$$\Psi(u) = [f(u+u_f)f(u)^M], \tag{24}$$

where $u = \ln t$, $u_f = J_0 N\beta$, $M = 2^N - 1$. Now Eq. (22) gives

$$\langle \ln Z\rangle = \Gamma'(1) + \int_{-\infty}^{\infty} u\frac{d\Psi(u)}{du}du. \tag{25}$$

$f(u)$ is a monotonic function. With exponential accuracy it equals 1 below 0; then becomes 0 above it. We need four asymptotic regimes [18,21]:

$$f(u) \approx \begin{cases} \dfrac{1}{\sqrt{\pi}\lambda}\Gamma\left(\dfrac{2u}{\lambda^2}\right)e^{-u^2/\lambda^2}, & \lambda \ll u, \\ \dfrac{1}{\sqrt{\pi}}\displaystyle\int_{u/\lambda}^{\infty} dx e^{-x^2}, & |u| \ll \lambda^2, \\ 1 - \dfrac{1}{\sqrt{\pi}\lambda}\Gamma\left(-\dfrac{2u}{\lambda^2}\right)e^{u^2/\lambda^2}, & -\dfrac{\lambda^2}{2} < u \ll -\lambda, \\ 1 - e^{u+\lambda^2/4}, & -\lambda^2 < u < -\dfrac{\lambda^2}{2}. \end{cases} \tag{26}$$

We are interested in those regimes asymptotic for $u \sim N$ or $u \sim \sqrt{N}$ and $\lambda \gg 1$. As $f(u+u_f)f(u)^M$ is like a step function, its derivative is like a δ function with the center at some $-u_0$. The vicinity of $-u_0$ contributes mainly to the integral in Eq. (25) (the bulk value is equal to u_0). A ferromagnetic (FM) phase appears when $-u_f$ [the center of the function $f(u+u_f)$] is further left than $-\sqrt{N}\lambda \ln 2$ [the center of $f(u)^M$]. The FM-SG border corresponds to

$$J_0 = \sqrt{\ln 2}, \quad \infty > \beta > \sqrt{\ln 2}. \tag{27}$$

When there is only the first level with distribution (18), $\langle \ln Z\rangle \equiv -\beta\langle E_1\rangle = u_f \equiv J_0 N\beta$. For that case $\Psi = f(u+u_f)$. Therefore Eq. (25) gives

$$\Gamma'(1) + \int_{-\infty}^{\infty} ud[f(u+u_f)] = u_f. \tag{28}$$

Using the last identity, we transform Eq. (25) into

$$\begin{aligned}\langle \ln Z\rangle &= \Gamma'(1) + \int_{-\infty}^{\infty} ud\Psi(u) \\ &= u_f - \int_{-\infty}^{\infty} ud\Psi_1(u) \\ &= u_f + \int_{-\infty}^{\infty} \Psi_1(u)du,\end{aligned} \tag{29}$$

where $\Psi_1(u) = f(u+u_f)[1-f(u)^M]du$.

B. Exact border of ferromagnetic and spin glass phases

Let us first consider the exact border of two phases $J_0 = \sqrt{\ln 2}$. $\Psi_1(u)$ is a product of two monotonic functions, decreasing (one to the left, the other to the right) from the point $u = -u_f$. We define an auxiliary function $F(u)$ by the differential equation

$$F'(u) = f(u+u_f). \tag{30}$$

Using the second equation in (26), we derive for $|u| \ll \lambda^2$

$$F(u-u_f) = \int_0^{u/\lambda} dx \frac{\lambda}{\sqrt{\pi}}\int_x^{\infty} e^{-y^2}dy. \tag{31}$$

Let us denote $\Psi_2(u) = [1-f(u)^M]$ and perform integration by parts in Eq. (29):

$$\begin{aligned}\langle \ln Z\rangle &= u_f + \int_{-\infty}^{\infty} F'(u)\Psi_2(u)du \\ &= u_f + F(\infty)\Psi_2(\infty) - F(-\infty)\Psi_2(-\infty) \\ &\quad - \int_{-\infty}^{\infty} F(u)\Psi_2'(u)du \\ &= u_f + [F(\infty) - F(-u_f)] - F'(-u_f) \\ &\quad \times \int_{-\infty}^{\infty} (u+u_f)\Psi_2'(u)du.\end{aligned} \tag{32}$$

We have truncated the expansion in degrees of $u+u_f$ as $\Psi_2'(u)$ is similar to a δ function near $-u_f$. We used $\Psi_2(\infty) = 1$, $\Psi_2(-\infty) = 0$, and $F(-u_f)\int_{-\infty}^{\infty}\Psi_2'(u)du = F(-u_f)$. Then Eq. (32) gives

$$\langle \ln Z\rangle - u_f \approx \frac{\beta\sqrt{N}}{\sqrt{\pi}}\int_0^\infty dx\int_x^\infty \exp(-y^2)dy \sim N^{1/2}. \quad (33)$$

Equation (33) is one of the main results of our investigation. It is obvious that in addition to the bulk term asymptotic in free energy there is a subdominant term proportional to the square root of the number of degrees. The importance of the subdominant term in the entropy has been underlined in [4], and has been well analyzed in [5]. These authors suggested identifying different universality classes of complex phenomena by the subdominant terms of entropy. In the 1D spin glass model with long-range interaction, $\langle J_{ij}^2\rangle \sim 1/(i-j)^2$, they derived Eq. (33) for the entropy. In [5] another object with a similar subdominant entropy has been mentioned, i.e., language [31].

