Universality of REM-like aging in mean field spin glasses

Gérard Ben Arous[1], **Anton Bovier**[2], **Jiří Černý**[3]

[1]Courant Institute, New York University, 251 Mercer St., New York, NY 10012-1185

[2]Weierstrass-Institut für Angewandte Analysis und Stochastik, Mohrenstrasse 39, 10117 Berlin, Germany, and
Institut für Mathematik, Technische Universität Berlin, Strasse des 17. Juni 136, 10623 Berlin, Germany

[3]ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland

E-mail: `benarous@cims.nyu.edu`, `bovier@wias-berlin.de`, `jiri.cerny@math.ethz.ch`

Abstract. Aging 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, aging occurs precisely as predicted by the REM-like trap model of Bouchaud and Dean. This is the first rigorous result about aging 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

A key concept that has emerged over the last 20 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 e.g. [1, 2, 3, 4] for excellent reviews). This phenomena occurs 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. On the one hand, one may study the *non-activated* regime, where the infinite volume limit at fixed time t is taken first, and *then* one analyzes the ensuing dynamics as t tends to infinity. This *non-activated* regime has been studied intensively for Langevin dynamics of various soft-spin versions of mean field spin glasses [5, 6, 7, 8].

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 crossings in the configuration space [9]. 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. [10, 11]). 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. The extreme situation here corresponds to considering time scales just before the equilibration time. While at these scales the relation to the equilibrium Gibbs distribution is most immediate, in many glassy systems these time scales appear to be beyond experimental or numerical reach.

In this letter we show that the mechanism for aging is universal for a class of Glauber dynamics of p-spin Spin Glasses (with $p \geq 3$), out of equilibrium, in a wide range of time scales. These time scales are exponentially long but still much shorter than the time scales needed to reach equilibrium. Thus this mechanism is essentially a transient one, linked to the exploration of the energy landscape long before the dynamics can feel the ground state. This mechanism has been first established for the simple case of the REM, hence our title.

To capture the features of activated dynamics, early on people introduced effective dynamics where the state space is reduced to the configuration with lowest energy [12, 13]. Bouchaud and others [14, 15, 16, 17, 1] systemized this approach by introducing the notion of "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. Mathematically, trap models are continuous time Markov chains whose state space is a (infinite or growing with some parameter) graph (e.g. $\mathbb{Z}^d$). To the vertices (= traps) of this graph one associates independent random variables whose common distribution is assumed to be heavy tailed, that is their mean is infinite. These variables represent the *mean waiting times* of the Markov chain in the corresponding trap.

Trap models involve the ad-hoc introduction of three major features that ultimately need justification. This is the independence of the waiting times associated to the traps, the heavy-tailed nature of their distribution, and finally the Markov property of the dynamics.

In a series of papers [18, 19, 20, 21] (see [22] for a comprehensive review) a systematic investigation of a variety of trap models was initiated. In this process, it emerged that the "slow-down" of the dynamics appears to be universal for these trap models (except in the exceptional dimension 1), and, more precisely, that it has a scaling limit that can be described using stable Lévy processes. Equivalently it was shown that the classical Dynkin-Lamperti picture for heavy-tailed renewal processes is universally satisfied for these trap models (in dimension larger than 1).

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, and only for a specific choice of the transition rates. In [23, 24, 25] this was done through 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 [20], 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 can 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 we present the first rigorous aging results in a model with *correlated energies*, the p-spin interaction Sherrington-Kirkpatrick (SK) model of spin glasses with $p \geq 3$. Quite surprisingly, the results obtained point again 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 defined as follows. A spin configuration σ is a vertex of the hypercube $\mathcal{S}_N \equiv \{-1,1\}^N$. The Hamiltonian is given by

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

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

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

where $R_N(\sigma,\tau)$ denotes the overlap, $R_N(\sigma,\tau) \equiv N^{-1}\sum_{i=1}^{N}\sigma_i\tau_i$. The (random) Gibbs measure on $\mathcal{S}_N$ is then given by $\mu_{\beta,N}(\sigma) \equiv Z_{\beta,N}^{-1}\mathrm{e}^{-\beta H_N(\sigma)}$. Note that in the limit $p \uparrow \infty$ one recovers the random energy model [26], 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

$$p_N(\sigma,\tau) = N^{-1}\mathrm{e}^{\beta H_N(\sigma)} \tag{3}$$

if σ and τ are related by flipping a single spin, and are zero otherwise. It is easy to see that this dynamics satisfies the detailed balance condition 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 useful to construct this dynamics as a time change of a simple unbiased discrete-time random walk, $Y_N(k)$, $k \in \mathbb{N}$, on $\mathcal{S}_N$ started *out of equilibrium*‡ at some fixed point of $\mathcal{S}_N$, say at $\{1,\dots,1\}$: We define the clock-process by

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

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

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

$S_N(k)$ is the time at which $\sigma_N(t)$ makes the k-th jump.

