Escape or switching at short times

G. Grinstein and R. H. Koch
IBM Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598, USA
(Received 23 November 2004; published 18 October 2005)

In the standard Arrhenius picture [S. Arrhenius, Z. Phys. Chem., Stoechiom. Verwandtschaftsl. **4**, 226 (1889); L. Néel, Ann. Geophys. (C.N.R.S.) **5**, 99 (1949)] of thermal switching or escape from a metastable to a stable state, the escape probability per unit time $P_s(t)$ decreases monotonically with time t as $P_s(t) \sim e^{-t/\tau_D}$, where the decay time $\tau_D = \tau_0 e^{U/k_B T}$, with U the energy barrier, $k_B T$ the thermal energy, and τ_0 the time between escape attempts. Here, we extend the Arrhenius picture to shorter times by deriving general conditions under which $P_s(t)$ is peaked rather than monotonic, and showing that in the simplest scenario the peak time τ_P diverges with τ_D as $\ln(\tau_D)$.

DOI: 10.1103/PhysRevE.72.046121 PACS number(s): 02.50.Ey, 05.40.−a

A basic paradigm for understanding the natural world is the escape or "switching" of a classical system from a metastable to a stable state through thermal activation over an energy barrier. In the standard Arrhenius picture [1,2], the escape probability $P_s(t)$ decays with time t as

$$P_s(t) \sim e^{-t/\tau_D}, \tag{1}$$

the decay time τ_D being given by $\tau_D = \tau_0 e^{U/k_B T}$. Here τ_0 measures the time between escape attempts, T is the temperature, U the energy barrier, and k_B Boltzmann's constant. These relations are among the best known in all of science, and are used ubiquitously to predict the decay of metastable systems. While formula (1) holds for long times, $t \gg \tau_D$, its legitimacy for shorter times is far from clear, however. For familiar problems such as the classic model of radioactive decay [3] or the switching of a single Ising spin [2,4], where escape is a Poisson process involving a single transition out of a metastable well, formula (1) holds for all $t \geq 0$. Recent work has suggested that this simple monontonic decay of $P_s(t)$ does not hold for all switching problems, however: Numerical experiments on spatially extended Ising systems [5–7] and laboratory measurements on switching in sub-micron sized magnetic tunnel junctions [8] found $P_s(t)$ curves with peaked shapes [Fig. 1(a)]. To complicate matters further, neither newer experiments on different tunnel junctions [9] nor experiments on switching in tiny single-domain magnetic particles [10] showed any deviation from the monotonic Arrhenius behavior. Thus the validity of Eq. (1) [or other more complex, monotonically decreasing forms for $P_s(t)$ [11]] at all but asymptotically long times remains unclear, the subject of apparently conflicting experimental claims.

In this paper, we attempt to resolve the problem by extending the Arrhenius picture to shorter times through simple derivation of the general conditions under which $P_s(t)$ is either peaked or monotonic, and by showing that, when $P_s(t)$ is peaked, the peak time τ_P commonly diverges with τ_D as $\ln(\tau_D)$. We also argue that switching curves found to be monotonic in some previous numerics on the two-dimensional Ising model and experiments on magnetic tunnel junctions actually have a peaked structure, verifying this claim explicitly in the Ising case through numerical study.

In switching problems [12], the states are divided into two groups: "switched" and "unswitched." The system starts in an unswitched state (or states), and $P_s(t)$ is defined as the

