

Field theory entropy, the H theorem, and the renormalization group

José Gaité

Instituto de Matemáticas y Física Fundamental, Consejo Superior de Investigaciones Científicas, Serrano 123, 28006 Madrid, Spain

Denjoe O'Connor

School of Theoretical Physics, Dublin Institute for Advanced Studies, 10 Burlington Rd., Dublin 4, Ireland

(Received 21 November 1995)

We consider entropy and relative entropy in field theory and establish relevant monotonicity properties with respect to the couplings. The relative entropy in a field theory with a hierarchy of renormalization-group fixed points ranks the fixed points, the lowest relative entropy being assigned to the highest multicritical point. We argue that as a consequence of a generalized H theorem Wilsonian RG flows induce an increase in entropy and propose the relative entropy as the natural quantity which increases from one fixed point to another in more than two dimensions. [S0556-2821(96)04620-6]

PACS number(s): 11.10.Gh, 05.70.Jk, 64.60.Ak, 65.50.+m

I. INTRODUCTION

The concept of entropy was introduced by Clausius through the study of thermodynamical systems. However it was Boltzmann's essential discovery that entropy is the natural quantity that bridges the microscopic and macroscopic descriptions of a system which gave it its modern interpretation. A more general definition, proposed by Gibbs allowed its extension to any system where probability theory plays a role. It is a variant of this entropy which we discuss in a field theoretic context. Boltzmann also defined, in kinetic theory, a quantity H , that decreases with time and for a noninteracting gas coincides with the entropy at equilibrium (H theorem). These ideas also admit generalization and in our context we will see that analogous "nonequilibrium" ideas can be associated with Wilsonian renormalization in our field theory entropic setting.

Probabilistic entropy can be defined for a field theory and in terms of appropriate variables is either a monotonic or convex function of those variables. A variant of it, the relative entropy, is suited to the study of systems where there is a distinguished point as in the case of critical phenomena, where a critical point is distinguished.

We shall see that monotonicity of the relative entropy along lines that depart from the distinguished point in coupling space entails its increase in the crossover from the critical behavior associated with one domain of scale invariance or fixed point to that associated with a "lower" fixed point, thus providing a quantity that naturally "ranks" the fixed points. This property is a consequence of convexity of the appropriate thermodynamic surface, which in turn is reflected in the general structure of the phase diagram [1]. The phase diagrams of lower critical points emerge as projections of the larger phase diagram. We shall see that the natural geometrical setting for these phase diagrams is projective geometry.

There have been many attempts to capture the irreversible nature of a Wilson renormalization group (RG) flow in some function which is intended to be monotonic under the iteration of a Wilson RG transformation [2]. These attempts have been successful in two dimensions where the Zamolodchikov

C function has the desired property. The monotonicity of the flow of the C function under scale transformations is reminiscent of Boltzmann's H function and this result has been accordingly called the C theorem. Boltzmann's H function was the generalization of entropy to nonequilibrium situations, in particular, to a gas with an arbitrary particle distribution in phase space. He proved that H increases whenever the gas evolves to its Maxwell-Boltzmann equilibrium distribution [3], effectively making this evolution an irreversible process. We will argue that an analogue "nonequilibrium" probabilistic entropy for a field theory provides a natural function that must increase under a Wilsonian RG flow. We shall consider a version of the H theorem suited to our needs, to see how the increase occurs. A differential increase along the RG trajectories demands detailed knowledge of the flow lines; however, statements about the ends of the flows are more robust and thus more easily established. It is such statements that we shall establish.

Among other attempts to apply the methods of entropy and irreversibility to quantum field theory, it was shown in [4] that an entropy defined from the quantum particle density, understood as a probability density, should increase as the field theory reaches its classical limit. If we regard this limit as a crossover between different theories, that result should be directly connected to ours. Regarding the connection with two-dimensional conformal field theories and Zamolodchikov's C theorem it is noteworthy that calculations of the geometrical or entanglement entropy (see [5] for background) give a quantity proportional to the central charge c [6]. We will not however pursue possible connections with the entanglement entropy here.

The structure of the paper is as follows: In Sec. II we review the definitions of entropy and relative entropy and adapt them to field theory. We study some of their properties, especially the property of monotonicity with respect to couplings, related with convexity. Section III discusses the crossover of the relative entropy between field theories. We provide some examples, ranging from the trivial crossover, in the Gaussian model as a function of mass, to the tricritical to critical crossover, which illustrates the generic features of this phenomenon. This section ends with a brief study of the

geometric structure of phase diagrams relevant to crossover phenomena. Although Sec. III heavily relies on RG constructs, the picture of the RG used is somewhat simple minded. In Sec. IV we improve on that picture, introducing Wilson's RG ideas. We see how these ideas naturally lead one to interpret crossover from *cutoff-dependent to cutoff-independent* degrees of freedom as an irreversible process in the sense of thermodynamics and therefore to consider a nonequilibrium field theoretic *H*-theorem-type entropy.

II. ENTROPY IN FIELD THEORY, DEFINITION AND PROPERTIES

For a normalized probability distribution \mathcal{P} , we take as our definition of probabilistic entropy,

$$S_a = -\text{Tr} \mathcal{P} \ln \mathcal{P} \tag{2.1}$$

and will refer to this as ‘‘absolute probabilistic entropy.’’ For example, for a single random variable ϕ governed by the normalized Gaussian probability distribution

$$\mathcal{P} = \exp(-\frac{1}{2} m^2 \phi^2 - j \phi + W[j, m^2]), \tag{2.2}$$

where $W[j, m^2] = -j^2/2m^2 + \frac{1}{2} \ln(m^2/2\pi)$ and Tr is understood to mean integration over ϕ . The absolute probabilistic entropy is given by

$$S_a = \frac{1}{2} - \frac{1}{2} \ln \frac{m^2}{2\pi}. \tag{2.3}$$

A natural generalization of this entropy known as the relative entropy [7] is given by

$$S[\mathcal{P}, \mathcal{P}_0] = \text{Tr}[\mathcal{P} \ln(\mathcal{P}/\mathcal{P}_0)], \tag{2.4}$$

where \mathcal{P}_0 specifies the *a priori* probabilities. The sign change relative to Eq. (2.1) is conventional. Relative entropy plays an important role in statistics and the theory of large deviations [8,9]. It is a convex function of \mathcal{P} with $S \geq 0$ and equality applying if and only if $\mathcal{P} = \mathcal{P}_0$. It measures the statistical distance between the probability distributions \mathcal{P} and \mathcal{P}_0 in the sense that the smaller $S[\mathcal{P}, \mathcal{P}_0]$ the harder it is to discriminate between \mathcal{P} and \mathcal{P}_0 . The infinitesimal form of this distance provides a metric known as the Fisher information matrix [10] and provides a curved metric on the space of parametrized probability distributions and the space of couplings in field theory [11]. For example, if we consider the probability distribution (2.2), with $j=0$ for simplicity, the entropy of the Gaussian distribution with standard deviation m^2 relative to the Gaussian distribution with standard deviation m_0^2 is given by

$$S[m^2, m_0^2] = \frac{1}{2} \ln \frac{m^2}{m_0^2} + \frac{m_0^2}{2m^2} - \frac{1}{2} \tag{2.5}$$

and can be easily seen to have the desired properties. By taking the *a priori* probabilities to be given by the uniform distribution we recover Eq. (2.1), modulo a sign. However, we see that Eq. (2.5) approaches Eq. (2.3) but modulo a divergent constant as $m_0 \rightarrow 0$. This reflects the fact that the uniform distribution is not normalizable. The uniform distribution in this setting does not strictly fit the criteria of a

suitable *a priori* distribution \mathcal{P}_0 and therefore violates the assumptions guaranteeing the positivity of the relative entropy. More generally for a continuously distributed random variable a more suitable distribution, with respect to which one can define the *a priori* probabilities, is one that resides in the same function space.

