Aging in the Random Energy Model

Gérard Ben Arous
DMA, EPFL, CH-1015 Lausanne, Switzerland

Anton Bovier
Weierstrass-Institut für Angewandte Analysis und Stochastik, Mohrenstrasse 39, D-10117 Berlin, Germany

Véronique Gayrard
DMA, EPFL, CH-1015 Lausanne, Switzerland
and CPT-CNRS, Luminy, Case 907, F-13288 Marseille Cedex 9, France
(Received 8 October 2001; published 11 February 2002)

The random energy model (REM) has become a key reference model for glassy systems. In particular, it is expected to provide a prime example of a system whose dynamics shows aging, a universal phenomenon characterizing the dynamics of complex systems. The analysis of its activated dynamics is based on so-called trap models, introduced by Bouchaud, that are also used to mimic the dynamics of more complex disordered systems. In this Letter we report the first results that justify rigorously the trap model predictions in the REM.

DOI: 10.1103/PhysRevLett.88.087201 PACS numbers: 75.10.Nr, 05.30.–d, 75.10.Hk, 75.10.Jm

A key concept that has emerged over the past years in the study of dynamical properties of complex systems is that of "aging." It is applied to systems whose dynamics are dominated by slow transients towards equilibrium (see [1] for an excellent review). This phenomena is manifest in a huge variety of systems, such as glasses, spin glasses, biomolecules, polymers, plastics, and has obvious practical implications in real-world applications.

When discussing aging dynamics, it is all important to specify the precise *time scales* considered in relation to the volume. The *nonactivated* regime is now well explored [2]. This note gives the first rigorous results for the *activated* aging regime, which corresponds to much longer time scales.

In "nonactivated dynamics" one considers the infinite volume limit at fixed time t, and *then* analyzes the ensuing dynamics as t tends to infinity. The typical setting where such a program has been carried out is Langevin dynamics of spherical mean field spin glasses, such as the p-spin Sherrington-Kirkpatrick (SK) model [3]. Note that even in this setting multiple and even infinitely many time-scales may be observed [e.g., in the SK model or the $(2 + 4)$-spin spherical SK model [1]]. In this context, mathematically rigorous results have been obtained recently in [4] only for the $p = 2$ spherical SK model.

Slow dynamics of complex systems is often attributed to the presence of "thermally activated" barrier crossing in the configuration space [5]. For instance, the standard picture of the spin glass phase typically involves a highly complex landscape of the free energy function exhibiting many nested valleys organized according to some hierarchical tree structure (see, e.g., [6,7]). To such a picture corresponds the heuristic image of a stochastic dynamics that, on time scales that diverge with the size of the system, can be described as performing a sequence of "jumps" between different valleys at random times whose rates are governed by the depths of the valleys and the heights of connecting passes or saddle points. To capture these Goldstein-type features, Bouchaud and others [1,8–11] have introduced an interesting ansatz, i.e., a mapping of the dynamics onto "trap models." These trap models are Markov jump processes on a state space that simply enumerates the valleys of the free energy landscape. While this picture is intuitively appealing, its derivation is based on knowledge obtained in much simpler contexts, such as diffusions in finite dimensional potential landscapes exhibiting a finite number of minima. In the systems one is interested in here, however, both dimension and number of minima are infinite or asymptotically growing to infinity.

Since trap models have become a key tool to model these dynamical phenomena in a large variety of systems, it is an important and interesting question to understand whether, how, and in which sense the long time dynamics of disordered systems such as spin glasses can really be described by trap models, and, in particular, to elucidate the precise time scales to which these models refer. To answer this question requires, of course, the study of the actual stochastic dynamics of the full system at diverging time scales, which is, in general, a very hard problem.

In this Letter, we report on the first rigorous results linking the long-time behavior of Glauber dynamics to a trap model in the context of the "simplest spin glass" [12] model, the random energy model (REM) [13,14]. While this model is surely far from "realistic," it offers a number of features that are "typical" for what one expects in real spin glasses, and its analysis involves already a good number of the problems one would expect to find in more realistic situations. The main advantage we will draw from this is, of course, the fact that the equilibrium properties of this model are perfectly understood.

