Metastability for Markov Processes with Detailed Balance

Hernán Larralde and François Leyvraz

Centro de Ciencias Físicas, UNAM, Apartado postal 48-3, CP 62251 Cuernavaca, Morelos, Mexico

(Received 15 December 2004; published 28 April 2005)

We present a definition for metastable states applicable to arbitrary finite state Markov processes satisfying detailed balance. In particular, we identify a crucial condition that distinguishes metastable states from other slow decaying modes and which allows us to show that our definition has several desirable properties similar to those postulated in the restricted ensemble approach. The intuitive physical meaning of this condition is simply that the total equilibrium probability of finding the system in the metastable state is negligible.

DOI: 10.1103/PhysRevLett.94.160201 PACS numbers: 02.50.Ga, 05.20.Gg, 05.70.Ln

Ever since the work of Gibbs, equilibrium statistical mechanics has rested on a secure theoretical foundation. This is due to the fact that we know the probability distribution describing the statistical properties of these systems. In nonequilibrium phenomena, to which metastability belongs, this is far from being the case in general. For the special case of metastability, however, it has been argued [1] that a description in terms of equilibrium states ought to be possible. In Ref. [1], this was done via analytic continuation techniques. In Ref. [2], an altogether different approach was suggested leading to similar conclusions: there it was argued that, due to the peculiar nature of metastable systems, it should be possible to define a metastable state by considering it as an equilibrium system on a restricted set of configurations. Significant progress in this direction has been realized by Davies [3,4]. In this Letter, following an earlier Letter [5], we take a stochastic process approach to the problem of metastability. We consider systems in which the dynamics of the system is given by a Markov process obeying detailed balance, and that are characterized by having a very slow eigenmode. Under certain conditions, which we specify later, we can give a precise definition of the restricted "metastable region" of the phase space in terms of the slow eigenmode. We then prove that, under these conditions, a system will always relax rapidly to either equilibrium or to a long-lasting pseudoequilibrium state in the metastable region, and, finally, we show that the probability distribution describing this pseudoequilibrium state is proportional to the equilibrium distribution restricted to the metastable subspace.

The phenomenon of metastability may be described informally as follows (see [2] for a much fuller discussion along similar lines): A system is said to be in a metastable state if, upon starting the system in a certain subset of initial conditions, it remains for a very long time in this limited subset, which has negligible measure in equilibrium. The requirement that the time during which the system remains in a metastable state be "large" means simply that it is sufficient to allow the system to relax to some kind of pseudoequilibrium state. Further, this subset is macroscopically distinct from the equilibrium state. Also, the return to equilibrium from a metastable state usually occurs in an abrupt fashion; i.e., the macroscopic variables do not change slowly from their metastable values to their equilibrium values, but rather they remain essentially constant and suddenly relax to their equilibrium value, by some relatively quick relaxation mechanism. This occurs in almost all systems in which an abrupt (first order) transition from one phase to another takes place. Indeed, near the coexistence curve of the two phases, it is usually possible to prepare a state of the unstable phase in such a way that it persists for a very long time: supercooled or superheated liquids are well-known examples. In many cases, such systems, although quite far from equilibrium, can be treated as if they were thermodynamically stable systems. This Letter attempts to present a reason for this surprising fact.

The results presented in this work are derived for arbitrary Markov processes satisfying certain quite general conditions. However, the typical systems we have in mind are models such as the finite Ising or Potts models in $d \geq 2$ dimensions with arbitrary short-range interactions with either Glauber or Kawasaki dynamics; such models can display a broad range of first order phase transitions, and it is well known that they usually show the typical metastable behavior associated with such transitions [6,7]. Furthermore, these systems can be used to model a large variety of physical systems beyond the ferromagnets they were originally meant to describe; thus the Ising model has often been interpreted as a lattice gas model, capable of giving a qualitative description of the liquid gas transition. Similarly, structural transitions in alloys can be modeled by appropriate spin models. In all these physical systems, metastability is in fact routinely observed. Deriving Markovian descriptions from a microscopic picture is not an easy task, but we shall not address this question here; we shall merely assume such a description to be adequate and attempt to define metastability within this framework.

The Glauber or Kawasaki dynamics mentioned above define Markov chains on the set of all spin configurations.

They are ergodic and acyclic, and they satisfy detailed balance with respect to the Gibbs measure [8], which are the generic features we will impose on our system.