C. Small deviation from the border of two phases

Consider a small deviation from Eq. (26) [scaling is reasonable, as we see in Eq. (35)]:

$$J_0 = \sqrt{\ln 2} + \frac{j_0}{\sqrt{N}}. \quad (34)$$

Now the finite size correction is less than in Eq. (33) and decreases exponentially at large values of j_0:

$$\langle \ln Z\rangle - \left(\sqrt{\ln 2} + \frac{j_0}{\sqrt{N}}\right)N \sim \beta\sqrt{N}\exp(-j_0^2). \quad (35)$$

Now calculate the magnetization. We define

$$m = \left\langle \frac{\exp(-\beta E_1)}{\sum_\alpha \exp(-\beta E_\alpha)}\right\rangle. \quad (36)$$

Using the identity $1/Z = \int_0^\infty dt\, e^{-tZ}$ we derive for m

$$m = -\int_0^\infty dt\frac{d}{dt}f(u+u_f)f(u)^M = 1 - \int_{-\infty}^\infty du\, f(u+u_f)\frac{d}{du}f^M(u), \quad (37)$$

where $u = \ln t$. Using the second equation in Eq. (26) we derive

$$m = \frac{d\langle \ln Z\rangle}{\beta\sqrt{N}dj_0} = \frac{1}{\sqrt{\pi}}\int_{-j_0}^\infty \exp(-y^2)dy, \quad (38)$$

and for its differential we have

$$\frac{dm}{dj_0} = \frac{1}{\sqrt{\pi}}\exp(-j_0^2). \quad (39)$$

The last expression could be represented also as

$$\frac{1}{\beta\sqrt{N}}\frac{d^2\langle \ln Z\rangle}{d^2 j_0} = \frac{1}{\sqrt{\pi}}\exp(-j_0^2). \quad (40)$$

Thus, at the exact border SG-FM ($j_0=0$) the dependence of the magnetization on the external (ordered) parameter is maximal (maximum instability principle). This is likely a characteristic property of every complex adaptive system. One has an ordered external parameter to manage the system as well as random parameters (the choice of "ordered" and "random" could be subjective). There is an emergent (essentially collective) property. If the interaction with the environment is defined via the emergent property, then the CAS drifts to the maximal instability point with maximal dependence of this emergent property on the ordered parameter.

One can consider dm/dj_0 as some degree of complexity. A close characteristic is the second derivative of the free energy via ordered coupling, Eq. (40). In our case, they coincide. However, there are possibly more complicated situations when they are different and both should be used.

Let us calculate the moments of $P_\alpha \equiv \exp(-\beta E_\alpha)/\sum_\delta \exp(-\beta E_\delta)$. Using the identity

$$\frac{1}{\sqrt{\pi}}\int_{-\infty}^\infty \exp(-y^2 - n\lambda y - e^{u-\lambda y})dy = \frac{d^n f_0(t)}{dt^n},$$

$$f_0(t) \equiv f(\ln t). \quad (41)$$

We have

$$\langle P_1^2\rangle = \int_0^\infty t\, dt\, f_0(t)^{M-1}\frac{d^2 f_0(te^{J_0N\beta})}{dt^2} = \int_{-j_0}^\infty \frac{e^{-x^2}}{\sqrt{\pi}}dx,$$

$$\langle P_\alpha^2\rangle = \int_0^\infty t\, dt\, f_0(t)^{M-2}\frac{d^2 f_0(t)}{dt^2}f_0(te^{J_0N\beta}),$$

$$\langle P_\alpha P_\gamma\rangle = \int_0^\infty t\, dt\, f_0(te^{J_0N\beta})f_0(t)^{M-3}\left(\frac{df_0(t)}{dt}\right)^2,$$

$$\sum_{\alpha,\gamma>1}\langle P_\alpha P_\gamma\rangle = 1 - \frac{1}{\sqrt{\pi}}\int_{-j_0}^\infty \exp(-x^2)dx. \quad (42)$$

For $T < T_c \equiv 2\sqrt{\ln 2}$:

$$\sum_{\alpha>1}\langle P_\alpha^2\rangle = \left(1 - \frac{1}{\sqrt{\pi}}\int_{-j_0}^\infty \exp(-x^2)dx\right)\left(1 - \frac{T}{T_c}\right),$$

$$\sum_{\alpha,\gamma>1,\alpha\neq\gamma}\langle P_\alpha P_\gamma\rangle = \left(1 - \frac{1}{\sqrt{\pi}}\int_{-j_0}^\infty \exp(-x^2)dx\right)\frac{T}{T_c}. \quad (43)$$

Define

$$C = \langle P_1^2\rangle \sum_{\alpha>1}\langle P_\alpha^2\rangle. \quad (44)$$

C takes the maximal value at the critical point $j_0=0$

$$C = \frac{T_c - T}{2T_c}. \quad (45)$$

For large j_0 we have that C decreases exponentially:

$$C \sim \exp(-j_0^2). \quad (46)$$

The more detailed investigation of the $j_0=0$ case states that

$$\langle P_1 \rangle = \frac{1}{2}, \quad \langle P_1^2 \rangle = \frac{1}{2}. \tag{47}$$

We see that $P_1=0$, 1 with probabilities $1/2$.

We can define C as the edge of chaos parameter. At the exact error threshold border it has a maximal value, equal to $1/2$ at zero temperature, i.e., the probabilities of ordered and random motions are equal. C decreases exponentially outside the region. What is the advantage of our choice Eq. (44) over another one, $\langle p_1 \rangle \Sigma_{\alpha>1} \langle P_\alpha \rangle$? Equation (44) distinguishes $\beta \to \infty$ as the optimal situation, and the last choice fails.

To define C, we have actually used the Tsallis entropy at $q=2$ [32]

$$I_q = -\frac{\left(\sum_r p_r^q - 1\right)}{q-1}. \tag{48}$$

In [3] Gell-Mann and Lloyd assumed the connection of I_q with systems at the edge of chaos.