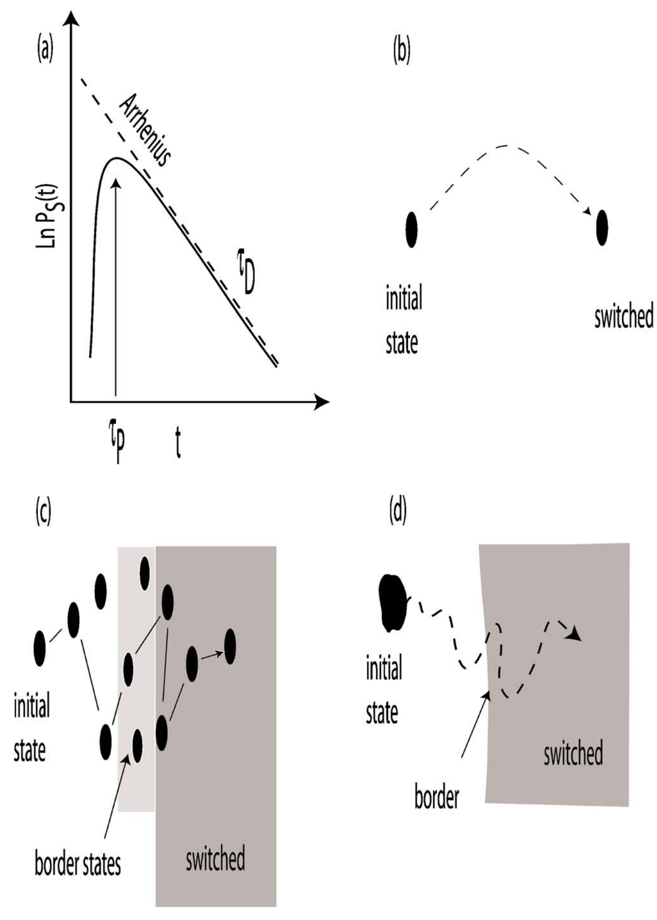

FIG. 1. (Color online) (a) Schematic semilog plot of $P_s(t)$ vs t, contrasting monotonic (e.g., Arrhenius) decay (dashed line) against peaked shape; (b) single, direct switching transition in system with discrete states produces $P_s(t)$ with pure exponential decay; (c) in system with discrete states, $P_s(t)$ has peak if switching must pass through initially unoccupied border states (shaded) to switch; (d) system with continuous states switches through biased diffusion across border, producing peak in $P_s(t)$.

1539-3755/2005/72(4)/046121(4)/$23.00

probability per unit time of the first passage to a switched state occurring at time t. We now argue that $P_s(t)$ can decrease monotonically for all $t \geq 0$ only in special cases wherein the system can make, with nonzero probability, a direct transition from the (unswitched) initial state(s) to the switched final state. Though some prominent examples of escape, such as the switching of a single binary variable (e.g., an Ising spin) between its two states, fall into this category [Fig. 1(b)], they are the exceptions rather than the rule. More typical are systems that must pass through a set of unswitched intermediate states in order to switch, giving rise to switching probabilities with peaked shapes.

Let us first consider systems with discrete state spaces, that is, consisting of a set of N variables S_α for $\alpha = 1, 2, \ldots, N$, each of which can assume a set of discrete values. Let there be a probability $Q(i'|i)$ per unit time of the system making a transition from state $i = \{S_\alpha\}$ to state $i' = \{S'_\alpha\}$. Typically, only pairs of states i and i' which differ from each other by a modest number of S_α's have $Q(i|i')$'s that are nonzero. (In Ising model dynamics, e.g., only single-spin or at most few-spin flips are generally allowed.) It is useful to define the subset, $\{i_B\}$, of "border" states: unswitched states from which direct transitions to switched states occur with nonzero probability. Then $P_s(t) = \Sigma_{i_B} P_{i_B}(t) q_{i_B}$, where $P_{i_B}(t)$ is the probability of the system occupying the border state i_B at time t, and q_{i_B} is the total probability per unit time of the system undergoing a transition into a switched state from the state i_B.

If none of the border states is occupied at $t=0$, then $P_{i_B}(t=0)=0$ for all i_B, so $P_s(t=0)=0$. The occupation probabilities of the border states must (barring pathological transition probabilites), increase continuously with t for sufficiently small t, and so, therefore, does $P_s(t)$. This initial increase of $P_s(t)$, together with its exponential decay for large t, implies that $P_s(t)$ must achieve a maximum at some $t = \tau_P > 0$ [Fig. 1(c)].