We begin by setting up a general framework applicable to a large variety of Markov processes: Given a *finite* set Γ of elements σ, we consider a Markov chain having transition probabilities $W_{\sigma\to\sigma'}$. We are interested in describing the behavior of a system that starts off at a given initial position and then jumps around Γ at rates given by $W_{\sigma\to\sigma'}$. Thus, if we define $P(\sigma, t)$ to be the probability to encounter the system at time t in the configuration σ, this probability obeys the master equation

$$\frac{\partial P}{\partial t}(\sigma, t) = \sum_{\sigma'}[W_{\sigma'\to\sigma}P(\sigma', t) - W_{\sigma\to\sigma'}P(\sigma, t)], \\ \equiv LP(\sigma, t), \tag{1}$$

where L is a linear operator on the space of all vectors $P(\sigma, t)$. Under very general conditions (namely, ergodicity and aperiodicity; see [8]), which are usually satisfied in the systems of interest [9], the probability distribution $P(\sigma, t)$ approaches a unique equilibrium distribution $P_0(\sigma)$ as $t \to \infty$. We will restrict our attention to the cases in which the rates obey detailed balance with respect to the equilibrium distribution [8]:

$$W_{\sigma\to\sigma'}P_0(\sigma) = W_{\sigma'\to\sigma}P_0(\sigma'). \tag{2}$$

Under these conditions, a scalar product of two vectors $\phi(\sigma)$ and $\psi(\sigma)$ can be defined as

$$(\phi, \psi) \equiv \sum_{\sigma}\frac{\phi(\sigma)\psi(\sigma)}{P_0(\sigma)}, \tag{3}$$

under which the operator L is self-adjoint [8]. Since the underlying vector space is finite dimensional, it then follows that there is a complete orthonormal set of eigenvectors P_n with eigenvalues $-\Omega_n$, where the Ω_n are by definition arranged in increasing order. The equilibrium distribution is the eigenvector with $\Omega_0 = 0$, and all other Ω_n are strictly positive.

Using the orthonormality of the P_n, we find

$$\sum_{\sigma}P_n(\sigma) = \delta_{n,0}, \tag{4}$$

implying that $P_0(\sigma)$ is normalized and that adding to it arbitrary multiples of $P_n(\sigma)$, when $n \geq 1$, does not alter this normalization.

From the completeness property of the eigenvectors follows

$$\delta_{\sigma,\sigma_0} = \sum_{n=0}^{\infty}\frac{P_n(\sigma)P_n(\sigma_0)}{P_0(\sigma_0)}. \tag{5}$$

This equation leads to a formal expression for the probability of going from σ_0 to σ in time t:

$$P(\sigma, t; \sigma_0, 0) = P_0(\sigma) + \sum_{n=1}^{\infty}\frac{P_n(\sigma)P_n(\sigma_0)}{P_0(\sigma_0)}e^{-\Omega_n t}. \tag{6}$$

We now turn to the characterization of a metastable state within the general setting outlined above. In view of the informal description of metastability sketched in the introduction, it is clear that, if any behavior different from equilibrium should occur over a large time scale, at least one of the Ω_n must be close to zero. However, it should be emphasized that this is by no means sufficient. For example, any diffusive system of sufficiently large size l will have relaxation rates tending to zero as D/l^2, where D is the diffusion constant. One of the main points of this Letter is to state a condition distinguishing a true metastable state from merely a low-lying eigenvalue of L, which corresponds, say, to a slow relaxation mode within equilibrium.

For clarity, we assume that $\Omega_1 \ll \Omega_n$ for all $n \geq 2$, which also explicitly excludes systems with many metastable states. Systems with a few metastable states present some technical difficulties, such as the possibility that a metastable state might be able to reach equilibrium only by passing through another metastable state, but are presumably similar to the case we treat here. On the other hand, systems with a very large number of metastable states, such as glasses or spin glasses, do not satisfy the simple separation of time scales we are assuming here. An extension of our approach to such systems presents considerable difficulties.

Now consider a process evolving from the initial condition σ_0. Then, following (6), in the relevant time range $\Omega_2^{-1} \ll t \ll \Omega_1^{-1}$, one finds that the configuration σ is occupied with the following (time-independent) probability:

$$P(\sigma) = P_0(\sigma) + \frac{P_1(\sigma_0)}{P_0(\sigma_0)}P_1(\sigma). \tag{7}$$

Note that, due to (4), this is normalized. It is also positive everywhere (except perhaps in some places where it may assume exponentially small negative values).

