Universality of REM-like aeging in mean field spin glasses

Gérard Ben Arous
Courant Institute, New York University, 251 Mercer St., New York, NY 10012-1185

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

Jiří Černý
Eidgenössische Technische Hochschule Zürich, Rämistrasse 101, 8092 Zürich, Switzerland
(Dated: November 8, 2007)

Ageing has become the paradigm to describe dynamical behavior of glassy systems, and in particular spin glasses. Trap models have been introduced as simple caricatures of effective dynamics of such systems. In this Letter we show that in a wide class of mean field models and on a wide range of time scales, ageing occurs precisely as predicted by the REM-like trap model of Bouchaud and Dean. This is the first rigorous result about ageing in mean field models except for the REM and the spherical model.

PACS numbers: 75.10.Nr, 75.10.Jm, 75.10.Hk, 05.30.-d XXX remove 05.30.-d and 75.10.Jm ? XXX

A key concept that has emerged over the last 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, bio-molecules, 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. In the *non-activated* regime one studies the infinite volume limit at fixed time t, and *then* analyzes the ensuing dynamics as t tends to infinity. This *non-activated* regime is now well explored [2] XXX More refs? XXX.

For longer time scales, that is times diverging with the volume of the system, a full picture is largely missing. The slow dynamics of complex systems in such time scales is often attributed to the presence of "thermally activated" barrier crossing in the configuration space [3]. 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. [4, 5]). 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 Goldstein's type features, Bouchaud and others [1, 6–9] have introduced an interesting ansatz, that is 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.

In a series of papers [10–13] a systematic investigation of a variety of trap models was initiated. In this process, it emerged that there appears to be an almost universal ageing mechanism based on α-stable subordinators that governs ageing in most of the trap models.

In contrast, very little has been done concerning the derivation of trap-model dynamics from stochastic dynamics of even moderately realistic spin-glasses, such as the p-spin interaction SK models. The only case where this has been achieved so far is the simplest of these models, the Random Energy Model (REM) of Derrida with a particular form of the transition rates. In [14–16] this was achieved by a very detailed analysis of the dynamics at time scales just before the equilibration time, and at temperatures below the critical one. This result relied, in particular, on the detailed understanding of the equilibrium distribution of this model. More recently, in [12], the same model was analyzed at much shorter (but still exponentially large) time scales. It emerged that the same aging mechanism is in place there and that aging could also occur above the critical temperature.

All these works made crucial use of the independence of energies of different spin configurations assumed in the definition of the REM. In the present Letter, for the first time, we present aging results in a model with correlated energies, the p-spin interaction spin glass model. Quite surprisingly, the results obtained point to the validity of the REM-like trap model as universal aging mechanism.

The p-spin SK model. We recall that the p-spin SK model is is defined as follows. A spin configuration σ is

---

*AB's research in supported in part by the DFG in the Durch-German Bilateral Research Group "Mathematics of random spatial models from physics and biology".

a vertex of the hypercube $\mathcal{S}_N \equiv \{-1,1\}^N$. The Hamiltonian is given by

$$H_N(\sigma) \equiv -\frac{1}{N^{(p-1)/2}} \sum_{i_1,\dots,i_p=1}^{N} J_{i_1\dots i_p}\sigma_{i_1}\dots\sigma_{i_p}, \qquad (1)$$

where $J_{i_1\dots i_p}$ are independent Gaussian random variables with mean zero and variance one. As a consequence, we can describe the Hamiltonian as a centered normal process indexed by $\mathcal{S}_N$ with covariance

$$\mathbb{E}[H_N(\sigma)H_N(\tau)] = NR_N(\sigma,\tau)^p, \qquad (2)$$

where $R_N(\sigma,\tau)$ denotes as usual the normalized overlap, $R_N(\sigma,\tau) \equiv N^{-1}\sum_{i=1}^N \sigma_i\tau_i$. We define a random Gibbs measure on $\mathcal{S}_N$, $\mu_{\beta,N} \equiv Z_{\beta,N}^{-1}e^{-\beta H_n(\sigma)}$. Note that in the limit $p \uparrow \infty$ one recovers the random energy model [17], where $H_N(\sigma)$ are i.i.d. Gaussian random variables with variance N.