In the case of a field theory Tr will be a path integral over the field configurations and just as when defining the partition function of a field theory an ultraviolet and an infrared regulator are, in general, necessary. Convenient infrared regulators will be to consider a massive field theory in a finite box. It is then convenient to deal with the entropy per unit volume or specific entropy $S = S/V$ where V is the volume of the manifold, \mathcal{M} , on which the field theory is defined. One would generally expect that S would contain divergent contributions as the regulators are removed. However, these contributions disappear in an appropriately defined relative entropy.

For a field theory consider

$$\mathcal{P}_z = \exp(-I^0[\phi, \{\lambda\}] - zI^c[\phi, \{l\}] + W[z, \{\lambda\}, \{l\}]), \tag{2.6}$$

where $W[z, \{\lambda\}, \{l\}] = -\ln Z[z, \{\lambda\}, \{l\}]$, with

$$Z[z, \{\lambda\}, \{l\}] = \int \mathcal{D}[\phi] e^{-I^0[\phi, \{\lambda\}] - zI^c[\phi, \{l\}]}, \tag{2.7}$$

i.e., the total action for the random field variable ϕ is given by $I = I^0[\phi, \{\lambda\}] + zI^c[\phi, \{l\}]$. We have divided the parameters of the theory into two sets: The set $\{\lambda\}$ is the set of coupling constants associated with the fixed distribution \mathcal{P}_0 and $\{l\}$ are those associated with the additional, or crossover, contribution to the action zI^c . The two sets are assumed to be distinct, the set $\{l\}$ may, however, incorporate changes to the couplings of the set $\{\lambda\}$.

We have introduced the variable z primarily for later convenience. For a given functional integral ‘‘measure,’’ associated with integration over a fixed function space (this may be made well defined by fixing, for example, ultraviolet and infrared cutoffs), $W[z, \{\lambda\}, \{l\}]$ reduces to $W^0[\{\lambda\}]$ when $z=0$. With the notation

$$\langle X \rangle = \int \mathcal{D}[\phi] X[\phi] e^{-I^0[\phi, \{\lambda\}] - zI^c[\phi, \{l\}] + W[z, \{\lambda\}, \{l\}]}, \tag{2.8}$$

assuming analyticity in z in the neighborhood of $z=1$, the value of principal interest to us, we have

$$\frac{dW[z, \{\lambda\}, \{l\}]}{dz} = \langle I^c \rangle, \tag{2.9}$$

and more generally

$$\frac{d\langle X \rangle}{dz} = -(\langle XI^c \rangle - \langle X \rangle \langle I^c \rangle).$$

We can therefore express the relative entropy as

$$S[z, \{\lambda\}, \{l\}] = W[z, \{\lambda\}, \{l\}] - W^0[\{\lambda\}] - z\langle I^c[\phi, \{l\}] \rangle. \tag{2.10}$$

It is the Legendre transform with respect to z of $W^c = W - W^0$:

$$\mathcal{S}[z, \{\lambda\}, \{l\}] = W^c[z, \{\lambda\}, \{l\}] - z \frac{dW^c[z, \{\lambda\}, \{l\}]}{dz}. \quad (2.11)$$

Next consider the derivative with respect to z of \mathcal{S} :

$$\frac{d\mathcal{S}[z, \{\lambda\}, \{l\}]}{dz} = -z \frac{d^2W[z, \{\lambda\}, \{l\}]}{dz^2}. \quad (2.12)$$

Reexpressing this in terms of expectation values we have

$$z \frac{d\mathcal{S}[z, \{\lambda\}, \{l\}]}{dz} = z^2 \langle (I^c - \langle I^c \rangle)^2 \rangle \quad (2.13)$$

implying that \mathcal{S} is a monotonic increasing function of $|z|$ which is zero at $z=0$. We also deduce from Eqs. (2.12) and (2.13) that W is a convex function of z .

Note that the expression (2.11) is amenable to standard treatment by field theoretic means. In perturbation theory, it is diagrammatically a sum of connected vacuum graphs. Furthermore, if the action is a linear combination of terms

$$I^c[\phi, \{l\}] = l^a f_a[\phi] \quad (2.14)$$

then with $z l^a = t^a$ (z is an overall factor) we have

$$\mathcal{S}[\{\lambda\}, \{t\}] = W[\{\lambda\}, \{t\}] - W[0] - t^a \partial_a W[\{\lambda\}, \{t\}], \quad (2.15)$$

where $\partial_a = \partial / \partial t^a$. Thus for this situation the relative entropy of the field theory is the complete Legendre transform of the generating function W with respect to all the couplings t^a . The negative of the ‘‘absolute’’ entropy or entropy relative to the uniform distribution (equivalent to $I^0[\phi, \{\lambda\}] = 0$) would be the complete Legendre transform with respect to all the couplings in such a field theory. In terms of its natural variables $\langle f_a \rangle = \partial_a W$ the relative entropy itself is a convex function (see below). It proves useful in what follows to regard it as a function of the couplings through $\langle f_a \rangle(t)$.

Let us consider the change in relative entropy due to an infinitesimal change in the couplings of the theory. This can be expressed as a one-form on the space of couplings. A little rearrangement shows that such a change can be expressed in the form

$$d\mathcal{S} = z(d\langle I^c \rangle - \langle dI^c \rangle) \quad (2.16)$$

which implies that z^{-1} performs the role of an integrating factor for the difference of infinitesimals $d\langle I^c \rangle - \langle dI^c \rangle$, just as temperature does for the absolute entropy. We could more generally consider different z 's for each of the composite operators $f_a[\phi]$ and obtain the generalization of (2.16):

$$d\mathcal{S} = \sum_a Z_{f_a} (d\langle f_a[\phi] \rangle - \langle df_a[\phi] \rangle).$$

In renormalization theory the Z_{f_a} play the role of composite operator renormalizations (e.g., $l^a f_a[\phi] = \frac{1}{2} \int t \phi^2$ the composite operator ϕ^2 gets renormalized by Z_{ϕ^2}). Thus one

could interpret composite operator renormalization factors Z_{f_a} (or in the example Z_{ϕ^2}) as integrating factors.