This focuses our attention on the value $P_1(\sigma_0)/P_0(\sigma_0)$, which characterizes the nature of the initial condition. This quantity will be central to understanding the conditions under which the initial condition can rightly be called metastable and the resulting probability distribution given by (7) can be identified with that of a metastable state. Let us be more specific: In what follows, we denote $P_1(\sigma)/P_0(\sigma)$ by $C(\sigma)$, and by C the maximum value of $C(\sigma)$. Next we define the two sets Γ_m and Γ_{eq} as

$$\Gamma_m := \left\{\sigma: \frac{C}{2} \leq \frac{P_1(\sigma)}{P_0(\sigma)} \leq C\right\}, \tag{8}$$

and Γ_{eq} is defined as the complement of Γ_m. The choice of the factor of $1/2$ to define the lower bound on $C(\sigma)$ in (8) is a matter of convention. We will show that, given the previous scenario, the system will have a metastable state, in the sense discussed in the introduction, if

$$\sum_{\sigma \in \Gamma_m} P_0(\sigma) \ll 1, \tag{9}$$

i.e., that the probability of being found in Γ_m in equilibrium is negligibly small. From a physical point of view, this is the *crucial* assumption, as it distinguishes metastable states from other slow decaying states. We shall show that the consequences that follow from (9) give rise to a behavior that can be identified as metastability and allow us to identify Γ_m with the metastable region. In particular, the main properties of the restricted state approach to the statistical description of metastable states can be derived.

Specifically, we will show that systems in which (9) holds have the following properties: (i) The probability that a state evolving from an initial condition σ_0, for which $C(\sigma_0) = C$, leaves Γ_m in a time less than t is of order $\Omega_1 t$. This justifies identifying such a state as at least a very persistent one. From this result it also follows that

$$\sum_{\sigma \in \Gamma_{eq}} [P_0(\sigma) + CP_1(\sigma)] \ll 1. \tag{10}$$

From this inequality and the positivity properties discussed above, we conclude that

$$P_1(\sigma) \approx -C^{-1} P_0(\sigma), \qquad \sigma \in \Gamma_{eq}. \tag{11}$$

(ii) The probability that a state is found in Γ_{eq} after a time of order Ω_2^{-1}, evolving from an initial condition σ_0 such that $C(\sigma_0) = (1 - p)C$, is p. From this we conclude that, after a time of order Ω_2^{-1} has elapsed, the system will either find itself in a state with $C(\sigma) = C$ (metastable state) or it will be in a state of equilibrium, for which $C(\sigma) = 0$. These results are important because they indicate that the systems that "relax" to a metastable state do so quickly, and once they are in the metastable state, they can be described by the probability distribution

$$P(\sigma) = P_0(\sigma) + CP_1(\sigma). \tag{12}$$

(iii) If we define a new process in which all transition rates connecting the metastable region Γ_m defined by (8) to Γ_{eq} are set equal to zero, we obtain another Markov process, also satisfying detailed balance with respect to the restriction of $P_0(\sigma)$ to Γ_m. We then show that this process is close to the original physical process, in the sense that the difference between the probabilities of reaching the same set X from the same initial conditions σ_0 is of order $\Omega_1 t$, if $C(\sigma_0) = C$. This result leads to

$$P_1(\sigma) \approx CP_0(\sigma), \qquad \sigma \in \Gamma_m, \tag{13}$$

and

$$2 \ln C = \ln \sum_{\sigma \in \Gamma_{eq}} P_0(\sigma) - \ln \sum_{\sigma \in \Gamma_m} P_0(\sigma), \tag{14}$$

which is interpreted in a natural way as the free energy difference between the two phases.