 0031-9007/02/88(8)/087201(4)$20.00

The REM.—We recall that the REM [13,14] is defined as follows. A spin configuration σ is a vertex of the hypercube $S_N \equiv \{-1,1\}^N$. We define the family of independent and identically distributed standard Gaussian random variables $\{E_\sigma\}_{\sigma \in S_N}$. We define a random (Gibbs) probability measure on S_N, $\mu_{\beta,N}$, by setting

$$\mu_{\beta,N}(\sigma) \equiv \frac{e^{-\beta\sqrt{N}E_\sigma}}{Z_{\beta,N}}, \tag{1}$$

where $Z_{\beta,N}$ is the normalizing partition function.

It is well known [13,14] that this model exhibits a phase transition at $\beta_c = \sqrt{2\ln 2}$. For $\beta \le \beta_c$, the Gibbs measure is supported, asymptotically as $N \uparrow \infty$, on the set of states σ for which $E_\sigma \sim -\sqrt{N}\beta$, and no single configuration has positive mass. For $\beta > \beta_c$, on the other hand, the Gibbs measure gives positive mass to the extreme (minimal) elements of the order statistics of the family E_σ; i.e., if we order the spin configurations according to the magnitude of their energies such that

$$E_{\sigma^{(1)}} \le E_{\sigma^{(2)}} \le E_{\sigma^{(3)}} \le \cdots \le E_{\sigma^{(2^N)}}, \tag{2}$$

then for any finite k the respective mass $\mu_{\beta,N}(\sigma^{(k)})$ will converge, as N tends to infinity, to some positive random variable ν_k. In fact, the entire family of masses $\mu_{\beta,N}(\sigma^{(k)}), k \in \mathbb{N}$ will converge to a random process $\{\nu_k\}_{k\in\mathbb{N}}$, called Ruelle's point process [8,15].

So far, the fact that σ are vertices of a hypercube has played no role in our considerations. It will enter only in the definition of the *dynamics of the model*. The dynamics we consider is a *continuous time Glauber dynamics* $\sigma(t)$ on S_N with transition rates,

$$p_N(\sigma \to \eta) = e^{+\beta\sqrt{N}E_\sigma},$$

when σ and η differ by a single spin flip. Note that the dynamics is also random; i.e., the law of the Markov chain is a measure valued random variable on Ω that takes values in the space of Markov measures on the path space $S_N^{\mathbb{N}}$. We will mostly take a quenched point of view; i.e., we consider the dynamics for a given fixed realization of the disorder.

It is easy to see that this dynamics is reversible with respect to the Gibbs measure $\mu_{\beta,N}$. One also sees that it represents a nearest neighbor random walk on the hypercube with traps of random depths.

The REM-like trap model.—The idea suggested by the known behavior of the equilibrium distribution is that this dynamics, for $\beta > \beta_c$, will spend long periods of time in the states $\sigma^{(1)}, \sigma^{(2)}, \ldots$, etc. and will move "quickly" from one of these configurations to the next. Based on this intuition, Bouchaud *et al.* [8,9] proposed the "REM-like" trap model: Consider a continuous time Markov process Z_M whose state space is the set $S_M \equiv \{1,\ldots,M\}$ of M points, representing the M "deepest" traps. Each of the states is assigned a random variable ε_k (representing minus the energy of the state k), which is taken to be exponentially distributed with rate one. If the process is in state k, it waits an exponentially distributed time with mean proportional to $e^{\varepsilon_k \alpha}$, and then jumps with equal probability in one of the other states $k' \in S_M$.

This process can be analyzed using techniques from *renewal* theory. The point is that, if one starts the process from the uniform distribution, it is possible to show that the counting process, $c(t)$, that counts the number of jumps in the time interval $(0,t]$, is a classical renewal (counting) process [16]; moreover, as $M \uparrow \infty$, this renewal process converges to a renewal process with a *deterministic* law for the renewal time with a heavy-tailed distribution whose density is proportional to $t^{-1-1/\alpha}$, where $\alpha = \beta/\beta_c$.

The quantity that is used to characterize the "aging" phenomenon is the probability $\Pi_M(t,s)$ that during a time interval $[t, t+s]$ the process does not jump. Bouchaud and Dean [9] showed that, for $\alpha > 1$,

$$\lim_{M\uparrow\infty} \frac{\Pi_M(t,s)}{H_0(s/t)} = 1,$$

where the function H_0 is defined by

$$H_0(w) \equiv \frac{1}{\pi \operatorname{cosec}(\pi/\alpha)} \int_w^\infty dx \, \frac{1}{(1+x)x^{1/\alpha}}. \tag{3}$$

Note that $H_0(w)$ behaves similar to $1 - Cw^{1-1/\alpha}$, for small w, and similar to $Cw^{-1/\alpha}$ for large w.