Again for the case (2.14), since

$$z^2 \langle (I^c - \langle I^c \rangle)^2 \rangle = t^a \langle (f_a - \langle f_a \rangle)(f_b - \langle f_b \rangle) \rangle t^b \quad (2.17)$$

and each of the l^a are arbitrary, we see that the quadratic form

$$Q_{ab} = \langle (f_a - \langle f_a \rangle)(f_b - \langle f_b \rangle) \rangle = - \frac{\partial^2 W}{\partial t^a \partial t^b} \quad (2.18)$$

is a positive definite matrix. This establishes the key property that W is a convex function of the couplings. \mathcal{S} is similarly a convex function of the $\langle f_a \rangle$, since

$$Q^{ab} = Q_{ab}^{-1} = \frac{\partial^2 \mathcal{S}}{\partial \langle f_a \rangle \partial \langle f_b \rangle}. \quad (2.19)$$

The matrix Q_{ab} is the Fisher information matrix and plays the role of a natural metric on the space of couplings $\{l\}$ measuring the infinitesimal distance between probability distributions.

We end this section by emphasizing that in the above we have established that W is a convex function of the l^a and \mathcal{S} is a convex function of the $\langle f_a \rangle$. Note that *the usual effective action can be viewed as the relative entropy* with $z I^c[\phi, \{l\}] = \int_{\mathcal{M}} J \phi$ and is therefore a convex function of $\langle \phi \rangle$. The relative entropy is equivalently a generalization of the effective action to a more general setting. A final observation is that the relations

$$\bar{f}_a = \langle f_a \rangle = \partial_a W(t) \quad (2.20)$$

are our field equations (on-shell conditions) and can be associated with equilibrium. If one releases these constraints by, for example, leaving the equilibrium setting, one can consider \mathcal{S} as a function of both the \bar{f}_a and l^a . The equilibrium conditions are then specified by Eq. (2.20).

III. CROSSOVER BETWEEN FIELD THEORIES

The concept of crossover arises in the physics of phase transitions, where it means the change from one type of critical behavior to another. This implies a change of critical exponents or any other quantity associated with critical behavior. In our context, a field theory (FT) is defined by a Lagrangian with a number of coupling constants. We will restrict our considerations to the case of superrenormalizable theories, in which case the theories can be taken to provide well-defined microscopic theories. The Lagrangian captures the universality class of a particular phase transition when the relevant couplings are tuned to appropriate values; these relevant couplings constitute a parametrization of the space of fields and couplings close to the associated fixed point (FP) of the RG. The functional integral provides global information, which can be depicted in a phase diagram, with variables W , $\{l\}$. The most unstable FP will therefore have the largest dimensional phase diagram and far from this FP may exist another where one (or more) of the maximal set of

couplings becomes irrelevant¹ and drops out. This implies the change to a universality class with fewer relevant couplings, hence a reduced phase diagram corresponding to projecting out the couplings which became irrelevant. The second FP and the reduced phase diagram define a new field theory.

It is fairly easy to see that in the region where homogeneous scaling holds and the RG trajectories satisfy linear RG equations there can be no more fixed points. One can define new coordinates called nonlinear scaling fields [12] where homogeneous scaling applies throughout the phase diagram. This possibility is also well known in the theory of ordinary differential equations (ODE's), where it is called Poincaré's theorem [13, p. 175]. In these coordinates, then, any other FP must be placed at infinity in a coordinate system adapted to the first FP. To study the crossover, when a FP is at infinity, we need to perform some kind of compactification of the phase diagram. Thus, we shall think of the total phase diagram as a compact manifold containing the maximum number of *generic* RG FP's. This point of view is especially sensible regarding the topological nature of RG flows. Furthermore, thinking of the RG as just an ODE indicates what type of compactification of phase diagrams is adequate: It is known in the theory of ODE's that the analysis of the flow at infinity and its possible singularities can be done by completing the affine space to projective space [14]. This as we shall see is also appropriate for phase diagrams.

We will restrict our considerations in what follows to scalar Z_2 symmetric field theories with polynomial potentials

and nonsymmetry breaking fields. For illustration, we will discuss some exact results pertaining to solvable statistical models, which illuminate the behavior of the field theories in the same universality classes.

A. Case (0): The Gaussian model and the zero to infinite mass crossover

Consider the action

$$I_0[\phi, \{\lambda(0)\}] = \int_{\mathcal{M}} \left\{ \frac{\alpha}{2} (\partial\phi)^2 + \frac{r_c}{2} \phi^2 \right\}. \quad (3.1)$$

The action associated with \mathcal{P}_z is then

$$I_0[\phi, \{\lambda(0)\}, t] = I_0[\phi, \{\lambda(0)\}] + \int_{\mathcal{M}} \frac{t}{2} \phi^2. \quad (3.2)$$

The crossover here is that associated with $z=t$. The model is pathological in that it is not well defined for $t < 0$ where there is no ground state, but our interest is in $t \geq 0$. The crossover of interest here is then from $t=0$ to large values of t . To make the model completely well defined we place it on a lattice and take the continuum limit.

For the Gaussian model on a square lattice with lattice spacing, taken for simplicity to be $a\sqrt{\alpha}$, and with periodic boundary conditions and sides of length $L = Ka\sqrt{\alpha}$, in d dimensions, we have, in the thermodynamic limit $K \rightarrow \infty$ [15],

$$W[a, r] = \frac{K^d}{2} \int_{-\pi}^{\pi} \frac{d\omega_1}{2\pi} \cdots \int_{-\pi}^{\pi} \frac{d\omega_d}{2\pi} \ln \left\{ \frac{(4/a^2)\sin^2(\omega_1/2) + \cdots + (4/a^2)\sin^2(\omega_d/2) + r_c + t}{2\pi} \right\}. \quad (3.3)$$

With the critical point of the model at $t=0$ we have $r_c=0$. The relative entropy is

$$S[a, t] = W[a, t] - W[a, 0] - t \frac{dW[a, t]}{dt} \quad (3.4)$$

so if $W[a, t]$ took the form $W[a, t] = \tilde{W}[a, t] + c + bt$ the linear term $c + bt$ would not contribute to the relative entropy. In the thermodynamic limit, if we restrict our considerations to a dimensionally regularized continuum model then for $d < 4$ the divergences that require subtraction are indeed of the linear form and we find that the relative entropy per unit volume is given by

$$S = \frac{(d-2)\pi}{2 \sin(\pi(d+2)/2) \Gamma[(d+2)/2] (4\pi)^{d/2}} t^{d/2}. \quad (3.5)$$

¹Here relevant and irrelevant have both their intuitive and RG meaning.

For $d > 2$ and sufficiently small t , in the neighborhood of the critical point, the relative entropy of both the continuum model and the lattice model agree. This can be seen by noting that the second derivative of W with respect to t diverges for small t and, for $d < 4$, the coefficient of divergence is the same for both the lattice and continuum expressions. Thus integrating back to obtain $W[t]$ will give expressions which differ by only a linear term in t for small t but this does not affect the relative entropy. From Eq. (3.5) the increase in relative entropy with t is manifest.