Dynamics. We consider a continuous time Markov dynamics $\sigma_N(t)$ on $\mathcal{S}_N$ whose transition rates are

$$p_N(\sigma,\sigma') = N^{-1}e^{\beta H_N(\sigma)} \qquad (3)$$

if σ and σ' are related by flipping a single spin, and are zero otherwise. 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.

It is thus useful to look at this dynamics as at a time change of a simple unbiased discrete-time random walk $Y_N(k)$, $k \in \mathbb{N}$, on $\mathcal{S}_N$ started at some fixed point of $\mathcal{S}_N$, say at $\{1,\dots,1\}$: We define the clock-process by

$$S_N(k) = \sum_{i=0}^{k-1} e_i \exp\big\{-\beta H_N\big(Y_N(i)\big)\big\}, \qquad (4)$$

where $(e_i, i \in \mathbb{N})$ is a sequence of mean-one i.i.d. exponential random variables. Then $\sigma_N(t)$ can be written as

$$\sigma_N(t) = Y_N(S_N^{-1}(t)). \qquad (5)$$

$S_N(k)$ is the instant of the k-th jump of $\sigma_N(t)$.

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 $\sigma^{(1)}, \sigma^{(2)}, \dots$ etc. and will move "quickly" from one of these configurations to the next. Based on this intuition, Bouchaud et al. [6, 7] proposed the "REM-like" trap model: Consider a continuous time Markov process Z_M whose state space is the set $K_M \equiv \{1,\dots,M\}$ 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 $e^{\beta\varepsilon_k}$, and then jumps with equal probability in one of the other states $k' \in K_M$.

The quantity that is used to characterize the "aging" phenomenon is the probability $\tilde{\Pi}_M(t,s)$ that during a time-interval $[t, t+s]$ the process does not jump. Bouchaud and Dean [7] showed that, for $\beta > 1$,

$$\lim_{s,t\uparrow\infty}\lim_{M\uparrow\infty} \frac{\tilde{\Pi}_M(t,s)}{H_{1/\beta}(s/t)} = 1, \qquad (6)$$

where the function H_α is defined by

$$H_\alpha(w) \equiv \frac{\sin(\pi\alpha)}{\pi}\int_w^\infty \frac{dx}{(1+x)x^\alpha} \qquad (7)$$

The dynamics of the REM-like trap model can be seen as a time change of a simple random walk $\tilde{Y}_M$ on the "complete graph" K_M by the clock process, $\tilde{S}_M(k) = \sum_{i=0}^{k-1} e_i \exp\{\beta\varepsilon_{Y_M(i)}\}$. As explained in [12], the result (6) can be deduced from the stronger claim

$$\lim_{n\uparrow\infty}\lim_{M\uparrow\infty} n^{-\beta}\tilde{S}_M(nt) = cV_{1/\beta}(t), \qquad t \ge 0, \qquad (8)$$

where $V_\alpha(t)$ is the α-stable subordinator with Laplace transform $\mathbb{E}[e^{-\lambda V_\alpha(t)}] = \exp(-t\lambda^\alpha)$.

The REM. In [14–16] it was confirmed that the REM-like picture is correct, at least for the dynamics defined in (3). This result was further extended to shorter time scales in [12] where the point of view of (8) was put in the foreground. Namely, it was shown that the clock process converges again to the stable subordinator: For every γ, $0 < \gamma < \min(\beta^2, \beta\sqrt{2\log 2})$,

$$\lim_{N\uparrow\infty} e^{-\gamma N} S_N(tN^{1/2}e^{N\gamma^2/2\beta^2}) = cV_{\gamma/\beta^2}(t). \qquad (9)$$

This implies then a similar aging result as in (6), $\Pi(te^{\gamma N}, se^{\gamma N}) \xrightarrow{N\to\infty} H_{\gamma/\beta^2}(s/t)$, as in the REM-like trap model, for the probability $\Pi_N(t,s)$ that $\sigma_N(t)$ does not jump between t and $t+s$.