Our purpose is to show, in a mathematically rigorous way, how and to what extent the REM-like trap model can be viewed as an approximation of what happens in the REM itself. To this end, we first introduce the set $T_M \equiv \{\sigma^{(1)},\ldots,\sigma^{(M)}\}$ of the first M states defined by the enumeration (2). Ideally, we would like to start with our original process $\sigma(t)$ and construct a new process Y_M as follows. Let $\tau_1, \tau_2, \ldots, \tau_n, \ldots$ be the sequence of times at which $\sigma(t)$ visits different elements of the (random) set T_M. Then set

$$X_{N,M}(t) \equiv \sum_{k=1}^{M} k \mathbf{1}_{\sigma^{(k)}=\sigma(\tau_n)} \mathbf{1}_{\tau_n \le t < \tau_{n+1}},$$

i.e., $X(t)$ takes the value k during time intervals at which the process $\sigma(t)$ "travels" from $\sigma^{(k)}$ to the next element of this set. We would like to say that Bouchaud's process Z_M approximates, after an N and an M dependent rescaling of the time, this process X, if N and M are large, i.e., that in some appropriate sense, for some function $c(N,M)$,

$$Z_M(t) \sim X_{N,M}(c(N,M)t),$$

when first N, then M, and finally t tend to infinity. This problem involves two main assumptions: (i) The process jumps with the uniform distribution from any state in T_M to any of the other states in T_M. (ii) The process observed on the set T_M is, at least asymptotically, a Markov process; in particular, the times between visits of two different states in T_M are asymptotically exponentially distributed. While it appears intuitively reasonable to accept these assumptions, (a) they are not at all easy to justify and (b) the second assumption is not even correct. In fact, we will see that such properties can be established only in a very weak asymptotic form, which is, however, just enough to imply that the predictions of Bouchaud's model apply to the long-time asymptotics of the process.

We will now present our main results. The full proofs of these results are given in [17,18]. We define instead of the set T_M introduced above the sets, for $E \in \mathbb{R}$,

$$T \equiv T_N(E) \equiv \{\sigma \in S_N | E_\sigma \leq -u_N(E)\},$$

where

$$u_N(x) \equiv \sqrt{2N \ln 2} - \frac{x}{\sqrt{2N \ln 2}} - \frac{\ln(N \ln 2) + \ln 4\pi}{2\sqrt{2N \ln 2}}.$$

We will call this set the "top." Note that $T_N(E) = T_{|T_N(E)|}$.

Our first result concerns just the "motion" of the process disregarding time. Let $\xi^1, \dots, \xi^{|T(E)|}$ be an enumeration of the elements of $T(E)$. Now define (for fixed N and E) the stochastic process X_ℓ with state space $\{1, \dots, |T(E)|\}$ and discrete time $\ell \in \mathbb{N}$ by

$$Y_\ell^{(N)} = X(\tau_\ell). \tag{4}$$

It is easy to see that $Y_\ell^{(N)}$ is a Markov process whose transition probabilities $p(i \to j)$ are nothing but the probabilities that the original process $\sigma(t)$ starting at ξ^i hits T first in the point ξ^j. It will be convenient to fix the random size of the state space of this process by conditioning. Thus, set $P_M(\cdot) \equiv P(\cdot \| T(E)| = M)$.

It is clear that under this law the processes $Y_\ell^{(N)}$ are Markov chains. The first result that is proven in [17] is that, if $\beta/\sqrt{2 \ln 2} > 1$, the transition probabilities of this chain converge, as $N \uparrow \infty$, to the uniform law:

$$p_M^*(i \to j) = \begin{cases} \frac{1}{M-1} & \text{if } i \neq j \\ 0, & \text{if } i = j. \end{cases}$$

In other words, the process that records the visits of the initial chain on the sites of $T_N(E)$ is asymptotically the same as in the trap model.