If, on the other hand, some of the border states are occupied initially, then direct transitions into switched states can occur at $t=0$. In this case $P_s(t=0)$ will have the nonzero value $P_s(t=0) = \Sigma_{i_B} P_{i_B}(t=0) q_{i_B}$. As t increases, $P_s(t)$ can then either decrease monotonically or increase initially and exhibit a peaked structure, depending upon the details of the initial occupation and transition probabilities. In simple models of radioactive decay, e.g., or the switching of a single binary variable, such as an Ising spin, there are only two states: the (unswitched) border state that is occupied initially, and the switched state. In such cases, switching is a Poisson process, and $P_s(t)$ decreases exponentially for all $t \geq 0$ [Fig. 1(b)].

Now consider problems with continuous- rather than discrete-valued variables, such as magnetic systems described by Brown's stochastic generalization [13] of the Landau-Lifshitz-Gilbert equations of micromagnetics [14,15]. Again one divides the (now continuous) state space, with N dimensions say, into "unswitched" and "switched" regions, separated by a border. Such problems are conveniently formulated in terms of the standard multivariate Fokker-Planck equation [16,17] for the probability density $p(\vec{x},t)$ of the system being at the point $\vec{x}$ in phase space at time t, where $\vec{x}$ is an N-component vector

$$\partial_t p(\vec{x},t) = -\partial_i[A_i(\vec{x})p(\vec{x},t)] + \frac{1}{2}\partial_i\partial_j[B_{ij}(\vec{x})p(\vec{x},t)]. \quad (2)$$

Here the functions A_i and B_{ij} describe the deterministic and noisy parts of the dynamic, respectively, and summation over repeated component indices i and j is implied. Starting from its initial state, the system executes through phase space the biased diffusive motion governed by Eq. (2), until it switches by crossing the border into the "switched" region. This is represented by imposing the "absorbing" boundary condition [17] $p(\vec{x},t)=0$ for every point $\vec{x}$ on the border.

$P_s(t)$ is then given by the integral over the entire border of the component of the probability current density normal to the border. In the Fokker-Planck formulation, this current is a linear function of $p(\vec{x},t)$ and its spatial derivatives. Since some distance in phase space typically separates the initial state(s) from the border, $p(\vec{x},t)$ and its spatial derivatives vanish at the border at $t=0$. Hence the current density also vanishes at the border at $t=0$, whereupon $P_s(t=0)=0$. This current density must increase from zero for small t, reflecting the initial growth from zero of the occupation probability of the points in phase space along all possible paths to the border from the initial point(s), and of the border states themselves. As the normal current at the border increases from zero with time, so does $P_s(t)$, which therefore must have a peaked structure as in the discrete case, barring pathologies in the functions that appear in the Fokker-Planck equation.

To reach the border and switch, in other words, the system must perform, through the state space, a continuous random walk biased adversely by the deterministic dynamics [18], and hence by the energy function if there is one [Fig. 1(d)]. Assuming that the noise is not pathological, this initial increase is presumably continuous and gradual [19].

To derive an asymptotic relation between τ_P and τ_D from simple arguments, we start with the case of discrete variables. The dynamics of the switching problem can be described by the familiar master equation [16] $d\mathcal{P}/dt = \mathcal{Q}\mathcal{P}$ for the unswitched states $\{i_U\}$; here the elements of the vector $\mathcal{P}$ and the matrix $\mathcal{Q}$ are $P_{i_U}(t)$ and $Q(i_U|j_U)$, respectively, with $Q(i_U|i_U)$ the negative of the total probability per unit time of the system undergoing a transition out of state i_U. When $\mathcal{Q}$ can be diagonalized, $P_{i_U}(t)$, and hence $P_s(t)$, can be expressed as the eigenvector expansion [17] $P_s(t) = \Sigma_i A_i e^{\lambda_i t}$, where λ_i is the ith eigenvector, $\lambda_{i+1} \leq \lambda_i < 0$ for all i, and the A_i are coefficients [20]. For large t, this expansion is dominated by the terms corresponding to the leading eigenvalues, all of which must be negative to reflect the vanishing of $P_s(t)$ at $t=\infty$.

