International Journal of Theoretical Physics, Vol. 25, No. 1, 1986

Asymptotic Thermodynamic Criteria for the Persistency of Metastable States

B. H. Lavenda[1] and C. Cardella[2]

Received May 20, 1985

The mean first exit time is logarithmically equivalent to the exponential of the smallest total free energy barrier separating metastable from stable states for a general class of discontinuous Markov processes in the thermodynamic limit. This asymptotic principle of minimum relative free energy difference constitutes a generalization of the corresponding entropy principle to systems in thermal contact with their environment and formalizes the results of a number of previous authors. The overwhelmingly most probable path of exit, in the thermodynamic limit, is the mirror image in time of the causal path. Along the anticausal path the free energy is an increasing function of time and hence does not provide any criterion of evolution. The kinetic mean-field model of a ferromagnetic serves for illustration.

1. INTRODUCTION

Cramér's (1938) pioneering investigation into large deviations of sums of independent random variables from their mean has recently been generalized to large deviations for families of random processes by Wentzell (1976) and Freidlin and Wentzell (1984). Asymptotics of events that have extremely small probabilities will nevertheless dominate over events which are much more probable in the long run. From a physical point of view this has been recognized for a long time beginning with the description of long-lived metastable states arising in second-order phase transitions as developed in the essentially equivalent theories of Curie and Weiss of spontaneous magnetization and Bragg and Williams of the order-disorder transition in binary alloys. In fact, Kramers' (1940) problem of the escape of a Brownian particle over a potential barrier is essentially of this nature. And it is more

[1]Dipartimento di Scienze Chimiche, Università di Camerino, via Sant' Agostino, 1, Camerino 62032 (MC) Italy.
[2]Dipartimento di Meccanica e Aeronautica, Università di Roma "La Sapienza," via Eudossiana, 18, 00184 Roma Italy.

0020-7748/86/0100-0095$05.00/0 © 1986 Plenum Publishing Company

than a mere coincidence that the same exponential factor involving the height of the free energy barrier separating the metastable from the stable states should be found in Kramers' expression for the asymptotic probability flow, Becker and Döring's (1935) calculation for the relaxation rate of a metastable state in the droplet model (Frenkel, 1946) of condensation, and in the relaxation time for metastable states in the mean-field model of a ferromagnet (Griffiths et al., 1966). In this paper, we shall establish that rough asymptotic estimates (i.e., up to a logarithmic equivalence) predict that the mean exit time from a metastable state in the thermodynamic ($N \to \infty$) limit is logarithmically equivalent to the total free energy barrier for a class of locally divisible, discontinuous Markov processes which are generated by jumplike perturbations. Moreover, we shall show that the path of maximum likelihood for the exit from a metastable state, in the thermodynamic limit on finite time intervals, is the mirror image in time of the causal path, or "anticausal" path, for processes satisfying detailed balance and are characterized by the principle of least dissipation of energy.

The problem is related to the limit behavior of the stationary distribution. It is well known that if there is only one stationary state, the invariant measure converges weakly to a measure concentrated on this state (Wentzell, 1976). If two or more stable stationary states exist, then the asymptotics depend upon the nature of the fluctuations. For discontinuous, jumplike processes we will again obtain the probability of unlikely events in the form of sums of factors of the form $\exp[-N\mathfrak{S}(\phi)]$, where $\mathfrak{S}(\phi)$ is an action functional (Wentzell, 1976) of a smooth path ϕ. As $N \to \infty$, it is only the smallest $\mathfrak{S}(\phi)$ which becomes important and the path will be seen to be related to the mirror image in time of the deterministic path (Lavenda and Santamato, 1982; Lavenda, 1985a). However, the form of the action functional will differ from the Onsager-Machlup functional (Onsager and Machlup, 1953) for diffusion processes.

The general results obtained by Cramér for large deviations for sums of independent variables are based on the assumption that there exist finite exponential moments of the form $E\{e^{\alpha x}\}$, where x is a random variable and for sufficiently small α. The generalization to stochastic processes consists in replacing the random variable x by a stochastic process x_t which, provided it has fixed rates of transitions, will turnout to be Poisson distributed. This we show by using the cluster method (Kikuchi, 1960) for deriving the probability distribution where α will turn out to be the Lagrange multiplier for the constraint that the transitions must conserve probability. As usual the Lagrange multiplier has a physical interpretation and we shall see that it is related to the intensity of the random fluctuations and, not unexpectedly, the causal path is obtained when it vanishes. This is one path of maximum likelihood—albeit the absolutely most probable path—but there may be other maximum likelihood paths for other "critical" values of the Lagrange multiplier (Lavenda and Cardella, 1985).