B. Case (i): The Ising universality class

Let us next consider the two-dimensional Ising model on a rectangular lattice. For simplicity we will restrict our considerations to equal couplings in the different directions. Since the random variables here (the Ising spins) take discrete values it is natural to consider the absolute entropy which corresponds to choosing entropy relative to the discrete counting measure and a sign change. This is the standard absolute entropy in this case. This model, as is well known, admits an exact solution [16] for the partition function with

$$W[k] = -\frac{1}{2} \ln[2 \sinh(2k)] - \int_0^\pi \frac{d\omega}{2\pi} \operatorname{arccosh}\{\cosh(2k)\cosh[2K(k)] - \cos(\omega)\} \quad (3.6)$$

for a rectangular lattice where $K(k) = \frac{1}{2} \ln \coth(k)$ and $k = J/k_B T$. The entropy is then

$$S_a = - \left(W(k) - k \frac{dW(k)}{dk} \right) \quad (3.7)$$

and plotted against k in Fig. 1(a). The monotonicity property of the entropy becomes one of convexity when the entropy is expressed in terms of the internal energy U as can be seen in Fig. 1(b).

Now, of course, we can also consider relative entropy in this setting. Since near its critical point the two-dimensional Ising model is in the universality class of a ϕ^4 field theory, to facilitate comparison with the field theory it is natural to choose an entropy relative to the critical point lattice Ising model. This is also natural since the critical point is a preferred point in the model. This relative entropy is given by

$$S = W(k) - W(k^*) - (k - k^*) \frac{dW(k)}{dk}, \quad (3.8)$$

where $k^* = \frac{1}{2} \ln(\sqrt{2} + 1) \sim 0.4406868$ is the critical coupling of the Ising model. We have plotted this in Fig. 2(a). We see that it is a monotonic increasing function of $|k - k^*|$ and is

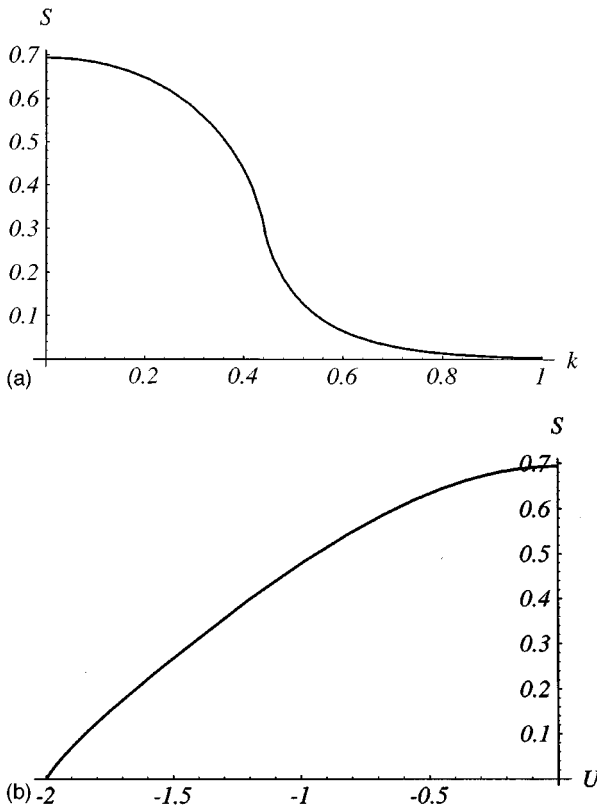


FIG. 1. (a) The entropy $S_a(k)$ for the two-dimensional (2D) Ising model. (b) The entropy $S_a(U)$ for the 2D Ising model.

zero at the critical point. In Fig. 2(b) we plot this entropy as a function of the relevant expectation value, the internal energy $U = dW/dk$, and set the origin at U^* , the internal energy at the critical point. Naturally, the graph is convex.

In more than two dimensions the Ising model has not been solved exactly. Its critical behavior is in the universality class of a ϕ^4 field theory, so we expect the general features of the two models to merge near the critical point. We will next consider the ϕ^4 theory.

We will choose the fixed probability distribution \mathcal{P}_0 for the ϕ^4 theory to be that associated with the critical point, or massless theory, which is described by the action

$$I_1^0[\phi, \{\lambda(1)\}] = \int_{\mathcal{M}} \left\{ \frac{\alpha}{2} (\partial\phi)^2 + \frac{r_c}{2} \phi^2 + \frac{\lambda}{4!} \phi^4 \right\} \quad (3.9)$$

with λ some arbitrary but fixed value of the bare coupling constant. We restrict our considerations to $d < 4$ where the theory is superrenormalizable. The parameter r_c depends on the cutoff (UV regulator) needed to render the theory at a path-integral level well defined, and is chosen such that the correlation length is infinite. The complete action associated with \mathcal{P}_z is

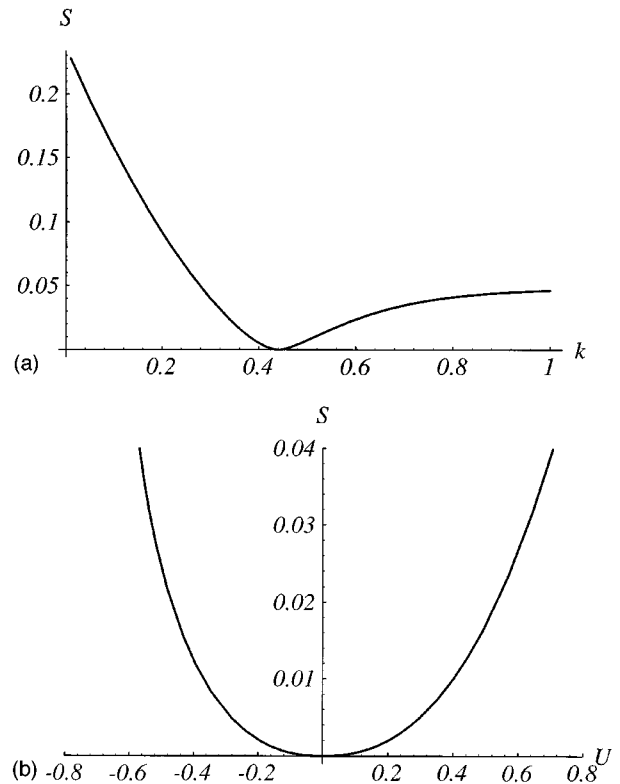


FIG. 2. (a) The relative entropy $S(k, k^*)$ for the 2D Ising model. (b) The relative entropy $S(U, U^*)$ for the 2D Ising model.

$$I_1[\phi, \{\lambda(1)\}, t] = I_1^0[\phi, \{\lambda(1)\}] + \int_{\mathcal{M}} \frac{t}{2} \phi^2. \quad (3.10)$$

The crossover of interest here is that associated with $z = |t|$. There are clearly two branches to the crossover, that for t positive and negative, respectively. We will restrict our considerations to the positive branch, corresponding to $\langle \phi \rangle = 0$, and the range of t is from 0 to ∞ . The identification of z with t allows us to use the arguments of the previous section. From Eq. (2.13) we conclude that the relative entropy is a monotonic function along this crossover line. This is the crossover line from the Wilson Fisher fixed point to the infinite mass Gaussian fixed point.