To show how these results come about, note the following basic property:

$$E[e^{\Omega_1 t'} C(\sigma(t')) | \sigma(t)] = e^{\Omega_1 t} C(\sigma(t)) \qquad (t < t'), \tag{15}$$

where $\sigma(t)$ denotes a path of the Markov process defined by (1) and E denotes the conditional expectation value. This relation is easily verified by a straightforward computation and means that $e^{\Omega_1 t} C(\sigma(t))$ is a martingale [8]. Let us now take T to be the random time defined by the first arrival of the path $\sigma(t)$ to Γ_{eq}. T is then a stopping time [8]. We define $\tau = \min(t, T)$; it then follows from (15) using standard theorems [8] that

$$E[e^{\Omega_1 \tau} C(\sigma(\tau))] = C(\sigma_0), \tag{16}$$

where σ_0 is the initial condition. Note that τ depends on the behavior of the path $\sigma(t)$ so that (16) does not follow from (15) by substitution. We first choose σ_0 so that $C(\sigma_0)$ is equal to the maximum possible value C. Noting that if $T < t$ then $C(\sigma(T)) \leq C/2$ and $e^{\Omega_1 T} < e^{\Omega_1 t}$, we can bound the left-hand side of (16) from above:

$$C = E[e^{\Omega_1 \tau} C(\sigma(\tau))] \leq C e^{\Omega_1 t} \left[1 - \frac{1}{2} \mathrm{Prob}(T < t) \right], \tag{17}$$

from which follows

$$\mathrm{Prob}\,(T < t) \leq 2(1 - e^{-\Omega_1 t}) = O(\Omega_1 t). \tag{18}$$

This is a basic result that confirms that, for times such that $\Omega_1 t \ll 1$, the probability that the process leaves Γ_m is negligible. This leads immediately to our first important result, namely, Eq. (10). Indeed, the left-hand side of (10) expresses the probability that the system has reached Γ_{eq} from an initial state having $C(\sigma^{(0)}) = C$ in a short time t (though larger than Ω_2^{-1}) and hence is, by our previous arguments, negligible. However, the result expressed in (18) is limited so far to a specific set of initial conditions, namely, those which have the maximal value of $C(\sigma)$.

We now need to show that if an arbitrary initial condition remains in Γ_m for times larger than Ω_2^{-1}, then it is very probable that it will behave similarly to the initial condition $\sigma^{(0)}$ having $C(\sigma^{(0)}) = C$. This is essential if we wish to argue that the states which form the metastable phase are macroscopically equivalent. For this we first need an intermediate result: Consider an initial condition $\sigma^{(p)}$ such that $C(\sigma^{(p)})$ equals $(1 - p)C$. The probability that this initial condition winds up in Γ_{eq} after a time t has elapsed is close to p if $\Omega_1^{-1} \gg t \gg \Omega_2^{-1}$. Indeed, defining $P(\mathrm{out}, t)$ for this initial condition as $P(\mathrm{out}, t) = \sum_{\sigma \in \Gamma_{eq}} P(\sigma, t; \sigma^{(p)}, 0)$, then, in the relevant time range,

$$P(\mathrm{out}, t) \approx \sum_{\sigma \in \Gamma_{eq}} [P_0(\sigma) + (1 - p)CP_1(\sigma)] \approx p, \tag{19}$$

where we have combined our basic assumption (9) and the result (10). Now let us define

$$F(p) \equiv \sum_{C(\sigma)\le(1-p)C} [P_0(\sigma) + CP_1(\sigma)]. \tag{20}$$

We show that $F(p) \ll 1$; indeed, $F(p)$ represents the probability that a system starting at a σ_0 with $C(\sigma_0) = C$ arrives at a σ satisfying $C(\sigma) \le (1-p)C$ at some fixed time $\Omega_2^{-1} \ll t_0 \ll \Omega_1^{-1}$. If this ever happens, the probability that the system will afterwards reach Γ_{eq} within a time of order Ω_2^{-1} is approximately equal to or greater than p, as follows from (19). But this would imply that a system originally satisfying $C(\sigma_0) = C$ would have nucleated in a time much less than Ω_1^{-1}, which has a negligible probability. Therefore the probability $F(p)$ must be negligibly small.

By the same argument, but now using the fact that the probability for a state which starts in Γ_{eq} to reach Γ_m in a short time is also negligible, we conclude that the support of $P_0(\sigma)$ is concentrated on states for which $C(\sigma) \approx 0$. This time consider

$$G(p) \equiv \sum_{(1-p)C\le C(\sigma)} P_0(\sigma), \tag{21}$$

and, upon using our basic hypothesis (9), we obtain that $G(p) \ll 1$ for $p < 1$. In other words, if (9) holds when Γ_m is defined by the inequalities (8), a similar claim can be shown when the prefactor $1/2$ is replaced by essentially any other number between 0 and 1.