Cramér introduced exponential factors of the type $e^{\alpha x}$ into the theory of associated random walks with defective distributions (Feller, 1971). These factors were used to transform the defective distribution $\mathfrak{F}$ into a proper distribution $\mathfrak{F}^{\#} = e^{\alpha x}\mathfrak{F}$ of the "associated" random walk, meaning that $\int_{-\infty}^{\infty} e^{\alpha x}\mathfrak{F}\{dx\} = 1$. Whether such an associated distribution exists depends on whether this integral equation has real roots. In the case where the random events are replaced by stochastic processes, the critical root of this equation was shown to be related to the anticausal path (Lavenda and Cardella, 1985). If such a path exists, which implies the property of detailed balance, then the critical Lagrange multiplier will turn out to be the gradient of the free energy. In this paper we show that it is the principle of minimum relative difference in free energy which is the criterion of stochastic exit for metastable states thereby establishing the generality of previous results (Kramers, 1940; Becker and Döring, 1935; Griffiths et al., 1966). And since we are dealing with rough limit theorems, where equivalence means logarithmic equivalence, we have no hope of obtaining the preexponential factors in the expressions for the mean exit and relaxation times. Yet, we share Freidlin and Wentzell's (1984) opinion that there are more interesting consequences which can be derived from rough theorems on large deviations than "sharp" theorems, which are valid up to an equivalence or better. Here we have an explicit case in which a rough limit theorem, for large deviations in a fairly general class of discontinuous Markov processes, is used to establish a criterion of stochastic exit. Namely, the principle of minimum relative free energy difference is a criterion of stochastic exit in systems which can exchange thermal energy with their environment and generalizes our previously derived principle of minimum relative entropy difference, valid for isolated processes (Lavenda and Santamato, 1982; Lavenda, 1985a). Furthermore, we are able to show that the class of maximum probability paths, consisting of the causal and anticausal paths, are characterized by the Rayleigh–Onsager principle of least dissipation of energy (Lavenda, 1978). Along all other paths, in the thermodynamic limit, the dissipation is greater than the rate of change of the free energy.

In Section 2 we derive an expression for the maximum probability distribution of a general, nonlinear one-step process in the thermodynamic limit and offer a physical interpretation of the so-called "step-operator" (van Kampen, 1981) in the master equation by relating it to the Lagrange multiplier appearing in the cluster method (Kikuchi, 1960). Probabilitistic estimates for large deviations and criteria for determining paths of maximum likelihood are given in Section 3. The problem of stochastic exit from the domain of attraction of the metastable state in the kinetic Weiss–Ising model of spontaneous magnetization in the phase transition from paramagnetism to ferromagnetism (Kac, 1968) serves for illustration in Section 4. Finally, in Section 5 we consider the weak noise limit, thereby obtaining the diffusion approximation. Paths of maximum likelihood are also shown to correspond to the minimum rate at which energy is being dissipated.

2. THE CLUSTER METHOD SOLUTION AND INTERPRETATION OF THE MASTER EQUATION FOR ONE-STEP PROCESSES

In this section, we derive a general expression for the maximum probability of a class of locally infinitely divisible, discontinuous Markov processes in the thermodynamic limit. This will provide a physical interpretation of the master equation for a "one-step" process, usually written as (van Kampen, 1981)

$$d/dt\, \mathfrak{P}(n,t) = [(\mathfrak{E}^{-1}-1)g(n) + (\mathfrak{E}-1)r(n)]\mathfrak{P}(n,t) \tag{1}$$

The probability per unit time that, being in state n, a jump occurs to $n-1$ is $r(n)$, while $g(n)$ is the transition probability per unit time for a jump to $n+1$. The "step operator" $\mathfrak{E}$ acts on continuous functions $\mathfrak{f}(n)$ in the following way (van Kampen, 1981):

$$\mathfrak{E}\mathfrak{f}(n) = \mathfrak{f}(n+1) \qquad \text{and} \qquad \mathfrak{E}^{-1}\mathfrak{f}(n) = \mathfrak{f}(n-1) \tag{2}$$

A solution to the Master equation (1) can be derived from the cluster method (Kikuchi, 1960) by considering a system of N states where N is a large number. Since the probabilities for upward and downward transitions are independent of one another, we can consider their distributions individually. Concerning downward transitions, the system, in a small time interval Δt, may either remain in the state n with probability $[1-\theta_r(n)\,\Delta t]$ or it may make a transition to $n-1$ with probability $\theta_r(n)\,\Delta t$, where $\theta_r(n)$ is the microscopic transition probability which can depend on the state of the system. The "path" probability that out of n $(=Ny)$ events, these will happen NP_1 and NP_2 times, respectively, is given by the multinomial expression:

$$\mathfrak{P}_r\{P_i(\Delta t)\} = [(Ny)!/(NP_1)!(NP_2)!][1-\theta_r(n)\,\Delta t]^{NP_1}[\theta_r(n)\,\Delta t]^{NP_2} \tag{3}$$

where $P_i(\Delta t)$ are the *path* parameters for the individual events that occur in the time interval Δt. In a similar manner, the path probability distribution for upward transitions is given by

$$\mathfrak{P}_g\{P_i(\Delta t)\} = [N(1-y)!/(NP_3)!(NP_4)!][\theta_g(n)\,\Delta t]^{NP_3}[1-\theta_g(n)\,\Delta t]^{NP_4} \tag{4}$$

which gives the probability that two independent events with *a priori* probabilities $\theta_g(n)\,\Delta t$ and $[1-\theta_g(n)\,\Delta t]$ will happen exactly NP_3 and NP_4

times, respectively. Since upward and downward transitions are independent, the total path probability $\mathfrak{P}\{P_i(\Delta t)\}$ will be the product of (3) and (4) whose logarithm is

$$(1/N)\ln \mathfrak{P}\{P_i(\Delta t)\} = y\ln y + (1-y)\ln(1-y) - \Sigma_i P_i \ln P_i + P_1 \ln[1-\theta_r(y)\,\Delta t] + P_2 \ln \theta_r(y)\,\Delta t + P_3 \ln \theta_g(y)\,\Delta t + P_4 \ln[1-\theta_g(y)\,\Delta t] \quad (5)$$

provided N is large enough so as to justify the use of Stirling's formula.