In the presence of a fixed UV cutoff one could consider the reference probability distribution to be that for which $\lambda = 0$ and then place λ into the crossover portion of the action. This provides us with another crossover and in this more complicated phase diagram there are in fact two Gaussian fixed points; a massless and infinite mass one, both associated with $\lambda = 0$ (see [17] for a description of the total phase diagram). The crossover between them is that associated with ‘‘case (0)’’ described above. If one further restricts to $\lambda = \infty$, this is equivalent to restricting to the fixed point coupling and is believed to be equivalent to the Ising model in the scaling region. The parameters t and k then should play equivalent roles, and describe the same crossover. In the ϕ^4 model one can further consider crossovers associated with varying λ at fixed t , by including a term $\int_{\mathcal{M}} (l/4!) \phi^4$ in I^c . In this family there will be a crossover curve at infinity which varies from one infinite mass Gaussian fixed point to another. Such crossovers can be viewed as a special case of the next example.

C. Case (ii): Models with two crossover parameters

Here the action for the fixed distribution from which we calculate the relative entropy is taken to be

$$I_2^0[\phi, \{\lambda(2)\}] = \int_{\mathcal{M}} \left\{ \frac{\alpha}{2} (\partial\phi)^2 + \frac{r_{1c}}{2} \phi^2 + \frac{\lambda_{1c}}{4!} \phi^4 + \frac{g}{6!} \phi^6 \right\} \quad (3.11)$$

(g fixed) and the action of the model is

$$I[\phi, \{\lambda(2)\}, t, l] = I_2^0[\phi, \{\lambda(2)\}] + \int_{\mathcal{M}} \left\{ \frac{t}{2} \phi^2 + \frac{l}{4!} \phi^4 \right\}. \quad (3.12)$$

The tricritical point corresponds to both t and l zero. There is now a plane to be considered. First consider the line formed setting $l = 0$ and ranging t from zero to infinity. This is a line leaving the tricritical point and going to an infinite mass Gaussian model. Again we see from the arguments of the previous section that the relative entropy is a monotonic function along this line. Similarly we can consider the line $t = 0$ and l ranging through different values. Again for positive l the relative entropy is a monotonic function of this variable. The critical line is a curve in this plane, since the critical temperature T_c should depend on l and one needs to change t as a function of l to track it.

It is interesting to consider the reduction of the two-dimensional phase diagram associated with the neighborhood of the tricritical point to the one-dimensional phase diagram of the critical point. This latter fixed point is associated with $l = \infty$ and the crossover from it to the infinite mass Gaussian fixed point at $t = \infty$ lies completely at infinity in the tricritical phase diagram. In the previous setting the crossover started from a finite location because we did not include the tricritical point. The reduction can be achieved as a projection from the tricritical phase diagram as follows: For any value of (t, l) we can let both go to infinity while keeping their ratio constant. The value of t/l parametrizes points on the line at infinity. Moreover, that projection is realized by letting z run to infinity, thus ensuring that the relative entropy increases in the process.

One can further appreciate the structure of the phase diagram commented on above in terms of the shape of RG trajectories, identified with scaling the nonlinear scaling field, where the phase diagram is presented in these coordinates. In the present case, the family of scaling curves is $t = cl^\varphi$ for various c , with only one parameter given by the ratio of scaling dimensions of the relevant fields $\varphi = \Delta_l/\Delta_t > 1$, called the crossover exponent. These curves have the property that they are all tangent to the t axis at the origin and any straight line $t = al$ intersects them at some finite point, $l_i = (a/c)^{1/(\varphi-1)}$ and $t_i = al_i$. For any given c the values of l_i and t_i increase as a decreases and go to infinity as $a \rightarrow 0$. This clearly shows that the stable fixed point of the flow is on the line at infinity and, in particular, its projective coordinate is $a = 0$. The point $a = \infty$ on the line at infinity is also fixed but unstable. In general, as the overall factor z is taken to infinity we shall hit some point on the separatrix connecting these two points at infinity.

The tricritical flow diagram that includes the separatrix can be obtained by a projective transformation (see Sec. III E). It is essentially of the same form as that considered by Nicoll, Chang, and Stanley [17], with the axes such that the tricritical point is at the origin (Fig. 3). The critical line is the vertical line (the l axis), and the crossover to the Gaussian fixed point which is the most stable fixed point is the line at infinity, in the positive quadrant of the (t, l) plane. The Gaussian fixed point is at the end of the horizontal t axis. Our variable z will parametrize radial lines in this (t, l) plane. As far as the parameter α is concerned, one could introduce another axis in the phase diagram, corresponding to this variable. This can be done for every crossover, and corresponds to crossover as the momentum is varied.

D. The general case of many crossovers

The question arises as to the *naturalness* of the choice of *a priori* distribution \mathcal{P}_0 . In the case of Z_2 models in dimension $4 > d > 2$ there is a natural choice for \mathcal{P}_0 . It is that field theory with the maximum polynomial potential that is superrenormalizable in this dimension. This theory admits the maximum number of nontrivial universal crossovers in this dimension. For this range of dimensions we, therefore, choose

$$I_k^0[\phi, \{\lambda\}] = \int_{\mathcal{M}} \left\{ \frac{\alpha}{2} (\partial\phi)^2 + \sum_{a=1}^{k+1} \frac{\lambda_{2a}}{(2a)!} \phi^a \right\} \quad (3.13)$$

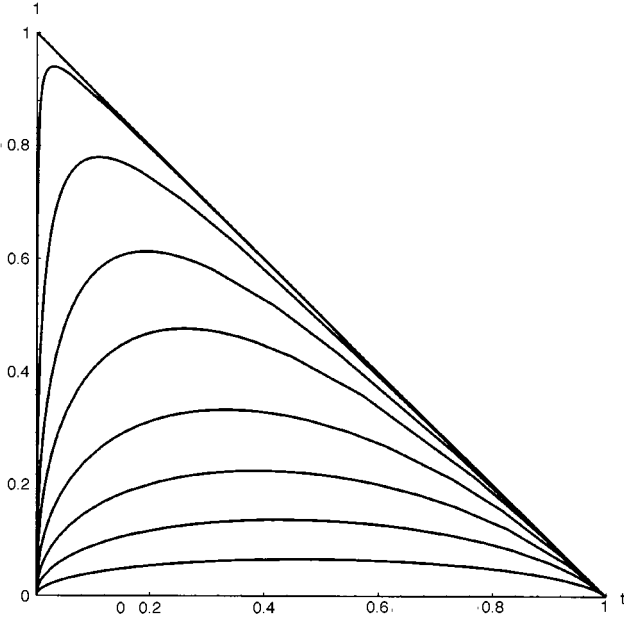


FIG. 3. Tricritical flow diagram showing the tricritical, critical, and Gaussian FP (with the mean-field crossover exponent $\varphi=2$).

and the full action is then

$$I_k[\phi, \{\lambda\}, l_2, \dots, l_{2k}] = I_k^0[\phi, \{\lambda\}] + \int_{\mathcal{M}} \left\{ \sum_{n=1}^k \frac{l_{2n}}{(2n)!} \phi^{2n} \right\}. \quad (3.14)$$

The different crossover lines from the multicritical point can then be arranged to correspond to flows from the origin along straight lines (in particular, the coordinate axes). From the general arguments of the previous section the relative entropy increases along those trajectories.

