Weierstraß–Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Aging in the random energy model

Gérard Ben Arous[1] Anton Bovier[2], Véronique Gayrard[3]

submitted: 16th October 2001

[1] DMA
EPFL
CH-1015 Lausanne
Switzerland
E-Mail:Gerard.Benarous@epfl.ch

[2] Weierstrass-Institut
für Angewandte Analysis
und Stochastik,
Mohrenstrasse 39,
D-10117 Berlin, Germany
E-Mail: bovier@wias-berlin.de

[3] DMA
EPFL
CH-1015 Lausanne, Switzerland,
and CPT-CNRS,
Luminy, case 907,
F-13288 Marseille, cedex 9, France
E-Mail: Veronique.Gayrard@epfl.ch
Veronique.Gayrard@cpt.univ-mrs.fr

Preprint No. 692
Berlin 2001

Key words and phrases. aging, Glauber dynamics, random energy model, trap models, metastability, extreme values.

PACS: 75.10.Nr, 75.10.Jm, 75.10.Hk, 05.30.-d.

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, Case907, F-13288 Marseille Cedex 9, France

(October 16, 2001)

In this letter we announce rigorous results on the phenomenon of aging in the Glauber dynamics of the random energy model and their relation to Bouchaud's 'REM-like' trap model. We show that, below the critical temperature, if we consider a time-scale that diverges with the system size in such a way that equilibrium is almost, but not quite reached on that scale, a suitably defined autocorrelation function has the same asymptotic behaviour than its analog in the trap model.

PACS: 75.10.Nr, 75.10.Jm, 75.10.Hk, 05.30.-d

Dynamical properties of complex disordered systems such as spin glasses, have been at the center of interest among theoretical physicists working in statistical mechanics. A key concept that has emerged over the last years is that of "aging": it is applied to systems whose dynamics are dominated by slow transients towards equilibrium. When discussing aging dynamics, it is all important to specify the precise *time-scales* considered in relation to the large volume limit. In "glassy dynamics" (see [6] for an excellent review) one considers the infinite volume limit at fixed time t, and *then* analyses the ensuing dynamics as t tends to infinity. The typical setting where such a program has been carried out is Langevin dynamics of sperical mean field spin glasses, such as the p-spin SK model [13]. 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 [6]). In this context, mathematically rigorous results have been obtained recently in [8] only for the $p=2$ spherical SK-model.

On a heuristic level, one expects that glassy dynamics describes the evolution of a system that is approaching "local equilibrium", or nearest local minima of the free energy. However, 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. [2,17]). 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 those rates are governed by the depths of the valleys and the heights of connecting passes or saddle points. To capture these features which are obviously beyond the realm of glassy dynamics, Bouchaud and others [1,7,11,12,6] have introduced an interesting ansatz. Their so-called "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.

It is thus 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 note we report on the first rigorous results linking the long-time behaviour of Glauber dynamics to a trap model in the context of the "simplest spin glass" [19] model, the random energy model (REM) [14,15]. 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.

The REM. We recall that the REM [14,15] is defined as follows. A spin configuration σ is a vertex of the hypercube $\mathcal{S}_N \equiv \{-1,1\}^N$. We define the family of i.i.d. standard Gaussian random variables $\{E_\sigma\}_{\sigma\in\mathcal{S}_N}$. We define a random (Gibbs) probability measure on $\mathcal{S}_N$, $\mu_{\beta,N}$, by setting

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

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

It is well-known [14,15] that this model exhibits a phase transition at $\beta_c = \sqrt{2\ln 2}$. For $\beta \leq \beta_c$, 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 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 s.t.

$$E_{\sigma^{(1)}} \geq E_{\sigma^{(2)}} \geq E_{\sigma^{(3)}} \geq \ldots \geq E_{\sigma^{(2^N)}} \tag{1}$$

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 $\mu_{\beta,N}(\sigma^{(k)}), \kappa \in I\!N$ will converge to a random process $\{\nu_k\}_{k\in I\!N}$, called Ruelle's point process [21,1].

So far the fact that σ are vertices of a hypercube has played no rôle in our considerations. It will enter only in the definition of the *dynamics of the model.* The dynamics we consider is a *Glauber dynamics* $\sigma(t)$ on $\mathcal{S}_N$ with transition rates

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

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 $\mathcal{S}_N^{I\!N}$. 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 $\mu_{\beta,N}$. On 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. [1,7] 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 positive random energy E_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 $e^{E_k\alpha}$, 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 [18]; moreover, as $n \uparrow \infty$, 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 [7] 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

$$H_0(w) \equiv \frac{1}{\pi \operatorname{cosec}(\pi/\alpha)} \int_w^\infty dx \frac{1}{(1+x)x^{1/\alpha}} \tag{2}$$

Note that $H_0(w)$ behaves like $1 - Cw^{1-1/\alpha}$, for small w, and like $Cw^{-1/\alpha}$ 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 $T_M \equiv \{\sigma^{(1)}, \dots, \sigma^{(M)}\}$ of the first M states defined by the enumeration 1. Ideally, we would like to start with our original process $\sigma(t)$ and construct a new process Y_M as follows. Let $\tau_1, \tau_2, \dots, \tau_n, \dots$ be the sequence of times at which $\sigma(t)$ visits different elements of the (random) set T_M. Then set

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

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 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:
1) The process jumps with the uniform distribution from any state in T_M to any of the other states in T_M, and
2) 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, they are a) 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 only established 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 will be given in two forthcoming papers [4,5].