In the thermodynamic limit as $N \to \infty$, the path of maximum likelihood is obtained by maximizing (5) with respect to the path parameters $\{P_i(\Delta t)\}$ subject to the constraint that the probability be conserved. At time t, we have

$$y(t) = P_1 + P_2 \quad \text{and} \quad 1 - y(t) = P_3 + P_4 \quad (6)$$

while at time $t + \Delta t$,

$$y(t+\Delta t) = P_1 + P_3 \quad \text{and} \quad 1 - y(t+\Delta t) = P_4 + P_2 \quad (7)$$

since $\sum_i P_i = 1$. Subtracting (6) from (7) gives the finite difference equation:

$$y(t+\Delta t) - y(t) - P_3 + P_2 = 0 \quad (8)$$

which is the constraint underwhich (5) is to be maximized. Note that the independent path parameters are P_2 and P_3, while P_1 and P_4 are dependent variables through (6) and (7). Multiplying (8) by the Lagrange multiplier α and adding it to (5), we vary the path parameters for given fixed initial, $y(t)$, and final, $y(t+\Delta t)$, states. We then obtain

$$P_2^* = P_1\, e^{-\alpha}\theta_r(y)\,\Delta t + o(\Delta t) \quad (9)$$

and

$$P_3^* = P_1\, e^{+\alpha}\theta_g(y)\,\Delta t + o(\Delta t) \quad (10)$$

Introducing the path parameter expressions (9) and (10) into (5) with constraint (8), we get

$$\ln \mathfrak{P}^*\{y(t), y(t+\Delta t)\} = \{r(y)(e^{-\alpha}-1) + g(y)(e^{+\alpha}-1)\}\,\Delta t - \alpha[y(t+\Delta t) - y(t)] + o(\Delta t) \quad (11)$$

subject to given final and initial states. Comparing the time derivative of (11) with the Master equation (1), we identify the macroscopic transition probabilities as

$$r(y) = Ny\theta_r(y) \quad (12)$$

and

$$g(y) = N(1-y)\,\theta_g(y) \quad (13)$$

with the step operator replaced by $e^{-\alpha}$. The probability $\mathfrak{P}^*\{y(t), y(t+\Delta t)\}$ is the maximum of the path probability, appearing in (5), subject to fixed endpoints of transition.

In fact, the master equation (1) can be interpreted as the equation of motion for the joint probability moment generating function of a bivariate Poisson distribution. Let us define

$$\mathfrak{G}(\alpha) = E\{e^{\alpha N[y(t+\Delta t)-y(t)]}\} \tag{14}$$

where the mathematical expectation stands for

$$\mathfrak{G}(\alpha) = \sum_{j,k} e^{\alpha(j-k)} p[Ny(t+\Delta t) = j;\, g(y)\,\Delta t]p[Ny(t) = k;\, r(y)\,\Delta t] \tag{15}$$

Since the p's are Poisson distributions with parameters $g(y)\,\Delta t$ and $r(y)\,\Delta t$ we obtain

$$\mathfrak{G}(\alpha) = e^{\mathscr{H}(\alpha,y)\,\Delta t} \tag{16}$$

The second characteristic function or Hamiltonian is

$$\mathscr{H}(\alpha, y) = g(y)(e^{+\alpha} - 1) + r(y)(e^{-\alpha} - 1) \tag{17}$$

Expanding $\mathscr{H}$ in a power series in α gives

$$\mathscr{H}(\alpha, y) = \sum_{k} (h_k(y)/k!)\,\alpha^k \tag{18}$$

where the coefficient h_k depends on the moments of the distribution and is known as the semi-invariant of order k (Feller, 1971). In the next section, we show that for certain values of the Lagrange multiplier—which is not a mere parameter in the theory as in the definition of the moment generating function—the distribution (11) will transform into well-known forms related to maximum likelihoods of transition.

3. LARGE DEVIATIONS AND CRITERIA FOR PATHS OF MAXIMUM LIKELIHOOD

The Hamiltonian defined by (17) is a convex function and positive semidefinite. To this function, the Young-Fenchel transformation assigns the function $\mathscr{L}(\beta, y)$ defined by

$$\mathscr{L}(\beta, y) = \sup_{\alpha} [\alpha\beta - \mathscr{H}(\alpha, y)] \tag{19}$$

where the Lagrangian $\mathscr{L}$ belongs to the same class as the Hamiltonian. The conjugate variable β is defined as

$$\beta = \partial\mathscr{H}/\partial\alpha \tag{20}$$

and provided the solution $\alpha = \alpha(\beta)$ exists, the Lagrangian is determined by the Young–Fenchel transform (19). Furthermore, the equality $\mathscr{H}(0, y) = 0$ implies that $\mathscr{L}(\beta, y)$ is nonnegative which vanishes when β coincides with the deterministic drift.