The picture that emerges then is that, after a relatively short transient time (namely, $t > \Omega_2^{-1}$), the system will be found only in states σ for which $C(\sigma) \approx 0$ (equilibrium) or $C(\sigma) \approx C$ (metastability), independently of the initial condition.

Let us now show how a Markov process corresponding to a restricted ensemble can actually be introduced and be shown to remain close to the original Markov process on time scales shorter than Ω_1^{-1}. To this end, define the following restricted transition rates:

$$W^R_{\sigma'\to\sigma} = \begin{cases} W_{\sigma'\to\sigma} & \sigma, \sigma' \in \Gamma_m \text{ or } \sigma, \sigma' \in \Gamma_{eq} \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

Since $P_0(\sigma)$ satisfies detailed balance in the original process, it is still the equilibrium distribution for this restricted process. But the system is no longer irreducible and therefore $P_0(\sigma)$ is not the unique stationary distribution. Indeed, $P_1^R(\sigma)$ defined by

$$P_1^R(\sigma) = \begin{cases} C'P_0(\sigma) & \sigma \in \Gamma_m \\ S'P_0(\sigma) & \sigma \in \Gamma_{eq} \end{cases} \tag{23}$$

is stationary for any constants C' and S'. In particular, we choose these constants so that $\Sigma_\sigma P_1^R(\sigma) = 0$ and $(P_1^R, P_1^R) = 1$ so as to have a correspondence with P_0 and P_1 of the physical system. This implies

$$C' = \left(\frac{\Sigma_{\Gamma_{eq}} P_0(\sigma)}{\Sigma_{\Gamma_m} P_0(\sigma)}\right)^{1/2}, \qquad S' = -1/C'. \tag{24}$$

It should be noted that, given our assumption (9), we will have $C' \gg 1$, as might have been expected intuitively. Of course, it is now tempting to identify P_1^R with P_1. To do this, we need to show that the process defined by (22), which we denoted by R (for restricted), remains close to the original Markov process defined by the rate $W_{\sigma\to\sigma'}$, which we denote by P (for physical). This can be done rigorously by techniques inspired by the coupling techniques of probability theory and will be discussed in greater detail in [10]. Here we restrict ourselves to the following argument: Any path connecting two conditions σ and σ' in time t has equal probability to occur in either process, except if it crosses the boundary between Γ_m and Γ_{eq}. However, such crossings are very unlikely for the range of times $t \ll \Omega_1$ under consideration, so that the two processes are unlikely to differ. From this, it can be shown that the state P_1^R is very close to the state P_1, from which the result $C = C'$ follows.

Summarizing, we have given a formal definition of metastability and displayed a particular condition which distinguishes true metastable states from other slow decaying modes. Using this condition, we show that the features of the restricted state approach to metastability can be derived in a rigorous way. It would be interesting to see whether the restriction of detailed balance could be lifted, so as to be able to define restricted ensembles in other systems far from equilibrium.

The support of DGAPA Project No. IN100803 is gratefully acknowledged.

[1] J. S. Langer, Ann. Phys. (N.Y.) **41**, 108 (1967).
[2] O. Penrose and J. L. Lebowitz, J. Stat. Phys. **3**, 211 (1971).
[3] E. B. Davies and Ph. A. Martin, Helv. Phys. Acta **54**, 106 (1981).
[4] E. B. Davies, Proc. London Math. Soc. **45**, 133 (1982); J. Lond. Math. Soc. **26**, 541 (1982); J. Stat. Phys. **27**, 657 (1982).
[5] G. Báez, H. Larralde, F. Leyvraz, and R. A. Méndez-Sánchez, Phys. Rev. Lett. **90**, 135701 (2003).
[6] K. Binder, Ann. Phys. (N.Y.) **98**, 390 (1976).
[7] M. Kalos, J. L. Lebowitz, O. Penrose, and A. Sur, J. Stat. Phys. **18**, 39 (1978).
[8] P. Brémaud, *Markov Chains, Gibbs Fields, Monte Carlo Simulations and Queues*, Texts in Applied Mathematics Vol. 31 (Springer, New York, 1999).
[9] Note that we are dealing with finite systems only, so that problems of ergodicity due to, say, phase transitions cannot arise.
[10] H. Larralde and F. Leyvraz (to be published).