The REM-like trap model. Let us briefly recall the known results in the dynamics of 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 states of exceptionally low energy and will move "quickly" from one of these configurations to the next. Based on this intuition, Bouchaud et al. [14, 15] (and earlier in a similar way de Dominicis [12] and Koper and Hilhorst [13]) 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 $\mathrm{e}^{\beta\varepsilon_k}$, 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 [15] showed that, for $\beta > 1$,

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

where the function H_α is defined by

$$H_\alpha(w) \equiv \frac{\sin(\pi\alpha)}{\pi} \int_w^\infty \frac{\mathrm{d}x}{(1+x)x^\alpha}. \tag{7}$$

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 [20], the result (6) can be deduced from the stronger claim

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

where $V_\alpha(t)$ is the α-stable subordinator (an α-stable subordinator is an increasing Lévy process that satisfies the scaling relation $V_\alpha(st) = s^{1/\alpha}V_\alpha(t)$. It is uniquely characterized by its Laplace transform which is given by $\mathbb{E}[\mathrm{e}^{-\lambda V_\alpha(t)}] = \exp(-t\lambda^\alpha)$. Stable Lévy processes with $\alpha < 1$ are the natural scaling limits of sums of independent random variables that do not have a finite expectation. For a recent pedagogical exposition, see [28]).

‡ Note that this is different from the situation considered in [27] where the equilibrated dynamics was studied numerically. This makes a comparison of these results with ours difficult.

The REM. In [23, 24, 25] 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 [20] where the point of view put forward in (8) was emphasized. Namely, it was shown that the clock process converges again to the stable subordinator: For every $0 < \varrho < 1$, if $\beta_\varrho \equiv \beta/\sqrt{\varrho} > \beta_c \equiv \sqrt{2\ln 2}$, $\gamma = \beta\sqrt{2\varrho\ln 2}$,

$$\lim_{N\uparrow\infty} \mathrm{e}^{-\gamma N} N^{\frac{\beta_\varrho}{2\beta_c}} S_N(t2^{N\varrho}) = cV_{\beta_c/\beta_\varrho}(t). \tag{9}$$

This implies then a similar aging result as in (6), $\lim_{N\to\infty}\Pi(t\mathrm{e}^{\gamma N}, s\mathrm{e}^{\gamma N}) = H_{\beta_c/\beta_\varrho}(s/t)$, as in the REM-like trap model for the probability $\Pi_N(t,s)$ that $\sigma_N(t)$ does not jump between t and $t+s$.

Note that this result has an interesting interpretation: at a time-scale $\mathrm{e}^{\gamma N}N^{-\frac{\beta_\varrho}{2\beta_c}}$ the process succeeds to make $2^{\varrho N}$ steps, that is it explores a subset of configuration space that corresponds to a "little REM" in volume $n = \varrho N$. At this time scale, the process feels an *effective inverse temperature* β_ϱ. If the effective temperature is below the critical one for the standard REM, the system shows aging, otherwise it does not. It may seem somewhat counter-intuitive that the systems is effectively "warming up" as time goes by.

Let us discuss the heuristics of this result. When the random walk has made $2^{\varrho N}$ steps, with $\varrho < 1$, it has only explored a small fraction of the total configuration space. In particular, it has not had time to find the absolute minima of H_N and hence is still out of equilibrium. Moreover, the random walk will essentially not have visited any configuration twice. Therefore, the minimum of H_N along on those configurations that were visited is the minimum of $2^{\varrho N}$independent Gaussian random variables of mean zero and variance N. It is well know (see e.g. [29]) that this is of the order $-N\sqrt{2\varrho\ln 2}$. Then the mean waiting time in this extreme trap is of order $\exp(\beta N\sqrt{2\varrho\ln 2}) = \mathrm{e}^{\gamma N}$, up to a polynomial correction. Now the condition that $\beta_\varrho > \beta_c$ implies that this time is of the same order as the total time the process has accumulated in all the other sites along its way, and, more precisely, the process will have spent all but a negligible fraction of its time in the "few" "deepest trap". Standard results of extreme value theory imply that the precise statistics of the times spent in the deepest traps are asymptotically governed by a Poisson process, and that the sum of these random times, after rescaling, converges to a stable subordinator, as claimed.