The crossovers in the above system can be organized in a natural hierarchical sequence, descending from any one multicritical fixed point to the one just below in order of criticality. In this way one loses one irrelevant coupling at each step. The reduced phase diagram at each step is the hyperplane at infinity of the previous diagram. Thus with our compactification they constitute a sequence of nested projective spaces, ending in a point. This structure deserves more detailed treatment.

E. The geometrical structure of the phase diagram

The phase diagrams for the critical models corresponding to different RG fixed points are nested in a natural way as projective spaces,

$$RP_k \supset RP_{k-1} \supset \dots \supset RP_1 \supset RP_0,$$

with RP_0 being just a point that represents the infinite mass Gaussian fixed point. In the action (3.14) the set of couplings l_{2n} together with the coupling λ_{2k+2} lend themselves to an interpretation as homogeneous coordinates for the projective space RP_k . The value of λ_{2k+2} is to be held fixed along any crossover so that the ratios $r_{2n} = l_{2n}/\lambda_{2k+2}$ become affine coordinates. Moreover, in the crossover from an upper criti-

cal point to a lower critical point, e.g., the tricritical to critical crossover, the phase diagram for the latter is realized as the codimension-one (hyper)plane at infinity, which is equivalent to $\lambda_{2k+2}=0$. Thus λ_{2k+2} effectively disappears from the action of the next critical point, which has l_{2k} as the highest coupling in the sequence. The set of couplings l_2, \dots, l_{2k} then constitute a system of homogeneous coordinates in the reduced phase diagram. One can reach a point of this phase diagram by making z go to infinity for different (fixed) values of l_{2i}/l_{2k} . This realization ensures that the relative entropy of points in this second phase diagram is lower than that of points of the first via monotonicity in z as discussed earlier.

One might, however, think that both phase diagrams cannot be incorporated in the same picture. This is not so: One can perform a projective change of coordinates so as to bring the (hyper)plane at infinity to a finite distance. This can be achieved by first rescaling to $\lambda_{2k+2}=1$. For example, in the tricritical to critical crossover of Sec. III B, the condition that g be fixed (e.g., $g=1$ where we now use dimensionless couplings, the original g , which we now label g_B , setting the scale) places the phase diagram of the critical fixed point at infinity. However, new homogeneous coordinates \bar{r} and $\bar{\lambda}$ and \bar{g} , defined so that the projective space is realized as the plane $r+\lambda+g=1$ rather than by $g=1$ can be specified by defining

$$\begin{aligned} \bar{r} &= r, \\ \bar{\lambda} &= \lambda, \\ \bar{g} &= r + \lambda + g. \end{aligned} \quad (3.15)$$

In these coordinates our previous ratios, that is, the affine coordinates, take the form

$$\begin{aligned} \frac{r}{g} &= \frac{\bar{r}/\bar{g}}{1 - \bar{r}/\bar{g} - \bar{\lambda}/\bar{g}}, \\ \frac{\lambda}{g} &= \frac{\bar{\lambda}/\bar{g}}{1 - \bar{r}/\bar{g} - \bar{\lambda}/\bar{g}}. \end{aligned} \quad (3.16)$$

The phase diagram in the new coordinates, drawn in Fig. 3, is patently compact. Transformations of the this type have been used before in global studies of the RG [17]. Another possible realization of the phase diagram would be to project onto the plane $\lambda+g=1$. The new coordinates are given by

$$\begin{aligned} \frac{r}{g} &= \frac{\bar{r}/\bar{g}}{1 - \bar{\lambda}/\bar{g}}, \\ \frac{\lambda}{g} &= \frac{\bar{\lambda}/\bar{g}}{1 - \bar{\lambda}/\bar{g}}. \end{aligned} \quad (3.17)$$

The resulting projective coordinate change converts the line at infinity into the line $\lambda=1$. The critical fixed point is on this line at $r=0$ but the infinite mass Gaussian point remains at $r=\infty$. Hence we can identify the resulting phase diagram as that of the critical model. Similar considerations apply quite generally to the entire hierarchy.

We see that the new ratios in Eq. (3.17) resemble the solution of typical one-loop RG equations. This is not necessarily accidental. In practice when one goes from bare to renormalized coordinates one defines the new coordinates in terms of normalization conditions [18], which can be chosen so that the range of these renormalized coordinates ranges over a finite domain, e.g., from zero to the fixed point value of the renormalized coupling. For example, in the ϕ^4 model the relation between bare and renormalized couplings at one loop is given by

$$\lambda_b = \frac{\lambda_r}{1 - a(d)\lambda_r R^{4-d}}$$

with R the IR cutoff and $a(d)$ a dimension-dependent factor. In terms of the dimensionless couplings $a(d)\lambda R^{4-d}$ we have precisely Eq. (3.17). However, at higher order in the loop expansion such normalization conditions may realize the projective space of the phase diagram in a more complicated fashion than Eq. (3.17). Nevertheless, one can think of the change from “bare” to renormalized coordinates as the transition from affine coordinates to a realization of the projective space.

IV. WILSON'S RG AND ENTROPY GROWTH

Field theoretic renormalization groups that are based on reparametrization of the couplings are a powerful tool for the study of crossovers and the calculation of crossover scaling functions, as discussed in [18]. In essence they can be viewed as implementing appropriate projective changes of coordinates implied by the above discussion. We now wish to discuss the relative entropy in a Wilsonian context. A Wilson RG transformation is such that it eliminates degrees of freedom of short wavelength and hence high energy. Typical examples are decimation or block spin transformations. It is intuitively clear that their action discards information on the system and therefore must produce an increase of entropy. Indeed, as remarked by Ma [19] iterating this type of transformation does not constitute a group but rather a semigroup, since the process cannot be uniquely reversed. In the language of statistical mechanics we can think of it as an irreversible process.

For concreteness we illustrate our approach by a very simple example, the Gaussian model with action

$$I = \frac{1}{2} \int_0^\Lambda d^d p \phi(p)(p^2 + r)\phi(-p), \quad (4.1)$$

which yields

$$W[z] = \frac{1}{2} \int_0^\Lambda \frac{d^d p}{(2\pi)^d} \ln \frac{p^2 + r}{\Lambda^2}. \quad (4.2)$$

This model has already been considered in Sec. III A but with a lattice cutoff instead of a momentum cutoff. The relevant coupling that effects the crossover is $z = t = r - r_c$. The corresponding relative entropy

$$S[z] = \frac{1}{2} \int_0^\Lambda \frac{d^d p}{(2\pi)^d} \left(\ln \frac{p^2 + r}{p^2 + r_c} - \frac{t}{p^2 + r} \right) \quad (4.3)$$

is finite when Λ goes to infinity, agreeing with Eq. (3.5), and vanishes for $t=0$. The Wilson RG is implemented by letting Λ run to lower values. Let us see that S is monotonic with Λ .