The second result concerns two key quantities: $\mathcal{T}(\eta)$, the mean time to reach $T_N(E)\backslash\eta$ starting from any point $\eta \in S_N$, and $Z(\eta, \bar{\eta})$, the mean time to reach $\bar{\eta} \in T_N(E)\backslash\eta$ before any other point in $T_N(E)$. First, one shows that, if $\alpha > 1$, for $\eta \in T_N(E)$, $Z(\eta, \bar{\eta}) \sim \mathcal{T}(\eta)$; i.e., the traveling time between two points in $T_N(E)$ depends only on the starting point. Second, $\mathcal{T}(\eta)$ are estimated very precisely as

$$\mathcal{T}(\eta) = \frac{M}{M-1} [e^{\beta\sqrt{N}E_\eta} + \mathcal{W}_{N,E}](1 + O(1/N)), \tag{5}$$

where $\mathcal{W}_{N,E}$ is a random variable of mean value,

$$\mathbb{E}\mathcal{W}_{N,E} = \frac{e^{-\beta\sqrt{N}\, u_N(E)}}{\alpha - 1},$$

and whose standard deviation is negligible compared to the mean, as E tends to $+\infty$. If $\alpha < 1$, on the other hand, the estimate on $\mathcal{T}(\eta)$ becomes independent on η as well and we get

$$\mathcal{T}(\sigma) = \frac{M}{M-1} e^{N(\beta^2/2 + \ln 2)}(1 + O(1/N)).$$

Remark: We see that for E very large, $\mathcal{W}_{N,E} \sim e^{\beta\sqrt{N}\, u_N(E)}$ represents a natural time scale for the process on $T(E)$. Thus (5) implies that for all $\sigma \notin T(E)$ the mean time of arrival in the top is proportional to $e^{\beta\sqrt{N} u_N(E)}$. On the other hand, there exists $\eta \in T(E)$ such that the mean time of the first exit from η, which is $e^{\beta\sqrt{N}E_\eta}$, is just of this order. Thus, the slowest times of exit from a state in $T(E)$ are of the same order as the time it takes to reach $T(E)$. This can be expressed by saying that on the average the process takes a time t to reach states that have an exit time t. This is a first manifestation of the aging phenomenon. In contrast, if $\alpha < 1$, for all $\sigma \in T(E)$, $\mathcal{T}(\sigma) \ll \sup_{\eta \in S_N} \mathcal{T}(\eta)$, and thus the time spent in the top states is irrelevant compared to the time between successive visits of such states. Thus, we see a clear distinction between the high and the low temperature phase of the REM on the dynamical level. Notice that, as has been observed in [19], the dynamical phase transition is not accompanied by a qualitative change in the spectral gap, which in all cases is related to the largest exit times. For related results on the high temperature dynamics, see also [20].

The fact that the mean times of passage from a state $\eta \in T(E)$ to another state $\bar{\eta} \in T(E)$ are asymptotically independent of the terminal state $\bar{\eta}$ confirms to some extent the heuristic picture of Bouchaud. Indeed, if in addition we added the hypothesis that the process observed on the top is Markovian, this would imply that the waiting times must be exponentially distributed with rates independent of the terminal state and given by $Z(\eta, \bar{\eta})$. This, however, cannot be justified. The reason for the failure of the Markov properties can be traced to the fact that the time spent outside of $T(E)$ when traveling between two states of $T(E)$ cannot be neglected in comparison to the waiting time in the starting point, which in turn is a manifestation of the *absence* of a true *separation of time scales*.

We now turn to a more precise analysis of the aging phenomenon.

The natural generalization of Bouchaud's correlation function $\Pi_M(t, s)$ is the probability that the process does not jump from a state in the top to another state in the top during a time interval of the form $[n, n + m]$. There is some ambiguity as to how this should be defined precisely, but the most convenient definition is to define $\Pi_\sigma(n, m)$ as the probability that the process starting at time 0 in σ does not jump during the time interval $[n, n + m]$ from one state in T to another state in T.

Of course we still have to specify the initial distribution. To be as close as possible to Bouchaud, the natural choice is the uniform distribution on $T_N(E)$ that we will denote by π_E. The natural correlation function is then

$$\Pi(n,m) \equiv \frac{1}{|T_N(E)|} \sum_{\sigma \in T_N(E)} \Pi_\sigma(n,m).$$

The following theorem establishes the connection to the trap model.

Theorem. Let $\beta > \sqrt{2 \ln 2}$. Then there is a sequence $c_{N,E} \sim \exp[\beta \sqrt{N} u_N(E)]$ such that, for any $\epsilon > 0$,

$$\lim_{t,s \uparrow \infty} \lim_{E \uparrow \infty} \lim_{N \uparrow \infty} P\left[\left| \frac{\Pi(c_{N,E}t, c_{N,E}s)}{H_0(s/t)} - 1 \right| > \epsilon \right] = 0,$$

where H_0 is defined in (3).