We define instead of the set T_M introduced above the sets, for $E \in \mathbb{R}$,

$$T \equiv T_N(E) \equiv \left\{\sigma \in \mathcal{S}_N \middle| E_\sigma \geq u_N(E)\right\}$$

where

$$u_N(x) \equiv \sqrt{2N \ln 2} + \frac{x}{\sqrt{2N \ln 2}} - \frac{1}{2} \frac{\ln(N \ln 2) + \ln 4\pi}{\sqrt{2N \ln 2}}$$

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

$$Y_\ell^{(N)} = X(\tau_\ell) \tag{3}$$

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

Theorem 1 *Let $\sigma(n)$ denote the Markov chain with transition matrix defined in (1.3) and whose initial distribution is the uniform distribution on $\mathcal{S}_N$. Let $Y_\ell^{(N)}$ be the Markov process defined by 3. Let Y_ℓ denote the Markov chain on $\{1, \ldots, M\}$ with transition matrix p^*_M given by*

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

and initial distribution $p^*_M(i) = 1/M$. Then, for all $M \in I\!N$,

$$Y^{(N)} \xrightarrow{\mathcal{D}} Y, \quad P_M\text{-a.s.}$$

Remark: Note that the statement of the theorem implies the convergence in law (w.r.t. P) of the probability distribution of $Y^{(N)}$.

The next results concern the mean times of the motion towards the top and between points of the top.

Theorem 2. *Assume that $\alpha \equiv \beta/\sqrt{2\ln 2} > 1$. Then on a set of probability one for all large enough N, the following holds:*

(i) For all $\eta \in \mathcal{S}_N$, let $\mathcal{T}(\eta)$ denote the mean time the process starting in η takes to reach $T_N(E)\backslash\eta$. Then

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

(ii) For all $\eta, \bar{\eta} \in T(E)$, the expected time, $Z(\eta, \bar{\eta})$, to reach $\bar{\eta}$ starting from η conditioned on the event that $\bar{\eta}$ is the first site different from η in $T_N(E)$ that is reached, is almost independent of $\bar{\eta}$, more precisely

$$|Z(\eta,\bar{\eta}) - \mathcal{T}(\eta)| \leq \frac{1}{1-\frac{1}{M}}\mathcal{W}_{N,E}O(1/N) \tag{4}$$

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

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

and whose standard deviation is negligible compared to the mean, as E tends to $-\infty$.

(iii) If $\alpha < 1$, then, for all $\sigma \in \mathcal{S}_N$,

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

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 (i) 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 first exit from η which is $\sqrt{N}E_\eta$, 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 $\sigma \in T(E)$, $I\!E\tau^\sigma_{T(E)\backslash\sigma} \ll \sup_{\eta\in\mathcal{S}_N} I\!E\tau^\eta_{\mathcal{S}_N\backslash\eta}$, and thus the time spent in 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 [16], 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].

Remark: Statement iii) of Theorem 2 expresses 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 $\bar{\eta}$. This 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, then the two preceeding theorems would immediately imply that the waiting times must be exponentially distributed with rates independent of the terminal state and given by 4. 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 travelling 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 3. *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 $\epsilon > 0$*

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

where H_0 is defined in 2.

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 minus 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 of "glassy dynamics", when the infinite volume limit is taken with a fixed time scale.

The proof of this theorem is rather involved and is the main content of [5]. 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.

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

[1] J.P. Bouchaud, J. Phys. I (France), **2**, 1705 (1992).
[2] G. Biroli and J. Kurchan, Phys. Rev. **E 64**, 6101 (2000).
[3] A. Bovier, "Statistical mechanics of disordered systems", MaPhySto Lecture Notes 10, Aarhus, 2001.
[4] G. Ben Arous, A. Bovier, and V. Gayrard, preprint (2001).
[5] G. Ben Arous, A. Bovier, and V. Gayrard, preprint (2001).

[6] J.P. Bouchaud, L. Cugliandolo, J. Kurchan, M. Mézard, "Out-of-equilibrium dynamics in spin-glasses and other glassy systems", in *Spin-glasses and random fields* (A.P. Young, Ed.). World Scientific, Singapore (1998).
[7] J.P. Bouchaud and D. Dean, J. Phys. I (France), **5**, 265 (1995).
[8] G. Ben Arous, A. Dembo and A. Guionnet, Probab. Theor. Rel. Fields **120**, 1-67 (2001).
[9] A. Bovier, M. Eckhoff, V. Gayrard and M. Klein, Prob. Theor. Rel. Fields **119**, 99-161 (2001).
[10] A. Bovier, M. Eckhoff, V. Gayrard and M. Klein, to appear in Commun. Math. Phys. (2001).
[11] J.P. Bouchaud, C. Monthus, J. Phys. A-Math. Gen. **29**, (14) 3847-3869 (1996).
[12] J.P. Bouchaud, P. Maass, B. Rinn, Phys. Rev. Letts. **84**, 5403-5406 (2000).
[13] L. Cugliandolo and D. Dean, J. Phys. A - Math. Gen. **28**, (15) 4213-4234 (1995).
[14] B. Derrida, Phys. Rev. Letts. **45**, 79-82 (1980).
[15] B. Derrida, Phys. Rev. B **24**, 2613-2626 (1981).
[16] L.R.G. Fontes, M. Isopi, Y. Kohayakawa, and P. Picco, Ann. Appl. Probab. **8**, 917-943 (2001).
[17] J.F. Fontanari, P.F. Stadler, preprint cond-mat/0106366 (2001).
[18] S. Karlin and H.M. Taylor, *A first course in stochastic processes*, Academic Press, Boston, 1975.
[19] D.J. Gross and M. Mézard, Nucl. Phys. **B 240**, 431-452 (1984).
[20] P. Mathieu and P. Picco, preprint CPT-2000/P.3982 (2000).
[21] D. Ruelle, Commun. Math. Phys. **108**, 225-239 (1987).