Suppose that the expectation

$$\int_0^T \mathscr{H}(\alpha)\, dt = (1/N) \ln E_x\left\{\exp\left[N \int_0^T \alpha\, dx\right]\right\} \tag{21}$$

is finite on $[0, T]$ for all α in some interval $|\alpha| \leq \tilde{\alpha}$. This says that the characteristic function, obtained by replacing real α by $i\alpha$, is analytic in the neighborhood of the origin. Based on an analogy with the method of associated distributions (Feller, 1971), we associate a new probability measure $P^{\#}(dx)$ with the original one $P(dx)$ according to

$$P^{\#}(dx) = \exp\left\{N \int_0^T [\alpha\, dx - \mathscr{H}(\alpha)\, dt]\right\} P(dx) \tag{22}$$

The idea of using such a change of measures is due to Cramer (1938). With respect to the original probability measure $P(dx)$, the probability of any event $\mathfrak{A}$ is

$$P(\mathfrak{A}) = \int_{\mathfrak{A}} \exp\left\{-N \int_0^T [\alpha\, dx - \mathscr{H}(\alpha)\, dt]\right\} P^{\#}(dx) \tag{23}$$

Suppose that we are interested in a particular event for which x_t does not differ by more than $\delta > 0$ from some smooth function ϕ_t on the entire interval $[0, T]$, namely,

$$\mathfrak{A} = \left\{\sup_{0 \leq t \leq T} |x_t - \phi_t| < \delta\right\} \tag{24}$$

Then writing (23) as

$$P(\mathfrak{A}) = \int_{\mathfrak{A}} \exp\left[-N \int_0^T \alpha\, d(x - \phi)\right] P^{\#}(dx) \exp[-N \mathfrak{S}_T(\phi)] \tag{25}$$

where $\mathfrak{S}_T(\phi)$ is the action

$$\mathfrak{S}_T(\phi) := \int_0^T \mathscr{L}[(\phi_t, \dot{\phi}_t)]\, dt \tag{26}$$

Freidlin and Wentzell (1984) estimate the expectation in (25) using the Kolmogorov and Chebyshev inequalities. They establish the lower bound,

$$P\left(\sup_{0 \leq t \leq T} |x_t - \phi_t| < \delta\right) \geq \exp\{-N[\mathfrak{S}_T(\phi) + \gamma]\} \tag{27}$$

for any $\gamma > 0$.

According to expressions (11), (17), and (19) the maximum probability will be determined when the action $\mathfrak{S}_T(\phi)$ achieves its infimum. Since in general

$$\mathscr{L}(\beta, y) \geq \alpha\beta - \mathscr{H}(\alpha, y) \tag{28}$$

and both $\mathscr{L}$ and $\mathscr{H}$ are nonnegative, then the infimum of the action (26) must be determined by the condition (Lavenda and Santamato, 1982; Freidlin and Wentzell, 1984):

$$\mathscr{H}(\tilde{\alpha}) = 0 \tag{29}$$

which defines a class of critical values $\tilde{\alpha}$ of the Lagrange multiplier.

Clearly, $\tilde{\alpha} = 0$ satisfies condition (29); this corresponds to the causal path. The Lagrange multiplier can therefore be related to the intensity of the random fluctuations. In addition to the extremum value $\mathfrak{S}_T(\phi^*) = 0$, corresponding to the causal path ϕ_t^*, there will be an additional critical value of the Lagrange multipler for systems manifesting detailed balance. In multidimensional random processes, detailed balance results when the circulatory probability current vanishes (Ito, 1984, and Section 5).

Let us consider the time interval $[T_1, T_2]$ and allow $T_1 \to -\infty$ thereby giving the system a sufficient amount of time to evolve to within any arbitrarily small neighborhood of the equilibrium state. The equilibrium state O is enclosed in an arbitrary domain Ω with a smooth boundary $\partial\Omega$. There it will remain for an overwhelming portion of its time. Yet, on account of the random fluctuations, there will always be a probability for the system to be found at a finite distance from the equilibrium state and we want to determine the maximum probability for such a spontaneous fluctuation to occur.

In the thermodynamic limit, the probability of this unlikely event occurring is $\exp[-N \min \mathfrak{V}(O, y)]$, where

$$\mathfrak{V}(O, y) = \inf\{\mathfrak{S}_{T_1,T_2}(\phi): \phi_T \in \Delta, \phi_T = y; -\infty \leq T_1 < T_2 < \infty\} \tag{30}$$

subject to the condition that the system was in some arbitrarily small neighborhood Δ of the equilibrium state at some distant time in the past. The critical value of the Lagrange multiplier is determined by condition (29) and this must also satisfy Hamilton's equation of motion [replacing β by $\dot{\phi}_t$, cf. equation (20)]

$$\dot{\phi}_t^\dagger = (\partial\mathscr{H}/\partial\alpha)(\alpha^\dagger, \phi_t^\dagger) \tag{31}$$

with the *final* condition $\phi_{T_2^\dagger} = y$. Since detailed balance holds, the maximum likelihood path $\phi_t^\dagger$ will be the mirror image in time of the causal path ϕ_t^* (Kikuchi, 1961; Lavenda and Cardella, 1985). The infimum of the action (30) is achieved along this extremal. For the critical value of the Lagrange multiplier $\alpha^\dagger$, the Hamiltonian vanishes and the Lagrangian reduces to

$$\mathscr{L}(\phi^\dagger, \dot{\phi}^\dagger) = \alpha^\dagger \dot{\phi}^\dagger \geq 0 \tag{32}$$

Evaluating the time integral of the Lagrangian along the extremal path $\phi_t^\dagger$ we obtain a difference in a function of state which implies that the critical value of the Lagrange multiplier be the derivative of a scalar function, viz.,

$$\alpha^\dagger = \partial \mathfrak{W}(\phi^\dagger)/\partial\phi \tag{33}$$

The action along the anticausal path is the difference in a function of state:

$$\mathfrak{S}_{T_1,T_2}(\phi^\dagger) = \mathfrak{W}(\phi_{T_2^\dagger}) - \mathfrak{W}(\phi_{T_1^\dagger}) \tag{34}$$