Further assuming that the two leading eigenvalues, λ_1 and λ_2 (where $\lambda_2 < \lambda_1 < 0$), provide an adequate description of the peaked $P_s(t)$ for all t, one has

$$P_s(t) \sim A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t}, \quad (3)$$

for coefficients A_1 and A_2. The dominance of the leading term $A_1 e^{\lambda_1 t}$ for asymptotically large t, clearly implies that $\tau_D = -1/\lambda_1$. The constraints that $P_s(t)$ be positive for all $t \geq 0$ and have a peaked structure [i.e., a positive time derivative, $P'_s(t) > 0$, for t near zero], further require that $A_1 > 0$,

$A_1+A_2 \geqslant 0$, and $A_1\lambda_1+A_2\lambda_2>0$; this last inequality implies $A_2<0$. One can then solve Eq. (3) for the peak time τ_P by setting $P_s'(t=\tau_P)=0$. This yields $\tau_P = \ln(-A_2\lambda_2/A_1\lambda_1)/(\lambda_1-\lambda_2)$.

In the limit $\lambda_1=0$, the eigenvalue λ_2 governs the long-time relaxation of probability towards the equilibrium or steady-state distribution, and so typically has a finite, nonzero value in this limit. In the simplest case, where the ratio A_2/A_1 does not vanish as $\lambda_1 \to 0$, the asymptotic result $\tau_P \sim \ln(-1/\lambda_1) = \ln(\tau_D)$ as $\lambda_1 \to 0$ follows immediately. (Note that A_2/A_1 cannot diverge, since that would violate the condition $A_1+A_2 \geqslant 0$.)

Even if A_2/A_1 vanishes as $\lambda_1 \to 0$, it will typically do so as a power of λ_1. In Ising-like systems, e.g., one anticipates A_2/A_1 vanishing exponentially in $1/T$ as $T \to 0$, i.e., as a power of λ_1. Then τ_P continues to diverge as $\ln(\tau_D)$. One can show explicitly that this is what occurs in the single-variable "ladder" model [8], which is described by the matrix $\mathcal{Q}$ with a particularly simple form.

Different outcomes and more complex scenarios may also be possible. If, for example, both A_1 and A_2 in Eq. (3) are positive, then the positivity of $P_s'(t)$ near $t=0$ requires a third term in the eigenvector expansion. Assuming that λ_2, λ_1, and the ratios of the A_i's approach finite constants as $\lambda_1 \to 0$, τ_P is then easily shown to approach a constant as τ_D diverges. While such behavior has not yet been observed in any solved or simulated models, including the Ising model discussed below, its occurrence has not been ruled out.

Turning to the case of continuous variables, the Ansatz $p(\vec{x},t) = P_{\lambda_i}(\vec{x})e^{-\lambda_i t}$ turns the Fokker-Planck equation for the switching problem (with absorbing boundary conditions at the border) into an eigenvalue problem with eigenvalues λ_i and eigenfunctions $P_{\lambda_i}(\vec{x})$ [16,17]. (Here we assume a discrete eigenvalue spectrum and the existence of a complete set of eigenfunctions.) $P_s(t)$, the integral over the border of the component of the probability current density normal to the border, can then be expressed in the same form as in the discrete-variable problem, viz., $P_s(t) = \Sigma_i A_i e^{-\lambda_i t}$, for coefficients A_i. Again, in the simplest scenario where only the two leading terms need be retained, τ_P diverges as $\ln(\tau_D)$.

The relation between τ_P and τ_D just derived is consistent with the results of existing calculations of $P_s(t)$ and τ_P on single-variable models with discrete [8] and continuous [21] variables, but remains to be tested on a many-body system. Accordingly, we performed Monte Carlo simulations of the classical two-dimensional nearest-neighbor ferromagnetic Ising model on a square lattice in a magnetic field. The model is described by the familiar Hamiltonian

$$H = -J\sum_{\langle ij \rangle} S_i S_j - h\sum_i S_i, \qquad (4)$$

with exchange constant $J(>0)$, and magnetic field h; here $S_i = \pm 1$ is a binary variable on the ith site of a two-dimensional square lattice of linear size L, $\Sigma_{\langle ij \rangle}$ designating a sum over nearest-neighbor pairs of sites.