p-spin models. We will now present our new results for the p-spin SK model. The full details of the proofs of these results are given in [30]. The first issue is the choice of the correlation function. Since the valleys in the free-energy landscape contains more than one configuration, the choice $\tilde{\Pi}$ from the REM is no longer appropriate. We set

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

that is we ask for the probability that the distance between the process at times $te^{\gamma N}$ and $(t+s)e^{\gamma N}$ is very small. Then, for $p \geq 3$, a similar result as in the REM holds. Namely, there is a p-dependent value ϱ_p, such that if ϱ and β satisfy the conditions

$$\beta_\varrho \equiv \beta/\sqrt{\varrho} > \sqrt{2\ln 2} \equiv \beta_c, \qquad \text{and} \qquad \varrho < \varrho_p, \tag{11}$$

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

$$\lim_{N\uparrow\infty} \Pi_N^\varepsilon(t,s) = H_{\beta_c/\beta_\varrho}(s/t). \tag{12}$$

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 ϱ_p used in (11) to limit the considered time scales is increasing and it satisfies

$$\varrho_3 \simeq 0.763 \qquad \text{and} \qquad \lim_{p\uparrow\infty} \varrho_p = 1, \tag{13}$$

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

Note the rôle of the two restrictions on β and ϱ in (11). The first one is again the statement that the effective temperature at the time scale considered is below the critical one. The second condition is related to the correlation of the energies. It implies that the REM-like behavior holds only up to time scales where the explored region is still so small that the process does not feel the correlations; essentially it ensures that the process does not have enough time to get close enough to a point it visited before so that it is able to feel the correlations.

The heuristic justification of the results in the p-spin model is rather similar to that of the REM. The difference here is that the energies at the sites that the walk has visited are correlated. Our assertion is that under the condition $\varrho < \varrho_p$, this has only a mild effect and does not change the overall picture. The reason for this relies on the geometric properties of typical trajectories of the random walk on the hypercube, and on the extreme value properties of correlated Gaussian processes. First, it has been shown (see e.g. [31, 32, 33]) that if p is larger than 2 and if K_ϱ is a totally random subset of the hypercube S_N of cardinality $2^{\varrho N}$, with ϱ sufficiently small (depending on p), then the extremal process of H_N restricted to K_ϱ are the same as if $H_N(\sigma)$ were independent random variables. Note that this is not true in the standard SK model with $p = 2$ which is the reason our results can be expected only for $p \geq 3$.

Now it is clear that the trajectories of the random walk cannot look exactly like a totally random set, since after all the trajectory is connected, while in K_ϱ essentially all points are isolated. However, a detailed analysis of the random walk reveals that its trajectories look very much like a random set K_ϱ with linear pieces between them joining the points up in a minimal way. Hence, the correlations have some impact only very locally in time, implying in particular that deep traps will not be made of single points 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 [34, 31]), 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.

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, and more usual, dynamics such as Metropolis or heat bath. Note that in the context of *trap models*, different kinds of dynamic rules have been studied already by Koper and Hilhorst [13], but this is quite a different matter from Glauber dynamics on the space of spin configurations. This problem remains one of the great challenges in the field.

Acknowledgments

Research of A.B. was supported in part by the DFG in the Dutch-German Bilateral Research Group "Mathematics of random spatial models from physics and biology".

References

[1] J.-P. Bouchaud, L. Cugliandolo, J. Kurchan, and M. Mézard. Out of equilibrium dynamics in spin-glasses and other glassy systems. In A. P. Young, editor, *Spin glasses and random fields.* World Scientific, Singapore, 1998.

[2] D. S. Fisher, P. Le Doussal, and C. Monthus. Nonequilibrium dynamics of random field Ising spin chains: exact results via real space renormalization group. *Phys. Rev. E (3)*, 64:066107, 2001.

[3] G. Biroli. A crash course on aging. *J. Stat. Mech.*, P05014, 2005.

[4] F. Iglói and C. Monthus. Strong disorder RG approach of random systems. *Phys. Rep.*, 412:227–431, 2005.