The expected first exit time $E_x\{\tau_\Omega\}$ from a domain Ω with a smooth boundary $\partial\Omega$, enclosing a metastable state O, with any initial condition $x \in \Omega$, can be used as an asymptotic measure of the persistency of the metastable state. This is especially true if the interval $[T_1, T_2]$ is not specified in advance. In general, this calculation is seldom feasible. However, in the thermodynamic limit only the principal term of this expression is necessary. Using the strong Markov property, Freidlin and Wentzell (1984) show that $E_x\{\tau_\Omega\}$ is logarithmically equivalent to

$$\exp\left[N\left(\min_y [\mathfrak{W}(y)] - \mathfrak{W}(O)\right)\right\} \tag{35}$$

in the thermodynamic limit. In other words,

$$\lim_{N\to\infty} (1/N) \ln E_x\{\tau_\Omega\} = \min_y[\mathfrak{W}(y) - \mathfrak{W}(O)] \tag{36}$$

The expected first exit time can be regarded as the relaxation time from metastable to stable states; the time spent outside of Ω is of order 1 since the system can evolve without the aid of fluctuations. In particular if O belongs to the interval (ϕ_1, ϕ_2) which are relative maxima of $\mathfrak{W}$ then the expected first exit time is logarithmically equivalent to (see also Labkovskii, 1972)

$$\exp[N\{\min[\mathfrak{W}(\phi_1), \mathfrak{W}(\phi_2)] - \mathfrak{W}(O)\}] \tag{37}$$

thereby providing a criterion where the first exit is most likely to occur. Furthermore, the asymptotic probability flow out of the domain of attraction of O is logarithmically equivalent to the inverse of (35).

4. THE MEAN-FIELD MODEL OF A FERROMAGNET

The kinetic Weiss–Ising model (Griffiths et al., 1966; Suzuki and Kubo, 1968; Kubo et al., 1973) will serve to illustrate the general criteria of the previous section. Consider a system of N Ising spins $\sigma_i = \pm 1$ placed in an external magnetic field H with a coupling strength $J > 0$. Suppose that there are

$$n = (1/2)\left(N + \sum_i \sigma_i\right) \tag{38}$$

spins pointing "up." The microscopic transition probabilities per unit time, $\theta_r(n)$ and $\theta_g(n)$, are given by the exponential Boltzmann factors (Griffiths et al., 1966):

$$\theta_r(n) = \exp\{-(1/2T)[U(n-1) - U(n)]\} \tag{39}$$

and

$$\theta_g(n) = \exp\{-(1/2T)[U(n+1) - U(n)]\} \tag{40}$$

where T is the temperature, in energy units, and $U(n)$ is the internal energy of the system:

$$U(n) = -J(2n-N)^2/2N - \mu_0 H(2n-N) \tag{41}$$

Assuming a quasicontinuous approach (Griffiths et al., 1966), we define the average magnetization per unit spin,

$$x := (2n-N)/N \tag{42}$$

The macroscopic transition probabilities (12) and (13) then become

$$r(x) = (1/2)N(1+x)\exp[(1/T)u'(x)] \tag{43}$$

and

$$g(x) = (1/2)N(1-x)\exp[-(1/T)u'(x)] \tag{44}$$

where

$$u(x) = -(J/2)x^2 - \mu_0 Hx$$

is the internal energy per spin and $u'(x) = du/dx$ is the energy required for flipping a spin.

Instead of using the cluster method of Section 2 to derive the expression for the maximum probability, it is simpler to write the moment generating function as

$$\mathfrak{G}(\alpha) = E\{e^{-2\alpha\Sigma_i\sigma_i}\} = E\{e^{-2\alpha Nx}\} = E\{e^{-2\alpha n}\, e^{2\alpha(N-n)}\}$$

or more explicitly as

$$\mathfrak{G}(\alpha) = \sum_{j,k} e^{[-2\alpha(j-k)]} p[n=j;\, r\,\Delta t]p[N-n=k;\, g\,\Delta t] \tag{45}$$

where the p's are Poisson distributions with parameters $r\,\Delta t$ and $g\,\Delta t$. The factor of 2 arises from the proper normalization of the distribution [cf. (58) below]. The closed-form expression for the moment generating function is therefore (16) with the Hamiltonian given by

$$\mathcal{H}(\alpha, x) = (N/2)[(1+x)\, e^{-(jx+h)}(e^{-2\alpha}-1) + (1-x)\, e^{(jx+h)}(e^{2\alpha}-1)] \tag{46}$$

The conjugate variable of the Lagrange multiplier, (20), is

$$\beta = N\{(1-x)\, e^{[jx+h+2\alpha]} - (1+x)\, e^{-[jx+h+2\alpha]}\} \tag{47}$$

where

$$h = \mu_0 NH/T, \qquad j = NJ/T \tag{48}$$

represent, respectively, the external magnetic field and molecular field scaled by the thermal energy T. A Legendre transform on the Hamiltonian can be performed so long as

$$\partial^2 \mathcal{H}/\partial\alpha^2 = 2N\{(1-x)\, e^{[jx+h+2\alpha]} + (1+x)\, e^{-[jx+h+2\alpha]}\} > 0 \tag{49}$$

implying a functional dependence between the conjugate variables.

