An introduction to ageing

Lectures given in Beijing

Anton Bovier
Weierstraı-Institut für Angewandte Analysis und Stochastik
Mohrenstraıe 39
10117 Berlin, Germany

Institut für Mathematik
Technische Universität Berlin
Straıe des 17. Juni 136
10623 Berlin, Germany

Contents

1

Introduction

These lectures will focus on some aspects of the dynamics of disordered systems. Our main interest is to understand the long-term behavior of these dynamics in situations when they the systems approaches equilibrium very slowly. In the physics literature, the paradigm of *ageing* has been introduced to characterize this behavior.

1.1 Characterisation of ageing

The term *ageing* refers to properties of a system out of equilibrium. In principle, this property refers to *real (physical) systems*. In the widest sense we can describe it as follows. Assume a systems is prepared (produced) at some initial time t_0. Then the system is left to itself. After some time t_w (called *waiting time*, an experimentalist may perform some measurement on the system. The question is, whether the experimentalist will be able to deduce the elapsed waiting time from his observation. If the answer is yes, we will say that the system ages, otherwise it does not.

Of course, this is a very general characterization and we will be interested in more specific situations. There are a number of clear real-world examples:

- many living beings, such as humans;
- wine....
- steel under stress
- plastics
- glasses
- unmagnetized iron placed in a magnetic field

- a magnetized material when the magnetic field is switched off

and many many more.

In these lectures we will be concerned with mathematical models that correspond to this behavior. Again, one could look at very general dynamical systems, but we will confine our interest exclusively to *Markov processes*.

Let us introduce some notation.

1.2 Trap models

The best studied models for aging are the so-called *trap models*, introduced essentially by Bouchaud and Dean [10, 11]. These models were introduced as caricatures of more realistic models, but they teach us something about how one would like to think about ageing systems. Let us give a general definition.

A trap model has the following ingredients:

(i) A graph, $\mathcal{G} = (\mathcal{E}, \mathcal{V})$; this can either be an infinite graph or a family, $\mathcal{G}_N$, of finite graphs such that $|\mathcal{G}_N| \uparrow \infty$.

(ii) A random environment of traps, i.e. a family of positive random variables, τ_i, $i \in \mathcal{V}$. The usual assumptions in trap models are that these are independent and identically distributed, and, moreover that they are in the domain of attraction of an α-stable distribution with $\alpha < 1$, i.e. $\lim_{t\uparrow\infty} t^\alpha \mathbb{P}[\tau_i > t] = 1$. In particular $\mathbb{E}\tau_i = +\infty$.

(iii) For any realization of the random variables τ_i, a continuous time Markov chain, X_t with state space $\mathcal{V}$ and transition rates

$$p(i,j) = \begin{cases} \tau_i^{1-a}\tau_j^a, & \text{if } (i,j) \in \mathcal{E}, \\ \sum_{k:(i,k)\in\mathcal{E}} \tau_i^{1-a}\tau_k^a, & \text{if } i = j) \\ 0, & \text{, else} \end{cases} \tag{1.1}$$

for some parameter $0 \leq a \leq 1$.

Note that with this choice, the Markov process is reversible with respect to the measure $\mu(i) \equiv \tau_i^{-1}$ on $\mathcal{V}$.

We will in these lectures mainly consider the case $a = 0$, which is the original choice of Bouchaud. In that case the dynamics has a simple description: starting in some site, i, the process waits an exponential time with mean τ_i, and then moves on uniformly to one of its neighbors in the graph $\mathcal{G}$. Now the random variables τ_i, i.e. the trapping times, have a very heavy-tailed distribution, so that as the process wanders about, it can find ever deeper traps, i.e. sites where it will wait longer and longer. So if it is the case that the process, by time T is with large probability in a trap whose waiting time is of order $g(T)$, then we can indeed determine the age of the process by studying its current typical sejourntimes. The nice feature of trap models in that respect is that the state space has site by site a temporal characteristic, a feature that more complicated models do not immediately show.

There has been a considerable amount of work done on trap models in the case when $\mathcal{G} = \mathcal{G}_N$ is the complete graph and when $\mathcal{G} = \mathbb{Z}^d$, mostly by Ben Arous and Černý [7, 2, 3, 3, 4].

We see that we will always be working with Markov process in random environments. It will be convenient to fix some notation once and for all.

Let $(\Omega, \mathcal{F})$ be the measure space on which the random variables τ_i are defined, and let $(\mathcal{D}_0(\mathcal{V}), \mathcal{B}(D_0(\mathcal{V})))$ be the measure space of càdlàgfunctions with values in $\mathcal{V}$. We consider the product space $(\Omega \times \mathcal{D}_0(\mathcal{V}), \mathcal{F} \times \mathcal{B}(D_0(\mathcal{V})))$, where $\mathcal{F} \times \mathcal{B}(D_0(\mathcal{V}))$ is of course the product sigma-algebra. On this space we define a probability measure, $\mathbb{P}$, as follows:

(i) The marginal distribution of the random variables τ_i under $\mathbb{P}$ is the product measure with identical one-dimensional marginals given by the distribution of the τ_i.
(ii) The conditional distribution of $\mathbb{P}$, given $\mathcal{F}$, $P_\tau \equiv \mathbb{P}[\cdot|\mathcal{B}(\mathcal{D}(\mathcal{V}))]$, is the law of the Markov chain described above.

One can easily check that this prescription fixes the joint law of the process X and the random variables τ_i. The measure P_τ is often called the *quenched law*. Rather abusively, the marginal of $\mathbb{P}$ in $(\mathcal{D}_0(\mathcal{V}), \mathcal{B}(D_0(\mathcal{V})))$ is called the annealed law by some authors, but this should be avoided.