III. DEFINITION OF COMPLEX ADAPTIVE PROPERTY AND UNIVERSALITY CLASSES

A. Definition of complexity

Free energy is the fundamental object in statistical mechanics. The bulk free energy is proportional to the number of particles (spins). It is well known that in the case of some defects on geometrical manifolds (lines, surfaces), in addition to the bulk term in the asymptote expression of free energy, there are subdominant terms proportional to some roots of N. Thus, the subdominant term in the free energy could be identified with the existence of some structures (much more involved than simple geometrical defects) in the system. In our case of the REM, the formulation of the model was homogeneous in space, but we got a square root subdominant term. In a complex system we assume the following hierarchy: bulk motion and some structures above it. The subdominant free energy is related to the structures. If we are interested just in structure, we can ignore the bulk free energy (an analogy in the physics of surfaces: to investigate the surface free energy we certainly miss the bulk energy). Therefore we have the following scheme.

(a) We define the complexity as the subdominant free energy. We have seen that in the case of the error threshold via the REM it scales as the square root of the number of spins. We assume that it is the most important class of complex phenomena, connected with living systems. In the complexity phase the intermediate scale free energy (or entropy, or Kolmogorov complexity) becomes strong, and the subdominant term scales as the square root of the number of degrees, Eqs. (33) and (35).

What do we mean by the intermediate scale? There is a minimal scale (ultraviolet cutoff) and maximal scale (infrared one). The intermediate scale is just their geometric average. In [33] the statistics of heartbeats were investigated. They found that healthy people can be differentiated by the coarse-grained entropy at the intermediate scale, which is coherent with the appearance of the middle scale free energy in our case. Therefore the situation is coherent with the criterion (a). The complexity in our definition is the free energy on a higher hierarchy level (connected with the structures). One should remember that the free energy itself is on the second level of the hierarchy. The energy is on the ground level. Due to thermodynamic motion, only its smaller part is manageable on the macroscopic level (only the free energy could be extracted as mechanical work while changing the global parameters). Therefore, complexity is a level on a hierarchy of the following modalities: energy, free energy, and subdominant free energy. Each higher level is more universal. It is explicit in the quantum field theory approach to critical phenomena [16]. Different renormalization schemes can give different bulk free energies, but the same logarithmic subdominant one. Thus we observe a hierarchy of modalities (a noncategorical statement about reality, see [34]). In principle, the hierarchy could be continued, and at some level life could appear. Our view (rather statistical mechanical than mathematical) is close to the one of Gell-Mann and Lloyd in [3], defining the system complexity as the "length of highly compressed description of its regularities."

Due to the above mentioned hierarchy, the identification of complexity with a subdominant free energy is more universal than the entropy approach of [4,5]. Sometimes the existence of structure could be identified in entropy or Kolmogorov complexity subdominant terms as well. In our case the free energy reveals a huge subdominant term, but not the entropy.

We assume that other features of our toy model are characteristic for complex adaptive systems.

(b) There is an emergent property, maximally unstable under the change of ordered external parameters, Eq. (39). Sometimes it can be characterized as a second derivative of free energy via an ordered parameter.

(c) The probability of ordered and disordered motions should be at the same level [like Eqs. (45) and (47)].

(d) The complex adaptive properties could be exponentially damped in the case of even a small deviation of the ordered parameter.

Let us discuss different complex systems, defining universality classes.

B. Critical theories

We assume that the subdominant term of free energy describes the number of real parameters of the system. In [5] a learning process for a model with finite K parameters was considered and a logarithmic subdominant term, proportional to K, was found. For 2D critical theories we can take either the total effective number of bosonic degrees (conformal charge c), or the number of primary fields as the number of parameters. According to [16], nature has chosen the first one, and the subdominant term of the free energy is proportional to the conformal charge and to the logarithm of the degrees of freedom.

We see that the complex phenomena analyzed in previous sections correspond to another class of universality than the models in [16]. In critical theories [16] magnetization disappears at the transition point (contrary to error threshold case).

Therefore, we admit that complex adaptive systems, while having some scaling (fat tails in markets), could not be described by critical theories.

In 2D percolation indices could be described by conformal field theory. Therefore, the percolation belongs to the class of critical theories. In our classification this situation is as complex as the class [16].

In the spin glass model of the REM Derrida found a logarithmic subdominant free energy. Therefore, the model belongs to the class of [16].

C. Financial markets

One can apply our criteria to financial markets [35]. To analyze the financial time series $y(t)$ (U.S. dollar–German mark exchange rate) the statistics of the price increment $y(t+\tau)-y(t)$ has been considered and the probability density function $p(x,\tau)$ has been constructed from the empirical data. A Fokker-Planck equation, where the role of time is played by $\ln \tau$, has been derived for the last distribution. We see a diffusion in the scale $\ln \tau$ as well as a drift. In [36] the ratio R of the ordered motion of y and the diffusion has been calculated. It is the tail exponent of $y : P(y) \sim \delta y^{-(1+\mu)}$ [36]. In practice, $R=\mu \sim 3$–5. In the situation, when the approach of [35,36] is correct, the more complex situation corresponds to the smaller values of μ. In the case of the error threshold model, considered in this paper, the subdominant term is larger in the region $R \sim 1$. Outside, it decreases exponentially like the one in Eq. (46).

For the markets something like this property can also be observed. There are fundamentalist traders who act in a deterministic way and the noisy ones [37]. In our model, they are similar to ordered and random couplings. In the case of (b) the fundamentalists' number is chosen to have a maximal influence on the market global characteristics. In the usual thermodynamics we have a fundamental notion of temperature, and the equilibrium is possible only when the temperature of different subsystems is the same. Now an edge of chaos parameter for complex adaptive systems is introduced. It is reasonable to assume that in the stable state it should be the same for different parts of the market (for example, for the traders and stocks). In this way it could be possible to predict future catastrophes. One can identify the edge of chaos parameters also by considering a correlation matrix of different stocks. According to the above mentioned data, there are both deterministic and stochastic parts. It is very important to identify the subdominant term in entropy by considering block entropies of financial data.