For the critical value of the Lagrange multiplier, $\alpha^* = 0$, (47) reduces to the causal equation of motion of the mean field:

$$\dot{x}_t^* = 2[\sinh(jx_t^* + h) - x_t^* \cosh(jx_t^* + h)] \tag{50}$$

A small and positive h gives the stationary state configuration shown in Figure 1 for $j > 1$ in the ferromagnetic region. The metastable stationary state x_m is separated from the stable state x_s by an unstable stationary state x_u. Any initial disturbance of the system to the right of x_u will approach x_s, whereas it will approach x_m for any initial disturbance to the left of that state. For a small external field, the relaxation time is

$$\tau_r = O[(j-1)^{-1}] \tag{51}$$

to both metastable and stationary states, while the transition from the metastable to stable stationary state will occur over a longer time scale, as we now show.

The paths of maximum likelihood are given by condition (29). In addition to the absolutely most probable path, which is the solution to the causal mean field equation (50), there is another path of maximum likelihood which is characterized by the critical value:

$$\alpha^\dagger = (1/2) \ln [(1 + x_t^\dagger)/(1 - x_t^\dagger)] - (jx_t^\dagger + h) \tag{52}$$

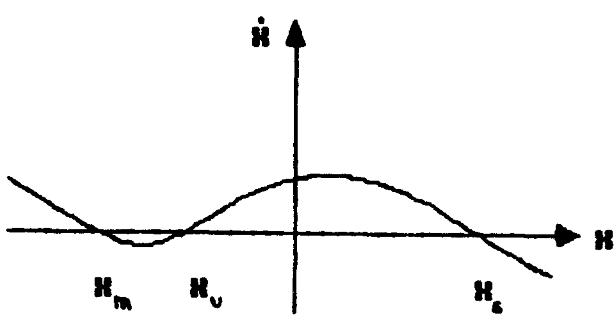

Fig. 1. The stationary state configuration for $J > 1$ and a small positive field h.

of the Lagrange multiplier. The equation of motion of this maximum likelihood path is the mirror image in time of the causal path (50), implying that the system satisfies detailed balance. For both the causal, x_t^*, and anticausal, $x_t^\dagger$, paths, criterion (49) is

$$\partial^2 \mathscr{H}/\partial \tilde{\alpha}^2 = 4N\{\cosh(jx_t + h) - x_t \sinh(jx_t + h)\} > 0. \tag{53}$$

A solution $x_t^\dagger$ exists in both neighborhoods of the metastable and stable states so long as it does not cross x_u. Since the process manifests detailed balance, (52) is the derivative of a scalar potential and hence

$$\begin{aligned}(d/dt)\mathfrak{W}(x_t^\dagger) &= (\partial \mathfrak{W}(x_t^\dagger)/\partial x)\dot{x}_t^\dagger \\ &= \mathscr{L}(x_t^\dagger, \dot{x}_t^\dagger) + \mathscr{H}(\partial \mathfrak{W}(x_t^\dagger)/\partial x, x_t^\dagger)\end{aligned} \tag{54}$$

where according to (29), the second term vanishes. The first term is positive since $\mathscr{L}(x, \beta) \geq 0$ and vanishes only when $\beta = \dot{x}_t^*$. The scalar potential $\mathfrak{W}(x_t^\dagger)$ increases with increasing t.

Consequently,

$$\min[\mathfrak{W}(x_u) - \mathfrak{W}(x_i)] = \int_{x_i}^{x_u} \{(1/2)\ln[(1+x)/(1-x)] - [jx + h]\}\, dx \tag{55}$$

is the criterion which determines from which of the two states exit will be made, namely, the metastable state x_m. After integration, we find that $\mathfrak{W}$ is simply T^{-1} times the Helmholtz free energy per spin, f, viz.,

$$\begin{aligned}\mathfrak{W}(n) &= (n/N)\ln(n/N) + [(N-n)/N]\ln[(N-n)/N] + u(n)/T \\ &= -s(n) + u(n)/T = f(n)/T\end{aligned} \tag{56}$$

where $u(n)$ and $s(n)$ are the internal energy and entropy per spin. *Since $\mathfrak{W}$ decreases for decreasing t, the Helmholtz free energy increases along the anticausal path and consequently does not provide any thermodynamic criterion of evolution along this path.* This form of the motion lies beyond the domain of validity of classical thermodynamics.

The form of the free energy curve is shown in Figure 2 for $j > 1$ and a small, positive external field. The fact that

$$f(x) - f(x_i) = (T/2)\int_{x_i}^{x} \ln[r(x)/g(x)]\, dx > 0 \tag{57}$$

for both $x > x_i$ $(r > g)$ and $x < x_i$ $(r < g)$, where x_i is either one of the two stationary states x_m or x_s means they are both *stable* in the deterministic sense. However, since the two wells of f are not of equal depth, random fluctuations will cause a "flow" in the probability from one side to the other.

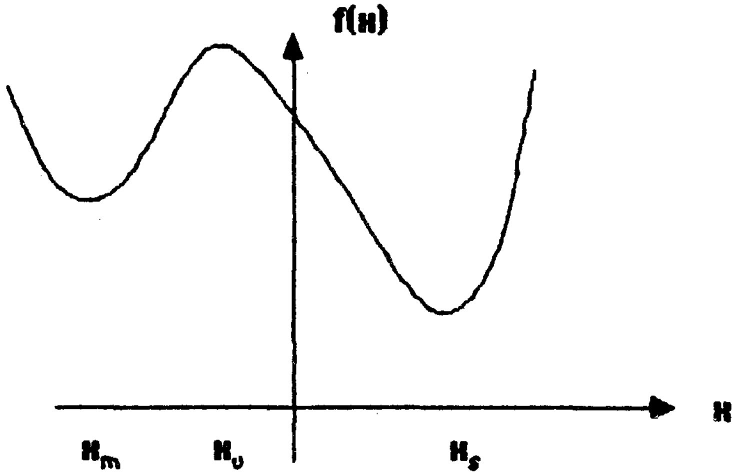

Fig. 2. The free energy curve for $j > 1$ and a small positive field h.