We will now present our main results for the p-spin SK model. The full proofs of these results are given in [18]. First, since the valleys in the free-energy landscape contains more than one configuration, we should change the two-point function Π. We set

$$\Pi_N^\varepsilon(t,s) = \mathbb{P}\{R_N\big(\sigma_N\big(te^{\gamma N}\big), \sigma_N\big((t+s)e^{\gamma N}\big)\big) \ge 1-\varepsilon\}, \qquad (10)$$

that is the overlap at two far-distant time instants is exceptionally large. Then, the similar result as in the REM holds, at least if $p \ge 3$. Namely, for all

$$0 < \gamma < \min(\beta^2, \zeta(p)\beta), \qquad (11)$$

$\varepsilon \in (0,1)$, $t > 0$, and $s > 0$,

$$\lim_{N\to\infty} \Pi_N^\varepsilon(t,s) = H_{\gamma/\beta^2}(s/t). \qquad (12)$$

The basis of this result is again the statement analogous to (9) that shows that the properly rescaled clock-process converges to a stable subordinator.

The function $\zeta(p)$ used in (11) to limit the considered time scales is increasing and it satisfies

$$\zeta(3) \simeq 1.0291 \quad \text{and} \quad \lim_{p\to\infty} \zeta(p) = \sqrt{2\log 2}, \quad (13)$$

hence in the limit $p \uparrow \infty$ we recover the result for the REM.

Let us discuss the meaning of these results. $e^{\gamma N}$ is the time-scale at which we want to observe the process. According to the theorem, at this time the random walk will make of the order of $N^{1/2}e^{N\gamma^2/2\beta^2} \ll e^{\gamma N}$ steps. This number is also much smaller than 2^N (as follows from (11)) which has two consequences. First, the process has not enough time to explore the state space and is still out of equilibrium. Second, and more importantly, for such number of steps, the simple random walk is extremely "transient", in the sense that (i) starting from a given point x, for a number of steps $\nu \sim N^\omega$, $\omega < 1$, the distance from x grows essentially linearly with speed one, that is there are no backtrackings with high probability; (ii) the SRW will *never* return to a neighborhood of the starting point, with high probability. The upshot is that we can think of the trajectory of the SRW essentially as of a straight line.

If the process H_N was i.i.d., as in the REM, then the maximum of H_N along the trajectory would be $\left(2N\ln(N^{1/2}e^{N\gamma^2/2\beta^2})\right)^{1/2} \sim \gamma N/\beta$, and the time spent in that site would be of order $e^{\gamma N}$, which is comparable to the total time. Formula (9) then states that the total accumulated time is composed of pieces of order $e^{\gamma N}$ that are collected along the trajectory. In fact, each jump of the subordinator corresponds to one visit to a "trap" that has waiting times of that order.

Our result states that in the p-spin model, the same is essentially true. The difference will be that the "traps" here will not consist of a single site, but consist of a deep valley (along the trajectory) that has approximately the same depth and whose shape and width we can describe quite precisely. Remarkably, each valley is essentially of a size independent of N (that is the number of sites contributing significantly to the residence time in the valley is essentially finite), and different valleys are statistically independent.

The fact that traps are finite may appear quite surprising to those familiar with the statics of p-spin models. From the results there (see [19, 20]), one knows that the Gibbs measure concentrates on "lumps" whose radius is of order $N\varepsilon_p$, with $\varepsilon_p > 0$. The mystery is however solved easily: Around a local minimum σ_0 with $H_N(\sigma_0) \sim \gamma N/\beta$, the process $H_N(\sigma)$ does grows essentially linearly with the distance $d(\sigma_0, \sigma)$ from the minimum, $\mathbb{E}[H_N(\sigma) - H_N(\sigma_0)] \sim c(p, \gamma, \beta)d(\sigma_0, \sigma)$. Therefore, the Gibbs mass decreases exponentially with $d(\sigma_0, \sigma)$. For the support of the Gibbs measure, one needs to take into account the entropy, that is that the volumes of the balls of radius r increases like $\exp(N(\ln 2 - I(1 - r/N)))$. For the dynamics, at least at our time-scales, this is, however, irrelevant, since the simple random walk leaves a local minimum essentially ballistically.