We have that

$$\frac{\partial S}{\partial \Lambda} = \frac{\Lambda^{d-1}}{2^d \pi^{d/2} \Gamma(d/2)} \left(\ln \frac{\Lambda^2 + r}{\Lambda^2 + r_c} - \frac{t}{\Lambda^2 + r} \right). \quad (4.4)$$

With the change of variable $x = \Lambda^2$, we have to show that the corresponding function of x is of the same sign everywhere. Then we want

$$\ln \frac{x+r}{x+r_c} - \frac{r-r_c}{x+r}$$

not to change sign. Interestingly, the properties of this expression are independent of x somehow for if one substitutes in $\ln \rho - (\rho-1)/\rho$ the value $\rho = (x+r)/(x+r_c)$ then one recovers the entire function. Now it is easy to show that $\ln \rho \geq 1 - 1/\rho$. (The equality holds for $\rho=1$ —the critical point.) This proof resembles the classical proofs of H theorems.

We plot in Fig. 4 the associated relative entropy for this model as a function of Λ to show that it is again a monotonic function. This behavior is actually closely related to the monotonicity with r considered before: The relative entropy as well as W is a function of the ratio r/Λ^2 , which is precisely the solution of the RG for this simple model.

There are certain features common to all formulations of Wilsonian RG's for a generic model. Even if the theory is simple at the scale of the cutoff, as may happen when we use a lattice model as our regularized theory, a Wilson RG transformation complicates it by introducing new couplings. Thus the action of Wilson's RG is defined in what is called theory space, typically of infinite dimension, comprising all possible theories generated by its action. In practice, one is interested in the critical behavior controlled by a given fixed point and the theory space reduces to the corresponding space spanned by the marginal and relevant operators. Under the action of the RG, the irrelevant coupling constants approach values which are functions of the relevant coupling constants. In the language of differential geometry, the RG flow converges to a manifold parametrized by the relevant couplings. Therefore, the information about the original trajectory or the value of the couplings at the scale of the cutoff is lost. In the language of FT, we can say that the nonrenormalizable couplings vanish (or, in general, approach predetermined values) when the cutoff is removed [20].

As described above, the action of the Wilson RG is reminiscent of the course of a typical nonequilibrium process in statistical physics. The initial state may be set up to be simple but if it is not in equilibrium then it evolves, getting increasingly complicated until an equilibrium state is reached, where the system can be described by a small number of thermodynamic variables. This idea can be formulated as Boltzmann's H theorem. In the modern version of this theorem [21] H is a function(al) of the probability distribution of the system defined as $H = -S_a$ of Eq. (2.1). It measures the information available to the system and has to be a minimum at equilibrium. To be precise, the actual probabil-

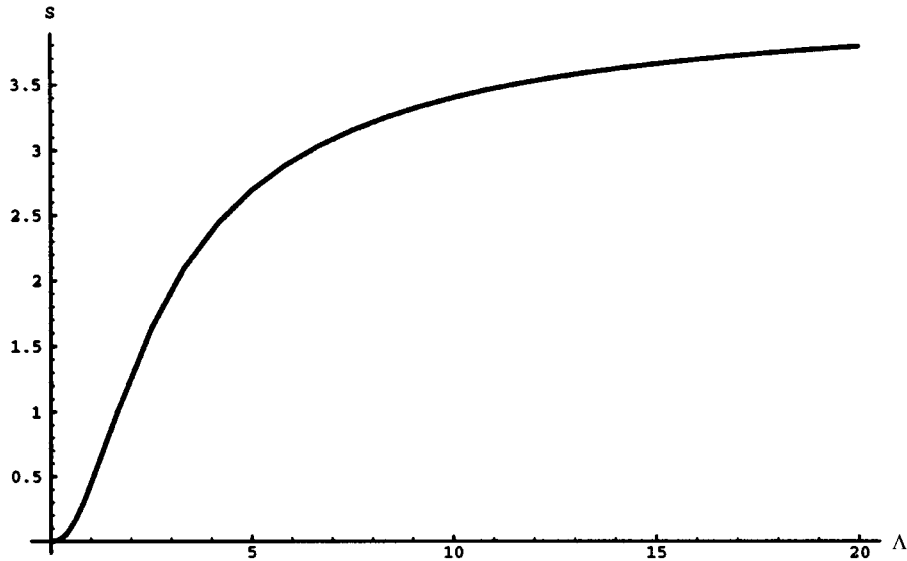


FIG. 4. The Wilsonian relative entropy of the Gaussian model.

ity distribution is such that it does not contain information other than that implied by the constraints or boundary conditions imposed at the outset.

The simplest case of the H theorem is when there is no constraint wherein H is a minimum for a uniform distribution. This is sometimes called the principle of equiprobability. From a philosophical standpoint, it is based in the more general principle of sufficient reason, introduced by Leibnitz. In our context, it can be quoted as stating that if to our knowledge no difference can be ascribed to two possible outcomes of an aleatory process, they must be regarded as equally probable. This is the case for an isolated system in statistical mechanics: all the states of a given energy have the same probability (microcanonical distribution). Another illustrative example is provided by a system thermally coupled to a heat reservoir at a given temperature where we want to impose that the average energy takes a particular value. Minimizing H then yields the canonical distribution.

In general, we may impose constraints on a system with states X_i that the average values of a set of functions of its state, $f_r(X_i)$, adopt predetermined values:

$$\langle f_r \rangle := \sum_i P_i f_r(X_i) = \bar{f}_r,$$

with $P_i = P(X_i)$. The maximum entropy formalism leads to the probability distribution [22]

$$P_i = Z^{-1} \exp\left(-\sum_r \lambda_r f_r(X_i)\right).$$

The λ_r are Lagrange multipliers determined in terms of \bar{f}_r through the constraints. In field theory a state is defined as a field configuration $\phi(x)$. One can define functionals of the field $\mathcal{F}_r[\phi(x)]$. These functionals are usually quasilocal and are called composite fields. The physical input of a theory can be given in two ways, either by specifying the microscopic couplings or by specifying the expectation values of

some composite fields, $\langle \mathcal{F}_r[\phi(x)] \rangle$. The maximum entropy condition provides an expression for the probability distribution,

$$P[\phi(x)] = Z^{-1} \exp\left(-\sum_r \lambda_r \mathcal{F}_r[\phi(x)]\right),$$

and therefore for the action,

$$I = \sum_r \lambda_r \mathcal{F}_r,$$

namely, a linear combination of relevant fields with coupling constants to be determined from the specified $\langle \mathcal{F}_r \rangle$.

The formulation of the H theorem described above is very general. The situation that concerns us here is the crossover from the critical behavior in the vicinity of a multicritical point to another more stable multicritical point under the action of the RG. As soon as a relevant field takes a nonvanishing value, the action of the RG drives the system away from the first fixed point towards the second. In our hierarchical sequence of critical points this was achieved by the couplings being sent to infinity relative to one another in a fashion that descended along this hierarchy. As described above, the condition represented by fixing the expectation value of the relevant field can be understood as imposing a constraint via the introduction of a Lagrange multiplier which appears as a coupling λ_i in the field theory. As in the case of the introduction of β (inverse temperature), when λ_i is sent to infinity we expect the entropy to decrease and thus our relative entropy should increase. Conversely, releasing the constraint is equivalent to sending the coupling to zero and the relative entropy decreases. In the above description the underlying theory is held fixed and only one parameter varied as one moves through a sequence of ‘‘quasistatic’’ states.