Criterion (55), for the preferential tendency of stochastic exit from the domain of attraction of metastable state, will be referred to as the principle of *minimimum relative free energy difference* for stochastic exit. It generalizes the corresponding entropy principle, valid for isolated systems, to systems in thermal contact with their environment (Lavenda and Santamato, 1982; Lavenda, 1985a). For N sufficiently large, so as to justify the use of Stirling's formula, the probability of finding n spins "up" is

$$\begin{aligned} P(n) &= Z^{-1} \exp[-N\mathfrak{W}(n)] \\ &= Z^{-1}[N!/n!(N-n)!] \exp[-U(n)/T] \end{aligned} \tag{58}$$

where

$$Z = \sum_n \binom{N}{n} e^{-U(n)/T} \tag{59}$$

is the partition function.

The expected first exit time from the domain of attraction of the metastable state is logarithmically equivalent to

$$\exp\{(N/T)[f(x_u) - f(x_m)]\} \tag{60}$$

or the total the free energy barrier (N times the free energy barrier per spin and unit temperature). This asymptotic formula, for large N, is the same as Kramers' (1940) result, for the probability current from one side to the other in the escape of a Brownian particle over a potential barrier, apart from a preexponential factor. This factor cannot be obtained from rough limit theorems on large deviations since they hold only up to a logarithmic equivalence (Wentzell, 1976; Freidlin and Wentzell, 1984).

If another stable stationary state were to lie to the left to x_m, separated by an unstable point x'_u, then the process would leave the domain of attraction of x_m toward the left as $N \to \infty$ provided $f(x'_u) < f(x_u)$. In the event that $f(x'_u) = f(x_u)$, which occurs for a vanishing external magnetic field, the problem of determining where the first expected exit will take place remains open.

5. MAXIMUM LIKELIHOOD PATHS AND LEAST DISSIPATION OF ENERGY

The paths of maximum likelihood may be characterized in terms of the principle of least dissipation of energy (Lavenda, 1978). Along the causal path, the rate of "dissipation" of free energy is

$$d/dt f(x_t^*) = (T/2)[g(x_t^*) - r(x_t^*)] \ln[r(x_t^*)/g(x_t^*)] \leq 0 \tag{61}$$

which cannot increase while along the anticausal path,

$$d/dt f(x_t^\dagger) = (T/2)[r(x_t^\dagger) - g(x_t^\dagger)] \ln[r(x_t^\dagger)/g(x_t^\dagger)] \geq 0 \tag{62}$$

the free energy cannot decrease when detailed balance holds. For small values of the Lagrange multiplier or a weak noise intensity, we may show that both (61) and (62) are the minimum rates of dissipation of energy for the decay and growth of fluctuations, respectively.

For small noise intensities, the Lagrangian is given in the Onsager–Machlup form (Onsager and Machlup, 1953; Lavenda and Santamato, 1982):

$$\mathscr{L}(x_t, \dot{x}_t) = (1/2)\{\dot{x}_t - [g(x_t) - r(x_t)]\}^2/[g(x_t) + r(x_t)] \geq 0 \tag{63}$$

Expanding the quadratic form, we obtain

$$\begin{aligned} \emptyset(\dot{x}_t) + \Psi(x_t) &\geq -\dot{x}_t[r(x_t) - g(x_t)]/[g(x_t) + r(x_t)] \\ &= -(1/T) d/dt f(x_t) \end{aligned} \tag{64}$$

where we have identified the Rayleigh-Onsager dissipation function as (Lavenda, 1978)

$$2\emptyset(\dot{x}_t) = [g(x_t) + r(x_t)]^{-1}\dot{x}_t^2 \tag{65}$$

and

$$2\Psi(x_t) = [g(x_t) + r(x_t)]^{-1}[g(x_t) - r(h_t)]^2 \tag{66}$$

is known as the "generating" function (Landau and Lifshitz, 1959; Lavenda, 1985a). In the derivation of (64), we have used the fact that for small values of the Lagrange multiplier, $(1/2)\ln[r(x_t)/g(x_t)] \approx [r(x_t) - g(x_t)]/[g(x_t) + r(x_t)]$. The criterion for paths of maximum likelihood, (29), is now given by the dissipation balance condition (Lavenda and Santamato, 1982; Lavenda, 1985a):

$$\emptyset(\dot{\tilde{x}}_t) - \Psi(\tilde{x}_t) \tag{67}$$

Along the causal path, the Lagrangian vanishes and (64) reduces to the equality

$$2T\emptyset(\dot{x}_t^*) = -d/dt f(x_t^*) \geq 0 \tag{68}$$

Expressed in words, (68) states that the *minimum* rate of dissipation of energy is given by the negative rate of change of the free energy per unit temperature. Along the causal path the free energy decreases in time and provides a thermodynamic criterion of evolution.