[5] L. F. Cugliandolo and J. Kurchan. Analytical solution of the off-equilibrium dynamics of a long-range spin-glass. *Phys. Rev. Lett.*, 71:173, 1993.

[6] L. F. Cugliandolo and D. S. Dean. Full dynamical solution for a spherical spin-glass model. *J. Phys. A*, 28:4213–4234, 1995.

[7] G. Semerjian, L. F. Cugliandolo, and A. Montanari. On the stochastic dynamics of disordered spin models. *J. Statist. Phys.*, 115:493–530, 2004.

[8] A. Guionnet. Dynamics for spherical models of spin-glass and aging. In *Spin glasses*, volume 1900 of *Lecture Notes in Math.*, pages 117–144. Springer, Berlin, 2007.

[9] M. Goldstein. Viscous liquids and the glass transition: A potential energy barrier picture. *The Journal of Chemical Physics*, 51:3728–3739, 1969.

[10] G. Biroli and J. Kurchan. Metastable states in glassy systems. *Physical Review E*, 64(1):016101, July 2001.

[11] J. F. Fontanari and P. F. Stadler. Fractal geometry of spin-glass models. *J. Phys. A.*, 35(7):1509–1516, February 2002.

[12] C. De Dominicis, H. Orland, and F. Lainée. *J. Phys. (Paris)*, 46:L463, 1985.

[13] G.M.J. Koper and H.J. Hilhorst. *Physica A*, 160:1–23, 1989.

[14] J.-P. Bouchaud. Weak ergodicity breaking and aging in disordered systems. *J. Phys. I (France)*, 2:1705–1713, september 1992.

[15] J.-P. Bouchaud and D. S. Dean. Aging on Parisi's tree. *J. Phys I(France)*, 5:265, 1995.

[16] C. Monthus and J.-P. Bouchaud. Models of traps and glass phenomenology. *J. Phys. A*, 29:3847–3869, 1996.

[17] B. Rinn, P. Maass, and J.-P. Bouchaud. Multiple scaling regimes in simple aging models. *Phys. Rev. Lett*, 84:5403–5406, 2000.

[18] G. Ben Arous and J. Černý. Bouchaud's model exhibits two aging regimes in dimension one. *Ann. Appl. Probab.*, 15(2):1161–1192, 2005.

[19] G. Ben Arous, J. Černý, and Th. Mountford. Aging in two-dimensional Bouchaud's model. *Probab. Theor. Rel. Fields*, 134(1):1–43, 2006.

[20] 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.

[21] G. Ben Arous and J. Černý. Scaling limit for trap models on $\mathbb{Z}^d$. *Ann. Probab.*, 35(6):2356–2384, 2007.

[22] G. Ben Arous and J. Černý. Dynamics of trap models. In *École d'Été de Physique des Houches, Session LXXXIII "Mathematical Statistical Physics"*, pages 331–394. Elsevier, 2006.

[23] G. Ben Arous, A. Bovier, and V. Gayrard. Aging in the random energy model. *Phys. Rev. Letts*, 88(8):087201, February 2002.

[24] G. Ben Arous, A. Bovier, and V. Gayrard. Glauber dynamics of the random energy model. I. Metastable motion on the extreme states. *Comm. Math. Phys.*, 235(3):379–425, 2003.

[25] G. Ben Arous, A. Bovier, and V. Gayrard. Glauber dynamics of the random energy model. II. Aging below the critical temperature. *Comm. Math. Phys.*, 236(1):1–54, 2003.

[26] D. J. Gross and M. Mezard. The simplest spin glass. *Nuclear Physics B*, 240:431–452, 1984.

[27] A. Billoire, L. Giomi, and E. Marinari. *Europhys. Lett.*, 71:824–830, 2005.

[28] A. E. Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications*. Universitext. Springer-Verlag, Berlin, 2006.

[29] M.R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer-Verlag, New York, 1983.

[30] G. Ben Arous, A. Bovier, and J. Černý. Universality of the REM for dynamics of mean-field spin glasses. arXiv:0706.2135.

[31] A. Bovier. *Statistical mechanics of disordered systems*. Cambridge University Press, Cambridge, 2006.

[32] G. Ben Arous, V. Gayrard, and A. Kuptsov. A new REM conjecture. arXiv:math/0612373, 2006.

[33] G. Ben Arous and A. Kuptsov. The limits of REM universality. In preparation, 2008.

[34] M. Talagrand. *Spin glasses: a challenge for mathematicians*. Springer-Verlag, Berlin, 2003.