D. Highly optimized tolerance design

This is the last crucial achievement of complex system theory, related to the robustness of engineering design [38,39]. In the simplest case one considers a forest fire model on a 2D lattice. There are trees at any site of the lattice, and there is a known probability of sparks. As a tree is fired, its nearest neighbors are also fired. One constructs firebreaks (sites without trees) to limit the size of the event (total number of fired trees). The goal is to construct a robust scheme against fire propagation for the given spark probability, using a minimal area of firebreaks. Scaling laws for the distribution of fire events have been found. The situation highly resembles the error threshold case. Actually, in [39] the connection of highly optimized tolerance (HOT) design with source coding has been directly stated. In the error threshold there is also scaling for the mean magnetization $m=1/2+c/\sqrt{N}$ [22]. It has been assumed in [38,39] that self-organized criticality SOC and HOT design are different classes of universality. We can adduce another argument. In a sandpile there is the analog of free energy, the number of recurrent states of the sandpile process. It is the number of spanning trees of the free-fermion model. Therefore the sandpile belongs to the class of universality [16] with central charge $c=-2$ [10]. We have used an important principle: the class of universality of the complex system should be the same in all of its representations. A very interesting feature of HOT design is that it gives a robustness against the originally given distribution of the noise. The robustness is very fragile: there is a large probability for the total crush (great fire) in the case of a change of the original conditions. This resembles property (d) in the definition of complex adaptive phenomenon. In [40] a constraint optimization with limited deviation (COLD) has been suggested design to avoid the large probability of total crush. They also mentioned the first known example of a HOT-design-like situation. It is the classic problem of gambler's ruin: optimizing the total return leads to ruin with probability 1 [41]. For a very complicated complex system with many hierarchies, the full optimization states a single simple principle for management of the system, as in this case the essence of different hierarchies should be the same. Only absolute optimization allows a full transformation of the content from one hierarchy level to another. This crucial feature has been lost in the COLD case. I think that the choice of COLD can be successful only for not too complicated systems. In the next section we will discuss the related concept of antiresonance for complex adaptive systems, exploring the property (d) in our approach.

E. Markov models and random networks

There are a lot of applications of Markov models in complex systems. Especially important are applications in bioinformatics [42]. There is some biological language in DNA and proteins, and hidden Markov models (the transition between states of the system is observed in a probabilistic way) have been applied to model this language. One can investigate the block entropies $S(N)$ for words with N letters in the stream of data and define the subdominant term. Such investigation has been thoroughly done in [6]. At large N, in the case of classic order R, the Markov process $S(N)$ gets an exact linear asymptote at $N>R$. For the case of a hidden Markov model a subdominant entropy, decreasing exponentially with N, has been found. This is very important. Those models, being very useful, do not share the class of universality of living systems, which we assume as corresponding to the subdominant term $\sim\sqrt{N}$.

Networks are very popular in complex system research [43,44]. How can these geometrical objects be classified into universality classes? In [45] has been introduced a statistical mechanics approach to describe the properties of a graph ensemble. The mean characteristics of the graph have been fixed, while maximizing the entropy of the ensemble. Now the number of pairs of vertices plays the role of the number of degrees of freedom. The free energy can be defined. For the case of a random graph there is no finite size correction in free energy expression. Therefore, the random graph corresponds to the Markov model class of complexity. Unfortunately I do not see a way to enlarge the method of [45] to scale-free networks.

F. Virus evolution near the error threshold

The evolution of the majority of viruses (RNA genome viruses) is described well by the Eigen model [27]. This brilliant model gives a simple and complete version of Darwin evolution theory. Information is represented here as a chain of spins taking $\lambda=2$ or $\lambda=4$ values. There are λ^N different configurations with corresponding probabilities p_i, $1 \leqslant i \leqslant \lambda^N$. At any moment, the virus is giving offsprings with some rate specific for his genome (fitness). Offsprings randomly change their mother genome to other ones (mutations). When the majority of individuals has a genome near one configuration ("wild" one), then genetic information is successfully transferred to future generations. Otherwise, there is a flat distribution of individuals in the genome space. It is interesting that the virus evolution is near the error threshold. In "quasispecies theory" [46] (a virus population with a distribution like a cloud around some "wild" genome configuration) there are equivalents of energy, i.e., fitness, and free energy, i.e., mean fitness, for the whole system (for one configuration a product of fitness and errorless copying probability). All of these (selective abilities) can be derived in this model just as a consequence of the Eigen equations. During evolution, the population is located mainly in a genome with high selection ability. Considering the evolution in dynamic environments, it is possible to define a different kind of selective ability, like a higher form of free energy (complexity?). This approach to defining complexity is a quite objective one. We assume that it is possible to calculate analytically also the ground state entropy (including the subdominant one) and define the complexity by [4,5].

It could be possible to investigate some aspects of optimal coding, impossible to do in an alternative way. Choosing as the REM's Hamiltonian a fitnesslike function, we can get analytical dynamics for optimal coding (the work is in progress). Thus, rigorous investigation of informational theoretical (complexity) aspects of evolution models could be very fruitful for both disciplines.

Virus evolution is often referred to as a typical example of a complex adaptive system. Another famous example is the immune system. Statistical mechanics has been successfully applied to this case [47]. I do not see a direct analogy with the error threshold phenomena here. But one should definitely choose a model from a high complexity class.