Similarly, the semidefinite quadratic form

$$\mathcal{L}(x_t, -\dot{x}_t) = (1/2)\{\dot{x}_t + [g(x_t) - r(x_t)]\}^2/[g(x_t) + r(x_t)] \geq 0 \tag{69}$$

which is the mirror image in time of the Lagrangian (63), can be written as the thermodynamic inequality:

$$\begin{aligned}\emptyset(\dot{x}_t) + \Psi(x_t) &\geq \dot{x}_t[r(x_t) - g(x_t)]/[g(x_t) + r(x_t)] \\ &= (1/T)\, d/dt f(x_t)\end{aligned} \tag{70}$$

Along the anticausal path, which satisfies the dissipation balance condition (67), we obtain

$$\begin{aligned}2\emptyset(\dot{x}_t^\dagger) &= [r(x_t^\dagger) - g(x_t^\dagger)]^2/[g(x_t^\dagger) + r(x_t^\dagger)] \\ &= (1/T)\, d/dt f(x_t^\dagger) \geq 0\end{aligned} \tag{71}$$

showing again that this path of maximum likelihood is characterized by a minimum rate at which energy is being dissipated.

In the case of m random processes with $m > 1$, the probability current density does not have to vanish inside the region under consideration even though it may satisfy zero boundary conditions (Stratonovich, 1963). The presence of a rotational flow will destroy detailed balancing of the individual processes (Lavenda, 1985b). The probability flow around a closed curve ϕ_t will have the asymptotic form (Ikeda and Watanabe, 1981)

$$\exp[-N\mathfrak{C}(\phi)] \tag{72}$$

for $0 \leq t \leq T$, where T is a period of the motion and $\phi_0 = \phi_T$. The circulation $\mathfrak{C}$ is defined as

$$\mathfrak{C}(\phi) := \int_0^T \{\mathcal{L}(\phi_t, \dot{\phi}_t) - \mathcal{L}(\phi_t, -\dot{\phi}_t)\}\, dt \tag{73}$$

The relation between the integrand of (73) and the critical vector $\alpha^\dagger = (\alpha_1^\dagger, \ldots, \alpha_m^\dagger)$ is

$$\mathcal{L}(\phi, \dot{\phi}) - \mathcal{L}(\phi, -\dot{\phi}) = 2\sum_i \alpha_i^\dagger \dot{\phi}_i \tag{74}$$

Therefore only in the event that the critical Lagrange multiplier is the gradient of a scalar function (e.g., the free energy) will the circulation (73) vanish. In the present case the integrand in (73) can be evaluated with the aid of (63) and (69). We then obtain

$$\mathfrak{C}(\emptyset) = 2 \int_0^T \{[r(\emptyset_t) - g(\emptyset_t)]/[r(\emptyset_t) + g(\emptyset_t)]\} \, d\emptyset_t \tag{75}$$

where we note that the integrand is the diffusion limit approximation for the critical value of the Lagrange multiplier, (52).

ACKNOWLEDGMENTS

This work has been partially supported by contributions from the Consiglio Nazionale delle Ricerche and the Ministero della Pubblica Istruzione.

REFERENCES

Becker, R., and Döring, W. (1935). *Annalen der Physik* **24**, 719.

Cramér, H. (1938). Sur un nouveau théorème de la théorie des probabilitités, *Acta Scientia et Industrie*, 736.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol II, pp. 548ff, Wiley, New York.

Freidlin, M. I., and Wentzell, A. D. (1984). *Random Perturbations of Dynamical Systems*, Chap. 5, Springer, New York.

Frenkel, J. (1946). *Kinetic Theory of Liquids*, Chap. 11, Oxford, London.

Griffiths, R. B., Weng, C-Y., and Langer, J. S. (1966). *Physical Review*, **149**, 301.

Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations*, p. 451, Kodansha, Tokyo.

Ito, H. M. (1984). *Journal of Physics A: Mathematical, Nuclear and General*, **17**, 1075.

Kac, M. (1968). In *Statistical Physics, Phase Transitions and Superfluidity*, Chrétien, M., Gross, E. P., and Deser, S. eds., Vol. 1, p. 243, Gordon & Breach, New York.

Kikuchi, R. (1960). *Annals Physics (New York)*, **10**, 127.

Kikuchi, R. (1961). *Physical Review*, **124**, 1682.

Kramers, H. A. (1940). *Physica*, **7**, 284.

Kubo, R., Matsuo, K., and Kitahara, K. (1973). *Journal Statistical Physics*, **9**, 51.

Labkovskii, V. A. (1972). *Akad. Nauk Gruzin. SSR*, **67**, 41 (in Russian).

Landau, L. D. and Lifshitz, E. M. (1959). *Statistical Physics*, p. 377. Pergamon, London.

Lavenda, B. H. (1978). *Thermodynamics of Irreversible Processes*, Chap. VI. Macmillan, London.

Lavenda, B. H. (1985a). *Nonequilibrium Statistical Thermodynamics*, Chap. 7. Wiley-Interscience, Chichester.

Lavenda, B. H. (1985b). *Journal of Physics A*, **18**, 293.

Lavenda, B. H., and Santamato, E. (1982). *Journal of Statistical Physics*, **29**, 345.

Lavenda, B. H., and Cardella, C. (1985). On the persistence of irreversible processes under the influence of random thermal fluctuations, *Journal of Physics A: Mathematical, Nuclear and General*, **18**.

Onsager, L., and Machlup, S. (1953). *Physical Review*, **91**, 1505.
Stratonovich, R. L. (1963). *Topics in the Theory of Random Noise*, Vol. 1, p. 76. Gordon & Breach, New York.
Suzuki, M., and Kubo, R. (1968). *Journal of the Physical Society of Japan*, **24**, 51.
van Kampen, N. G. (1981). *Stochastic Processes in Physics and Chemistry*, Chap. VI. North-Holland, Amsterdam.
Wentzell, A. D. (1976). *Theory of Probability and Applications*, **21**, 227; **21**, 449.