In the Wilson RG picture certain expectation values are held fixed while the microscopic theory is allowed to evolve.

This involves the crossover from cutoff-dependent degrees of freedom to cutoff-independent ones and generically falls into the nonequilibrium situation described above. In this process one expects that the entropy will actually increase as the system evolves. This means that our relative entropy should decrease. One can easily see from Fig. 4 in the example described at the beginning of this section that this is indeed the case. In terms of renormalized couplings for given values of the couplings, we can start with any value of λ_i and let the RG act. All the trajectories converge to the critical manifold where λ_i is determined by the other couplings, $\lambda_i(\lambda_r)$. The trajectories approach each other in a sort of reverse chaotic process. In a chaotic process there is great sensitivity to the initial conditions, however, in the RG flow there is great insensitivity to the initial values of the irrelevant couplings which diminish as the flow progresses and in fact vanish at the end of the flow.

V. CONCLUSIONS

We have established that the field theoretic relative entropy provides a natural function which ranks the different critical points in a model. It grows as one descends the hierarchy in the crossovers between scalar field theories corresponding to different multicritical points. This is a consequence of general properties of the entropy and, in particular, of the relative entropy.

We have further established that the phase diagrams of the hierarchy of critical points are associated with a nested sequence of projective spaces. It is convenient to use coordinates adapted to a particular phase diagram in the hierarchy. Hence a crossover implies a coordinate change. The transition from bare to renormalized coordinates provides a method of compactifying the phase diagram. By changing from the bare coordinates, in which the phase diagram naturally ranges over entire hyperplanes to appropriate renormal-

ized ones the phase diagram can be rendered compact.

We discussed the action of the Wilson RG and argued that the relative entropy increases as more degrees of freedom are integrated out, when the underlying Hamiltonian is held fixed. However, when the Hamiltonian is allowed to flow, as it generically is in a Wilson RG, the resulting flow corresponds to a nonequilibrium process in thermodynamics. Nevertheless, the general formulation of the H theorem provided by Jaynes allows us to conclude that the entropy increases in such a process and that the relative entropy (due to our choice of signs) decreases. In contrast, the field theoretic crossover wherein one moves from one point in a phase diagram to another by varying one of the underlying parameters (such as temperature) corresponds to a sequence of quasi-static states and in the case of our hierarchical sequence as one descends the sequence by sending various parameters to infinity one is gradually placing tighter constraints much as reducing the temperature does in the canonical ensemble. Thus one expects the entropy should reduce and the relative entropy increase. This is indeed what we find.

One might wonder as to the connection between our entropy function and the Zamolodchikov C function. It is unlikely that in two dimensions the two are the same. Zamolodchikov's C function is built from correlation data and in the case of a free-field theory it is easy to check that the two functions do not coincide.

ACKNOWLEDGMENTS

The authors express their gratitude to the Institute for Physics, University of Amsterdam, where much of this work was done, for its hospitality and financial support. Denjoe O'Connor is grateful to John Lewis for several helpful conversations. J. Gaité acknowledges support under Grant No. PB 92-1092.

-
- [1] M. E. Fisher, in *Proceedings of the Gibbs Symposium*, New Haven, Connecticut, 1989, edited by D. G. Cali and G. D. Mostow (American Mathematical Society, Providence, 1990).
 - [2] D. J. Wallace and R. K. P. Zia, *Ann. Phys. (N.Y.)* **92**, 142 (1975); J. L. Cardy, *Phys. Lett. B* **215**, 749 (1988); A. Cappelli, D. Friedan, and J. Latorre, *Nucl. Phys.* **B352**, 616 (1991); A. Cappelli, J. I. Latorre, and X. Vilasís-Cardona, *ibid.* **B376**, 510 (1992); V. Periwal, *Mod. Phys. Lett. A* **10**, 1543 (1995).
 - [3] R. Balian, *From Micro-Physics to Macro-Physics*, Monographs in Physics Vol. 1 (Springer-Verlag, Berlin, 1991).
 - [4] J. Pérez-Mercader, in *The Physical Origins of Time Asymmetry*, edited by J. J. Halliwell, J. Pérez-Mercader, and W. Zurek (Cambridge University Press, Cambridge, England, 1994).
 - [5] L. Bombelli, R. Koul, J. Lee, and R. Sorkin, *Phys. Rev. D* **34**, 373 (1986); M. Srednicki, *Phys. Rev. Lett.* **71**, 666 (1993).
 - [6] C. Holzhey, F. Larsen, and F. Wilczek, *Nucl. Phys.* **B424**, 443 (1994).
 - [7] Richard S. Ellis, *Entropy, Large Deviations and Statistical Mechanics* (Springer, Berlin, 1995).
 - [8] J. T. Lewis, C. E. Pfister, and W. G. Sullivan, *J. Stat. Phys.* **77**, 397 (1994).
 - [9] J. T. Lewis and C. E. Pfister, *Russ. Math. Surveys* **50**:2, 279 (1995).
 - [10] Shun-ichi Amari, *Differential-Geometric Methods in Statistics*, Lecture Notes in Statistics Vol. 28 (Springer-Verlag, Berlin, 1985).
 - [11] Denjoe O'Connor and C. R. Stephens, in *Directions in General Relativity*, Proceedings of the 1993 International Symposium, Maryland, edited by B. L. Hu, M. P. Ryan, and C. V. Vishevshawara (Cambridge University Press, Cambridge, England 1993), Vol. 1.
 - [12] F. J. Wegner, *Phys. Rev. B* **5**, 4529 (1972).
 - [13] V. I. Arnold, *Chapitres Supplémentaires de la Théorie des Équations Différentielles Ordinaires* (MIR, Moscow, 1980).
 - [14] S. Lefschetz, *Differential Equations: Geometric Theory*, Pure and Applied Mathematics Vol. V1 (Interscience, New York, 1962).
 - [15] T. H. Berlin and M. Kac, *Phys. Rev.* **86**, 821 (1952).
 - [16] L. Onsager, *Phys. Rev.* **65**, 117 (1944).
 - [17] J. F. Nicoll, T. S. Chang, and H. E. Stanley, *Phys. Rev. B* **12**, 458 (1975).

- [18] Denjoe O'Connor and C. R. Stephens, *Int. J. Mod. Phys. A* **9**, 2805 (1994).
- [19] S. Ma, *Modern Theory of Critical Phenomena*, *Frontiers in Physics* (Springer, Berlin, 1982).
- [20] J. Polchinski, *Nucl. Phys.* **B231**, 269 (1984).
- [21] E. T. Jaynes, *Papers on Probability Theory, Statistics and Statistical Physics* (Reidel, Dordrecht, 1983).
- [22] M. Tribus, *Thermostatistics and Thermodynamics* (Van Nostrand, Princeton, 1961).