G. Sherrington-Kirkpatrick model

Usually one defines the logarithm of different ground states [49] (solutions of Thouless-Anderson-Palmer equations) as a complexity. It is a reasonable characteristic to be investigated (although a very complicated one). I think that to identify the universality class of the model it is enough to calculate finite size corrections of the free energy or energy. Such calculations have been done for a Sherrington-Kirkpatrick model [50]. The subdominant energy scales as $N^{1/3}$. Therefore, it is an additional class of complexity. For different spin glasses other subdominant term scalings are possible as well, and finite dimensional spin glasses are likely to have another universality class. We have mentioned the Sherrington-Kirkpatrick model [48], because it is connected with neuron networks.

IV. ANTIRESONANCE IN COMPLEX SYSTEMS

A. Complex resonance

The concept of resonance is probably the most noticeable phenomenon in nature, culture, and science. The close notion of synchronization in complex systems is becoming more and more popular [51]. We are going to analyze the idea of resonance in complex systems, to look for the possibility of, in some sense, the inverse situation with an exponential damping of motion (antiresonance). We suppose that this notion will complement our view of complex systems in the previous section.

Originally, the simplest resonance situation has been investigated in the mechanics of a classical deterministic system with some resonance frequency, driven by an external harmonic force. When two frequencies coincide, the reaction of the system to an external force increases drastically. Even in this simple case we can observe two features of the phenomenon. Frequency is an essence of motion, and there is a sharp peak in the ratio of output to force.

The next step was parametric resonance in classical mechanics. There is a hierarchy here. We observe a motion at given values of parameters, and the resonance frequency depends on the values of external parameters. If one changes the external parameter with the same frequency as the frequency of the pendulum, there appears the famous parametric resonance—the flow of energy from a higher level of the hierarchy to a lower level. Let us generalize this situation to other complex systems to define complex resonance.

If there is a hierarchy in the system, and states at different hierarchic levels have some essence (comparable logically with each other), generalized resonance happens when these essences coincide. What about the essence of the state? In classical mechanics, there is only one real number characterizing the total state, i.e., frequency. In general one should look for other total parameters of the system. In modern physics these are the following: temperature in statistical mechanics, the replica system breaking scheme in spin glasses (edge of chaos parameter), and the wave function phase in quantum mechanics.

The next famous example of such a (generalized parametric resonance) situation is related to the Nishimori line in disordered systems [52,53]. A hierarchy (quenched disorder) is present here. Sometimes it is possible to introduce some formal temperature to describe this disorder. If two temperatures (the real one for the spins and the formal one for the quenched disorder) coincide, the system reveals some interesting properties, becoming maximally analytic in some sense.

So we can define a hierarchy for the resonance. In the trivial case, the system is not hierarchic, it is logically homogeneous. The more involved case corresponds to the situation with very different kinds of motions or (and) hierarchy. It is reasonable to define the second case as a complex resonance. In several situations (i.e., stochastic resonance), when it is impossible to define and compare clearly the essence of a state, one considers a situation when there is a sharp peak in the ratio output-input at an optimal value of the external parameter.

An important moment should be mentioned regarding our concept. If we consider some functional having different parameters, functions, and logical structures, and we optimize it over the entire variables (in addition to some fixed group of parameters or functions), it could be stated that the essence of the whole system is the same as that of a fixed group.

B. Antiresonance

Let me now analyze the resonance situation with the opposite goal: to use the high levels of hierarchy to achieve a maximal negative effect. This is a situation not too rare in living systems.

We define antiresonance as a situation, when (1) a resonance is possible for some value of the external parameter; (2) it is possible to define the opposite phase transformation of the parameter; and (3) at the opposite phase values of the parameter there is either (a) an exponential damping of a motion or (b) a new feature (opposite in some sense to those at the resonant parameter case) arising in a resonant way.

The phenomenon is very complex. Thus, we are investigating the simplest models, trying to reveal those situations in complex systems, when such a phenomenon is possible. Let us consider the pendulum with $x(0)=x_0$,$x'(0)=0$, when the frequency varies with some small amplitude h [54]:

$$\frac{d^2x}{dt^2} = -w^2[1 + h\cos(2wt+\phi)]x. \tag{49}$$

Here $h \ll 1$, and w is a frequency. Taking $\cos(2wt+\phi) = \sin(2wt)$, we get an exponentially amplified solution:

$$x(t) = \exp\left(\frac{hw}{4}t\right)[\cos(wt)]. \tag{50}$$

Choosing $\cos(2wt+\phi) = -\sin(2wt)$, we have an exponential damping

$$x(t) = \exp\left(-\frac{hw}{4}t\right)[\cos(wt)]. \tag{51}$$

For the original amplitude A the damping period T is

$$T \sim \frac{4\ln A}{hw}. \tag{52}$$

In this situation the picture is symmetric (both amplification and damping are possible). The other situation is possible with only resonant damping (like the domino effect).

C. Nishimori line

One considers [52,53] N spins s_i with interaction Hamiltonian

$$H = -\sum_{i_1,\ldots,i_p} j_{i_1,\ldots,i_p} s_{i_1} s_{i_2},\ldots,s_{i_p}. \tag{53}$$

There is a p-spin interaction here; the couplings $j_{i_1,\ldots,i_p}$ are random quenched variables ±1 with probability $(1+m_0)/2$ for the values 1 and $(1-m_0)/2$ for the values -1. It is possible to write the following probability distribution:

$$P(j_{i_1,\ldots,i_p}) = \frac{\exp(\beta_0 J_{i_1,\ldots,i_p})}{2\cosh(\beta_0)}. \tag{54}$$

The parameter β_0 resembles an inverse temperature. Using the invariance of the Hamiltonian under the transformation

$$s_i \to s_i v_i, \quad j_{i_1,\ldots,i_p} \to j_{i_1,\ldots,i_p} v_{i_1} v_{i_2},\ldots,v_{i_p}, \tag{55}$$

in [52,53] the exact energy of the model at $\beta_0=\beta$ has been calculated. At $\beta=\beta_0$ our system has the best ferromagnetic properties in the sense that the number of up spins $\Sigma_i \langle s_i \rangle / |\langle s_i \rangle|$ is maximal at the Nishimori temperature [53]. In the opposite phase, we can take $\beta=-\beta_0$. While the order parameters are different in the ferromagnetic and antiferromagnetic phases, the free energy is the same in both models, as $Z(j,\beta)=Z(j,-\beta)$ for the Hamiltonian (53). For the odd values of p one has optimal properties for the configuration $s_i=-1$. Thus, there is a trivial antiresonance according to our definition. For the even values of p (i.e., $p=2$) and bonds on the links of hypercubic lattices in d-dimensional space, there is an antiferromagnetic ordering (an antiresonance situation).