Simulations were performed at temperature T, following the protocol described in Ref. [7]: The system is initialized with all spins pointing down, then updated according to the standard Metropolis Monte Carlo algorithm [22], with single spin flips at randomly chosen sites as the only allowed moves. After each attempted pass through the system (i.e., L^2 spin updates), one computes the total magnetization, M, halting the simulation and recording the time elapsed, i.e., the number of passes completed, when M first exceeds 0. This time is defined as the switching time. Many repetitions of this procedure generate the switching probability $P_s(t)$.

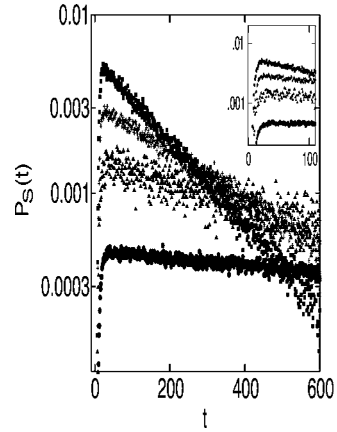

FIG. 2. Semilog plot of $P_s(t)$ vs t for Ising model with parameters $L=20$, $J=1$, $k_BT=1.81535$, and (top to bottom) $h/J=0.125$, 0.115, 0.105, 0.90. The inset shows detail of short-time behavior.

One can also iterate the master equation of the system directly, thereby calculating the probability of the system remaining unswitched, and hence $P_s(t)$, as a function of elapsed time t. This obviates the need for any random numbers or repetitions, allowing one to achieve much longer times and reduced error. Even though one need only keep track of the unswitched states, the exponential growth with L of their number limits this method to very small L.

Figure 2 shows the semilog $P_s(t)$ vs t curves produced by these simulations, for fixed values of the linear system size ($L=20$), exchange strength, and temperature, and a series of different magnetic field values. These curves indeed have the

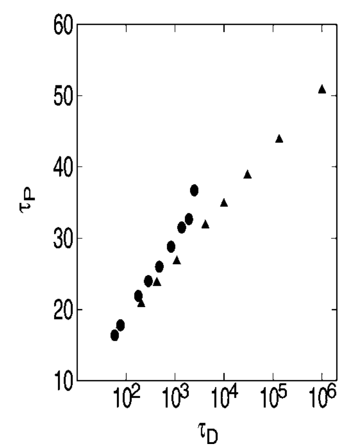

FIG. 3. Semilog plot of τ_P vs τ_D for (a) $L=20$, $J=1.0$, $h=0.15$, and k_BT taking values between 1.81535 and 1.60 (circular symbols, from simulations); (b) $L=3$, $J=1.0$, $h=2.0$, and k_BT between 1.2 and 0.3 (triangular symbols, from master equation).

peaked shape predicted by our arguments, the linear decrease at large t indicating the expected exponential decay.

Figure 3 shows a semilog plot of τ_P vs τ_D covering roughly two orders of magnitude in τ_D, for simulations performed with $L=20$ and a series of temperature values. The data seem consistent with τ_P growing as $\ln(\tau_D)$, as per the prediction above. Figure 3 also shows data acquired for $L=3$ and a series of different temperatures, obtained from direct numerical solution of the master equation for the system, rather than from simulations. These data obey $\tau_P \sim \ln(\tau_D)$ over four decades in τ_D.

In earlier Ising model simulations with parameters very similar to those of our Fig. 2, the probability, $P(t)=1-\int_0^t P_s(u)du$, of the system not having switched at time t was found to exhibit pure exponential decay [7]. This implies a monotonic exponential decrease of $P_s(t)$ and no peak, which seems to contradict our findings. Since, however, an obvious peak in $P_S(t)$ manifests itself as a much less obvious inflection point in $P(t)$, we feel confident that $P_s(t)$ in Ref. [7] actually does have a peak, so there is no real discrepancy between the two sets of results. A similar statement applies to measurements of $P(t)$ in magnetic tunnel junctions [9] similar to those studied in Ref. [8]: We believe that direct measurements of $P_s(t)$ on the sample of Ref. [9] would in fact show the peaks found in Ref. [8].