Remark: Note that the rescaling of the time by the factor $c_{N,E}$ ensures that we are observing the system first on the proper time of equilibration as N goes to infinity, and that then, as E tends to infinity, we measure time on a scale at which the equilibration time diverges. Thus, the trap model describes the large time asymptotics on the "last scale" before equilibrium is reached. This is to be contrasted to the other extreme where the infinite volume limit is taken with a fixed time scale, as in [2,3].

The proof of this theorem is based on the techniques of [21,22]; it is rather involved and is the main content of [18]. It may be instructive to see a brief outline of our approach. Basically, the idea is to mimic the proof in the trap model and to set up a renewal equation. Now it is easy to derive a renewal equation for the quantities $\Pi_\sigma(n,m)$. However, in contrast to the situation of the trap model, it is not possible to obtain a single closed equation for $\Pi(m,n)$. This means that we actually have to study a system of renewal equations which renders the proof rather complicated. The key ingredients then are precise estimates of the Laplace transforms of the probability distributions entering the renewal equations in the complex plane. What makes the final result emerge is then the fact that, in the neighborhood of the origin (which corresponds to large times), the Laplace transforms have almost the desired properties that would lead to such a closed equation. This makes it possible to employ perturbation expansions to prove the theorem.

Remark: We conclude the paper with a remark on the role of the particular choice of the transition probabilities (1) depending only on starting points. Clearly, these favor the proximity to Bouchaud's model. However, our reason to choose them is that they avoid having to solve the problem of determining very precisely the barrier heights between two points in T_N, which is in general a tremendous random geometric problem to which we have no answer. It is possible, although not certain, that different choices of a reversible dynamics might lead to different "trap models," all of which should, however, show aging.

We thank J.-Ph. Bouchaud, L. Cugliandolo, and J. Kurchan for helpful discussions and encouraging comments.

[1] J. P. Bouchaud, L. Cugliandolo, J. Kurchan, and M. Mézard, in *Spin-Glasses and Random Fields,* edited by A. P. Young (World Scientific, Singapore, 1998).
[2] L. Cugliandolo and J. Kurchan, Phys. Rev. Lett. **71**, 173 (1993).
[3] L. Cugliandolo and D. Dean, J. Phys. A **28**, 4213 (1995).
[4] G. Ben Arous, A. Dembo, and A. Guionnet, Probab. Theory Relat. Fields **120**, 1 (2001).
[5] M. Goldstein, J. Chem. Phys. **51**, 3728 (1969).
[6] G. Biroli and J. Kurchan, Phys. Rev. E **64**, 6101 (2001).
[7] J. F. Fontanari and P. F. Stadler, cond-mat/0106366, 2001.
[8] J. P. Bouchaud, J. Phys. I (France) **2**, 1705 (1992).
[9] J. P. Bouchaud and D. Dean, J. Phys. I (France) **5**, 265 (1995).
[10] C. Monthus and J. P. Bouchaud, J. Phys. A **29**, 3847 (1996).
[11] B. Rinn, P. Maass, and J. P. Bouchaud, Phys. Rev. Lett. **84**, 5403 (2000).
[12] D. J. Gross and M. Mézard, Nucl. Phys. **B240**, 431 (1984).
[13] B. Derrida, Phys. Rev. Lett. **45**, 79 (1980).
[14] B. Derrida, Phys. Rev. B **24**, 2613 (1981).
[15] D. Ruelle, Commun. Math. Phys. **108**, 225 (1987).
[16] S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes* (Academic, Boston, 1975).
[17] G. Ben Arous, A. Bovier, and V. Gayrard, WIAS-preprint 690, 2001 (http://mpej.unige.ch/mp_arc-bin/mpa?yn=01-374).
[18] G. Ben Arous, A. Bovier, and V. Gayrard, WIAS-preprint 691, 2001 (http://mpej.unige.ch/mp_arc-bin/mpa?yn=01-375).
[19] L. R. G. Fontes, M. Isopi, Y. Kohayakawa, and P. Picco, Ann. Appl. Probab. **8**, 917 (2001).
[20] P. Mathieu and P. Picco, CPT-2000/P.3982, 2000.
[21] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein, Probab. Theory Relat. Fields **119**, 99 (2001).
[22] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein (to be published).