D. Antiresonance in complex systems

A search for antiresonance in stochastic resonance [55] is a very interesting issue. The resonance is certainly a complex one, when the deterministic harmonic motion has the same period as the transition because of noise. To construct the antiresonance is problematic, as stochastic resonance has no phase to reverse the resonance situation. In [56] a stochastic resonance explanation for the crashes and bubbles in financial markets (using the Ising spin model) was considered. There is no phase for the noise to be reversed in stochastic resonance, but the information for the agents can certainly be positive or negative, thus moving the market from the border of two phases to one side.

During the last decade, the idea of evolution or development at the edge of chaos [7,8], related to complex adaptive systems, was very popular. What about the antiresonance aspect of the origin of life? It is the case, at least, for the hypercycle model by Eigen and Schuster [57]. One tries to construct a self-replicating system that bypasses simple problems, i.e., mutations, but, unavoidably, parasite creatures appear. As a result, there is a chance to consume all the information via those parasite creatures. We see, in some sense, a resonance picture with a chance for antiresonance. Virus evolution is also often near the error threshold (mutation catastrophe) [58]. At the top level of life there is the phenomenon of apostasis, when the cell can be killed by a simple command.

Our point of view is the following: even if complex systems are walking at the edge of chaos and climbing the mountains of fitness landscapes (in case of biological evolution), it is often a walk near the precipice. For evolution it is not so dangerous, as only the survival of the species is crucial. One should be much more careful with a rare or single system, like humanity.

We observe the antiresonance phenomenon in history. It is known from the experience of the history of mankind that in tense situations when a category connected with some symbol becomes urgent (paramount) the defying of that symbol unavoidably leads to the reverse reaction (often during the three years). Very dangerous is a situation when such phenomenon proceeds widely over the world simultaneously in many countries (global antiresonance).

E. Complexity parameters and stability of complex systems

What parameters could be applied to analyze complex systems? In addition to the edge of chaos parameter, reasonable for error-threshold-like systems, we can use the Nishimori temperature as a parameter. In principle, Parisi's replica symmetry scheme also could be considered as a complexity parameter. For the stability of the complex system it is important that those parameters coincide in different subsystems or hierarchial levels.

An interesting example of a complexity parameter in protein physics is the protein design temperature (see the review [59]). Here the Hamiltonian $H(j,s)$ is a function of j_i (amino acid types in a sequence) and s_l (conformations). The couplings j have a distribution like the one in Eq. (54):

$$P(j) \sim \exp[-H(j,s)\beta_d], \tag{56}$$

where s is some ferromagneticlike "native" configuration. Perhaps the methodology of the Nishimori line could be applied to the protein case.

V. CONCLUSION

We have rigorously solved the error threshold for optimal codes using the random energy model, calculating the magnetization and finite size corrections to the free energy. This approach was applied in our previous work where many results of Shannon information theory about optimal coding were derived. There is an alternative method (replica approach with Nishimiori line), working well also in the case of realistic low-density parity check codes [60] (see the review [26]). The REM approach could not be applied directly in the case of finite block coding, but it is much simpler. The main results of information theory were derived in the REM approach about 6–9 years before those found through alternative methods: error threshold for finite rate of information transmission [19] versus [61], reliability exponent [25] versus [62], data compression [23] versus [63]. Multichannel coding was analyzed first in [24]. It is very interesting to check the universality class of codes with finite block length [60], optimal codes with a finite number spin interaction. Unfortunately, the alternative method of [62,63] could not be applied here directly.

Carefully investigating error threshold phenomena in the REM, we have found several criteria of complexity: Eqs. (33), (35), (39), (44), and (46), which could be applied for complex adaptive systems. In [4,5] it was already suggested to consider the subdominant part of the entropy as a measure of complexity. We have enlarged that idea, suggesting the use of a subdominant part of the free energy as a measure of complexity. It is more universal than the bulk free energy, and could be considered as the next step in the hierarchy energy–free energy–subdominant term in free energy. This hierarchy could be continued. Complexity appears on the third level; at some higher levels life could appear. We admit that our approach includes a qualitative idea about the edge of chaos: in the complex phase, the probabilities of ordered and disordered motions are equal [Eq. (44)], and complexity properties damp exponentially outside the error threshold point [Eqs. (35), (39), and (46)]. We adduced arguments that, unlike SOC or ordinary critical theories, HOT design belongs to the error threshold universality class of complexity. There are a few classes of subdominant term behavior: zero or exponentially decreasing subdominant terms for Markov and hidden Markov models [6]; logarithmic corrections for critical theories [16]; cubic root corrections for the Sherrington-Kirkpatrick model; square root corrections for the error threshold, long-range SG model [5] and, maybe, language. First, a complexity class should be identified from the empirical data, to model a complex phenomenon. As percolation or SOC models belong to the same universality class [16], it is improbable that they can describe financial markets. Originally, only SOC criticality was identified with the qualitative idea of "edge of chaos." But we see that the error threshold class is higher than SOC, and this complexity class is likely connected with lifelike systems [7]. We have introduced also the concept of antiresonance, a phenomenon, perhaps, typical for the creation (and existence) of life and for advanced complex adaptive systems.