In summary, the picture of switching that emerges from our arguments is one wherein, apart from well known cases that turn out to be rather special, $P_s(t)$ has a peaked structure with the peak time τ_P commonly growing as $\ln(\tau_D)$ asymptotically, where $\tau_D \sim e^{U/k_BT}$ is the characteristic time for the exponential decay of $P_s(t)$ at large t. Empirically, moreover, τ_P seems to set a rough time scale beyond which $P_s(t)$ decreases exponentially. This scenario generalizes the venerable results of Arrhenius and Néel to shorter times, providing at least a qualitative picture of switching on all time scales. Though typically $\tau_P \ll \tau_D$, recent experiments [8,9] make clear that modern magnetic and other devices have evolved to the point where even very modest peak times have practical relevance.

[1] S. Arrhenius, Z. Phys. Chem., Stoechiom. Verwandtschaftsl. **4**, 226 (1889).

[2] L. Néel, Ann. Geophys. (C.N.R.S.) **5**, 99 (1949).

[3] T. A. Littlefield and N. Thorley, *Atomic and Nuclear Physics*, 3rd ed. (Van Nostrand Reinhold, New York, 1979).

[4] E. C. Stoner and E. P. Wohlfarth, Philos. Trans. R. Soc. London, Ser. A **240**, 599 (1948).

[5] J. M. González *et al.*, Phys. Rev. B **52**, 16034 (1995).

[6] D. García *et al.*, J. Appl. Phys. **79**, 6019 (1996).

[7] H. L. Richards *et al.*, J. Magn. Magn. Mater. **150**, 37 (1995).

[8] R. H. Koch *et al.*, Phys. Rev. Lett. **84**, 5419 (2000).

[9] N. D. Rizzo *et al.*, Appl. Phys. Lett. **80**, 2335 (2002).

[10] W. Wernsdorfer *et al.*, Phys. Rev. Lett. **78**, 1791 (1997).

[11] See I. Eisenstein and A. Aharoni, Phys. Rev. B **16**, 1285 (1977), for an example of a problem where $P_s(t)$ exhibits multiexponential decay. Problems involving an ensemble of particles with a distribution of decay times also exhibit $P_s(t)$'s that do not decay as a single exponential, e.g. R. Street and J. C. Woolley, Proc. Phys. Soc., London, Sect. A **62**, 562 (1949); A. Aharoni, J. Appl. Phys. **57**, 4702 (1984); M. El-Hilo, K. O'Grady, and R. W. Chantrell, J. Magn. Magn. Mater. **109**, L164 (1992).

[12] Only static switching problems are considered here, not switching under time-dependent conditions, such as those produced by pulsed magnetic fields. See R. V. Kohn, M. G. Reznikoff, and E. Vanden-Eijnden (unpublished) and M. G. Reznikoff, Ph.D. thesis, New York University, New York 2004.

[13] W. F. Brown, Phys. Rev. **130**, 1677 (1963).

[14] L. D. Landau and E. M. Lifshitz, Phys. Z. Sowjetunion **8**, 153 (1935).

[15] T. L. Gilbert, Phys. Rev. **100**, 1243 (1955).

[16] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).

[17] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1997).

[18] W. E, W. Ren, and E. Vanden-Eijnden, Phys. Rev. B **66**, 052301 (2002); J. Appl. Phys. **93**, 2275 (2003).

[19] One might speculate that the initial increase of $P_s(t)$ from zero is bounded above by the function $e^{-C/t}/t^{\gamma}$ that describes the unbiased diffusion of probability density, where C and γ are t-independent parameters.

[20] This expansion makes clear that even if $P_s(t)$ does not have a peak, its decay will typically be multiexponential (i.e., described by the sum of two or more exponentials), until the long-time, single-exponential regime is reached. See Ref. [11].

[21] G. Grinstein and R. H. Koch, Phys. Rev. B **71**, 184427 (2005).

[22] N. Metropolis *et al.*, J. Chem. Phys. **21**, 1087 (1953).