XXX I stoped here XXX Next we consider the Gaussian process restricted to the SRW trajectory. We expect that the main contributions to the sums $S_N(k)$ come from places where Y_N is maximal (on the trajectory). We expect that the distribution of these extremes do not feel the correlation between points a distance farther than ν apart. On the other hand, for points closer than ν, the correlation function $R_N(Y_N(i), Y_N(j))^p$ can be well approximated by a linear function $1 - 2p|i - j|/N$ (using that $R_N(Y_N(i), Y_N(j)) \sim 1 - 2|i - j|/N$). This is convenient since this process has an explicit representation in terms of i.i.d. random variables that allow for explicit computations (in fact, this is one of the famous Slepian processes for which the extremal distribution can be computed explicitly [24, 25]). Thus the idea is to cut the SRW trajectory into blocks of length ν and to replace the original process $H_N(Y_N(i))$ by a new one U_i, where U_i and U_j are independent, if i, j are not in the same block, and $\mathbb{E}[U_iU_j] = 1 - 2p|i - j|/N$ if they are. For the new process, the convergence result of formula 9 is relatively straightforward. The main step is the computation of Laplace transforms. Comparing the real process with the auxiliary one is the bulk of the work.

The proof of our results relies on the combination of detailed information on the properties of simple random walk on the hypercube, which is provided in Section 4 (but see also [21–23]), and comparison of the process H_N on the trajectory of the SRW to a simpler Gaussian process using interpolation techniques la Slepian, familiar from extreme value theory of Gaussian processes.

Remark: We conclude the Letter with a remark on the rôle of the particular choice of the transition probabilities (3) depending only on starting points. Clearly these favor the proximity to Bouchaud's model. For us, on a technical level, the independence of the random walk trajectory of the random environment defined by the Hamiltonian is crucial. Even in the case of the REM, we do not know at this point how to deal with different dynamics. This problem remains one of the great challenges in the field.

[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. F. Cugliandolo and J. Kurchan, Phys. Rev. Lett. **71**, 173 (1993).

[3] M. Goldstein, The Journal of Chemical Physics **51**, 3728 (1969).

[4] G. Biroli and J. Kurchan, Physical Review E **6401**, 016101 (2001).

[5] J. F. Fontanari and P. F. Stadler, J. Phys. A. **35**, 1509

(2002).
[6] J.-P. Bouchaud, J. Phys. I (France) **2**, 1705 (1992).
[7] J.-P. Bouchaud and D. S. Dean, J. Phys I(France) **5**, 265 (1995).
[8] C. Monthus and J.-P. Bouchaud, J. Phys. A **29**, 3847 (1996).
[9] B. Rinn, P. Maass, and J.-P. Bouchaud, Phys. Rev. Lett **84**, 5403 (2000).
[10] G. Ben Arous and J. Černý, Ann. Appl. Probab. **15**, 1161 (2005).
[11] G. Ben Arous, J. Černý, and T. Mountford, Probab. Theory Related Fields **134**, 1 (2006).
[12] G. Ben Arous and J. Černý, *The arcsine law as a universal aging scheme for trap models*, to appear in Comm. Pure Appl. Math. (2008).
[13] G. Ben Arous and J. Černý, Ann. Probab. **35**, 2356 (2007).
[14] G. Ben Arous, A. Bovier, and V. Gayrard, Physical Review Letters **88**, 087201 (2002).
[15] G. Ben Arous, A. Bovier, and V. Gayrard, Comm. Math. Phys. **235**, 379 (2003).
[16] G. Ben Arous, A. Bovier, and V. Gayrard, Comm. Math. Phys. **236**, 1 (2003).
[17] D. J. Gross and M. Mezard, Nuclear Physics B **240**, 431 (1984).
[18] G. Ben Arous, A. Bovier, and J. Černý, *Universality of the REM for dynamics of mean-field spin glasses*, arXiv:0706.2135.
[19] A. Bovier, *Statistical mechanics of disordered systems* (Cambridge University Press, Cambridge, 2006).
[20] M. Talagrand, *Spin glasses: a challenge for mathematicians* (Springer-Verlag, Berlin, 2003).
[21] G. Ben Arous and V. Gayrard, *Elementary potential theory on the hypercube*, arXiv:math/0611178.
[22] P. Matthews, J. Theoret. Probab. **2**, 129 (1989).
[23] J. Černý and V. Gayrard, *Hitting time of large subsets of the hypercube*, arXiv:math/0611242.
[24] L. A. Shepp, Ann. Math. Statist. **42**, 946 (1971).
[25] D. Slepian, Ann. Math. Statist. **32**, 610 (1961).