We have suggested investigating, first, the main features of complexity to identify the large universality classes. What other characteristics could be used for the further characterization of complex phenomena? Perhaps the language of the system with its grammar, or, in physical systems, the existence of local gauge invariance. In the case of the REM, formulated as a spin model, there is a local gauge invariance [see Eq. (55)]. There is local scale invariance for the models of [16]. Therefore the two theories could be connected, according to our complexity analysis. This is really the case, as has been proved in [64]. We hope that other applications of this analysis are possible. The spin glass phase and error threshold border in the REM reveal the advantage of the subdominant free energy approach to complexity compared with the subdominant entropy one. The latter, if used as a complexity measure, produces lower classes [$\sim O(1)$ instead of $\ln N$ or $\sqrt{N}$]. We have used the free energy to define the complexity. In general, when a direct statistical mechanics formulation of the problem is impossible, one can use a variable describing a manageable amount of motion on the macroscopic level. The context of the problem can contribute greatly to making a proper choice. For example, in the Eigen model the equivalent of energy is fitness (with a minus sign). Free energy is automatically defined as a negative selective ability (mean fitness) of a group of configurations.

In Sec. IV the idea of the essence of a complex system state was used several times. In the case of spin glasses, the real state of the system is defined in the replica space with some probability followed by projection to zero replicas (in Parisi's theory). In the case of hidden Markov models the state is not directly observable again, as we get information via probabilistic processes. In quantitative linguistics, an abstract linear space has been applied to catch the meaning of the words [65]. Perhaps the first example is quantum mechanics: there is a unitary evolution of the state in Hilbert space, and during the measurement we have some probabilistic results. In all those examples the state of the complex system is not formulated directly via observables, but instead in some hidden abstract space, where the interpretation of the system (its motion) is rather simple (the formulation of spin glass statistical physics in replica space is much easier than in the zero replica limit, and formulation of the Schrödinger equation is easier than a quantum theory of measurement). We assume that it is an important feature of complex systems: the real state of the system is hidden in abstract space, and can be observed in reality only in a probabilistic way. Therefore we suggest a "principle of expanded prereality:" to solve a complex problem one should reformulate the problem in some internal, hidden, wider space ("prereality"), and then return back to the observable space ("reality") in a probabilistic way.

In view of our results, it is very important to look for antiresonance phenomena in stochastic resonance. Unfortunately, early attempts to find it have not been successful [66]. Another important problem is to identify the universality class of turbulence. An accurate numerical analysis to identify the universality class [67] is likely possible for the case of Burgers turbulence. According to the whole experience of complex systems and our "prereality" principle, to succeed in a turbulence solution one should formulate the problem in a wider abstract space, and then return back to the observable. It is very important to investigate the language models [65] and latent semantic analysis [68,69] in our approach. As we mentioned, the results of [31] (by means of entropy analysis) already support the idea that language belongs to the error threshold class. The investigation of the semantic class is much deeper. The singular value decomposition in [65,68,69] qualitatively resembles the fracturing of couplings into ferromagnetic and noisy ones.

Note added in proof. Recently, J. Crutchfield sent me some of his articles [70–72]. In Ref. [70] the Renyi entropy has been applied to investigate the complexity (much before Ref. [3]). In Ref. [71] the edge of chaos phenomenon has been investigated rigorously in dynamic systems; later the edge of chaos situation has been clarified in cellular automata [72]. This suggests that the names of J. P. Crutchfield and K. Young should be added to the list of the founders of the edge of chaos phenomenon.

ACKNOWLEDGMENTS

I am grateful to C. Biebricher, S. Chen, V. Priezjev, Y. Sinai, D. Saad, and H. Wio for discussion. I greatly appreciate P. Bak's earlier support of my idea to connect complex phenomena with the error threshold. This work was partially supported by the National Science Council of Taiwan under Grant No. NSC 91-2112-M-001-056.

[1] C. H. Bennet, in *Complexity, Entropy and the Physics of Information*, edited by W. H. Zurek (Addison-Wesley, Reading, MA, 1990), p. 215.
[2] M. Gell-Mann, in *Complexity: Metaphors, Models and Reality*, edited by G. Cowan, D. Pines, and D. Meltzer (Addison-Wesley, Reading, MA, 1994), p. 17.
[3] M. Gell-Mann and S. Lloyd, in *Nonextensive Entropy: Interdisciplinary Applications*, edited by M. Gell-Mann and C. Tsallis (Oxford University Press, New York, 2004), p. 387.
[4] P. Grassberger, Int. J. Theor. Phys. **25**, 907 (1986).
[5] W. Bialek, I. Nemenman, and N. Tishby, Neural Comput. **13**, 2409 (2001).
[6] J. P. Crutchfield and D. P. Feldman, e-print cond-mat/0102181.
[7] S. A. Kauffmann, *The Origns of Order, Self-Organization and Selection in Evolution* (Oxford University Press, New York, 1993).
[8] P. Bak, C. Tang, and K. Wisenfield, Phys. Rev. Lett. **59**, 381 (1987).
[9] P. Bak (private communication).
[10] D. Dhar, Phys. Rev. Lett. **67**, 1713 (1987).
[11] I. Chisar and J. Korner, *Information Theory* (Mir, Moscow, 1985).
[12] J. S. Shiner and M. Davison, Phys. Rev. E **59**, 1459 (1999).
[13] J. P. Crutchfield and C. R. Shalizi, Phys. Rev. E **59**, 275 (1999).
[14] E. T. Jayenes, in *Papers on Probability, Statistics and Statistical Physics*, edited by R. D. Rosenkrantz (Reidel, Dordrecht, 1982).
[15] E. J. Chaisson, *Cosmic Evolution* (Harvard University Press, London, 2001).
[16] J. L. Cardy and I. Peschel, Nucl. Phys. B: Field Theory Stat. Syst. **300**[FS22], 377 (1988).
[17] N. Sourlas, Nature (London) **239**, 693 (1989).
[18] B. Derrida, Phys. Rev. B **24**, 2613 (1981).
[19] D. B. Saakian, JETP Lett. **55**, 198 (1992).
[20] P. Rujan, Phys. Rev. Lett. **70**, 2968 (1993).
[21] D. B. Saakian, JETP Lett. **61**, 610 (1995).
[22] A. E. Allahverdyan and D. B. Saakian, Teor. Mat. Fiz. **109**, 1574 (1996).
[23] D. B. Saakian, Teor. Mat. Fiz. **97**, 1199 (1993).
[24] A. E. Allakhverdyan and D. B. Saakian, JETP **111**, 1410 (1997).
[25] A. Allakhverdian and D. Saakian, Nucl. Phys. B **498**, 604 (1997).

[26] Y. Kabashima and D. Saad, J. Phys. A **37**, R1 (2004).
[27] M. Eigen, Naturwissenschaften **58**, 465 (1971).
[28] E. Baake, M. Baake, and H. Wagner, Phys. Rev. Lett. **78**, 559 (1997).
[29] D. B. Saakian and C. K. Hu, Phys. Rev. E **69**, 021913 (2004).
[30] D. B. Saakian, and C. K. Hu, Phys. Rev. E **69**, 046121 (2004).
[31] W. Hilberg, Frequenz **44**, 243 (1990).
[32] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
[33] M. Costa, A. Goldberger, and C.-K. Peng, Phys. Rev. Lett. **89**, 068102 (2002).
[34] R. S. Ingarden, A. Kossakowski, and M. Ohya, *Information Dynamics and Open Systems* (Kluwer, Dordrecht, 1997).
[35] Ch. Renner, J. Peinke, and R. Friedrich, e-print cond-mat/0102494.
[36] D. Sornette, Physica A **290**, 211 (2001).
[37] T. Lux and M. Marchesi, Nature (London) **297**, 498 (1999).
[38] J. Doyle and J. M. Carlson, Phys. Rev. Lett. **84**, 2529 (2000).
[39] J. Doyle and J. M. Carlson, Phys. Rev. Lett. **84**, 5656 (2000).
[40] M. E. Newman, M. Girvan, and J. D. Farmer, Phys. Rev. Lett. **89**, 028301 (2002).
[41] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[42] T. Koski, *Hidden Markov Models for Bioinformatics* (Kluwer, Dordrecht, 2001).
[43] S. H. Strogatz, Nature (London) **410**, 268 (2001).
[44] R. Albert and A. L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).
[45] J. Park and M. E. J. Newman, e-print cond-mat/0405566.
[46] M. Eigen, J. McCaskill, and P. Schuster, Adv. Chem. Phys. **75**, 149 (1989).
[47] A. S. Perelson, Rev. Mod. Phys. **69**, 1219 (1997).
[48] D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).
[49] A. Crisanti, L. Leuzzi, G. Parisi, and T. Rizzo, Phys. Rev. Lett. **92**, 127203 (2004).
[50] G. Parisi, F. Ritort, and F. Slanina, J. Phys. A **26**, 3775 (1993).
[51] A. T. Winfree, Science **298**, 2336 (2002).
[52] H. Nishimori, Prog. Theor. Phys. **66**, 1169 (1981).
[53] H. Nishimori, Physica A **306**, 68 (2002).
[54] L. D. Landau and E. M. Lifshitz, *Classical Mechanics* (Nauka, Moscow, 1973).
[55] R. Benzi, S. Sutera, and A. Vulpiani, J. Phys. A **14**, L453 (1981).
[56] A. Krawiecki and J. A. Holyst, Physica A **317**, 597 (2003).
[57] M. Eigen and P. Schuster, *The Hypercycle-A Principle of Natural Self-Organization* (Springer, Heidelberg, 1979).
[58] M. Eigen, Proc. Natl. Acad. Sci. U.S.A. **99**, 13 374 (2002).
[59] V. S. Pande, A. Yu. Grossberg, and T. Tanaka, Rev. Mod. Phys. **72**, 260 (2000).
[60] D. MacKay and R. Neal, *Cryptography and Coding: 5th IAM Conference*, edited by C. Boyd, Lecture Notes in Computer Science Vol. 1025 (Springer, Berlin, 1995), p. 100.
[61] Y. Kabashima and D. Saad, Europhys. Lett. **45**, 97 (1999).
[62] N. Skantzos, J. Mourik, and D. Saad, J. Phys. A **36**, 11 131 (2003).
[63] T. Murayama, J. Phys. A **35**, L95 (2002).
[64] D. Saakian, Phys. Rev. E **65**, 067104 (2002).
[65] D. Aerts and M. Czachor, J. Phys. A **37**, L123 (2004).
[66] H. Wio (private communication).
[67] Y. Sinai (private communication).
[68] S. Deerwester *et al.*, J. Am. Soc. Inf. Sci. **41**, 391 (1990).
[69] T. K. Landauer, K. P. Foltz, and D. Laham, Discourse Process. **25**, 259 (1998).
[70] J. P. Crutchfield and K. Young, Phys. Rev. Lett. **63**, 105 (1989).
[71] J. P. Crutchfield and K. Young, in *Complexity, Entropy and the Physics of Information*, edited by W. H. Zurek (Addison-Wesley, Reading, MA, 1990).
[72] M. Mitchell, P. T. Hraber, and J. P. Crutchfield, Complex Syst. **7**, 89 (1993).