When studying trap models, the most commonly used *correlation functions* are

$$R[t_w, t] \equiv \mathbb{P}[X(t_w + t) = X(t_w)], \tag{1.2}$$

respectively its *quenched* version

$$R_\tau[t_w, t] \equiv P_\tau[X(t_w + t) = X(t_w)]. \tag{1.3}$$

Another correlation function is

$$\Pi[t_w, t] \equiv \mathbb{P}[X(t_w + s) = X(t_w), \forall 0 \le s \le t], \tag{1.4}$$

respectively

$$\Pi_\tau[t_w, t] \equiv P_\tau[X(t_w + t) = X(t_w), \forall 0 \le s \le t]. \tag{1.5}$$

One could of course, instead of just asking that $X(s) = X(t_w)$ ask for a milder version, like $\text{dist}\,(X(s), X(t_w))$, for some distance, or one might ask for the distribution of such a distance. However, as again we will see later, it is in the spirit of the trap model to use the strict definition above: for large times, very deep traps are quite isolated, and so the right thing to realize the event is for the process to be in the same deep trap (most of) all the time.

One now speaks about ageing systems, if these functions, as t and t_w become large, do not become independent of t_w.

1.3 Glauber dynamics

Trap models may reproduce ageing behavior, but they are in some sense ad hoc models, that are not motivated by microscopic physical models. In particular, they have two features that seem artificial built in: the independence of the traps and the heavy tails of the distribution of the traps.

Models that are a step closer to reality are Glauber dynamics of (random) spin systems. Here we consider as state space the hypercube $\mathcal{S}_N \equiv \{-1, 1\}^N$ (we could also be more general), and defined on this an energy function (Hamiltonian) $H_N(\sigma)$ which may depend on a random parameter, i.e. may be considered as a random process indexed by $\mathcal{S}_N$. The examples we will be concerned with here are so-called *mean-field* spin glasses, where H_N is a centered Gaussian process with some covariance

$$\text{cov}(H_N(\sigma), H_N(\sigma')) = N f(\,\text{dist}\,_N(\sigma, \sigma')),$$

for some function f such that $f(0) = 1$ and $\text{dist}\,_N$ a normalized distance. The most prominent examples are the p-spin interaction Sherrington-Kirkpatrick models, where

$$\text{cov}(H_N(\sigma), H_N(\sigma')) = N R_N(\sigma, \sigma')^p, \tag{1.6}$$

with $R_N(\sigma, \sigma') \equiv N^{-1} \sum_{i=1}^n \sigma_i \sigma_i'$. Given such a Hamiltonian, one constructs a *Gibbs measure*

$$\mu_{\beta,N}(\sigma) \equiv \frac{2^{-N} \exp(-H_N(\sigma))}{Z_{\beta,N}}, \tag{1.7}$$

where $Z_{\beta,N}$ is such that $\mu_{\beta,N}$ is a probability.

A *Glauber dynamics* is then a (discrete or continuous time) Markov chain that is reversible with respect to this measure. In most cases, one assumes also that only transitions are allowed in which a single spin is flipped at a time. Popular rates are: *Metropolis rates*:

$$p(\sigma, \sigma') = \exp\left(-\beta[H_N(\sigma') - H_N(\sigma)]_+\right), \quad \text{if } |\sigma - \sigma'| = 2 \tag{1.8}$$

and zero else; particularly nice are *random time change rates*:

$$p(\sigma, \sigma') = \exp\left(\beta H_N(\sigma)\right), \quad \text{if } |\sigma - \sigma'| = 2 \tag{1.9}$$

We see that in these dynamics, neither independence nor heavy tails appear. Nonetheless, one expects that under suitable conditions, trap model dynamics emerges as appropriate description of the long time behavior of these models (when $N \uparrow \infty$).

2

Some equilibrium results: The REM

In this chapter I will briefly present some of the ideas surrounding the equilibrium properties of spin-glasses in the simplest case the REM. This will then motivate the REM-like trap model in a heuristic way, and we will later see to what extend this heuristics can be justified.

We will work on the configuration space $\mathcal{S}_N \equiv \{-1,1\}^N$, although for some time that may look quite artificial. The simplest (from a probabilistic point of view) energy landscape we can define are of course iid random variables. Thus we set

$$H_N(\sigma) = -\sqrt{N} X_\sigma \tag{2.1}$$

where X_σ, $\sigma \in \mathcal{S}_N$, are 2^N i.i.d. standard normal random variables.

The first things to look for are the minima of this function, i.e. the maxima of the X_σ. Now for iid random variables, this is a well studied problem in the field of *extreme value statistics*. We begin with a simple observation.

Lemma 2.0.1 *The family of random variables introduced above satisfies*

$$\lim_{N\uparrow\infty} \max_{\sigma\in\mathcal{S}_N} N^{-1/2} X_\sigma = \sqrt{2\ln 2} \tag{2.2}$$

both almost surely and in mean.

Proof Since everything is independent,

$$\mathbb{P}\left[\max_{\sigma\in\mathcal{S}_N} X_\sigma \leq u\right] = \left(1 - \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-x^2/2} dx\right)^{2^N} \tag{2.3}$$

and we just need to know how to estimate the integral appearing here. This is something we should get used to quickly, as it will occur all over the place. It will always be done using the fact that, for $u > 0$,

$$\frac{1}{u}e^{-u^2/2}\left(1-2u^{-2}\right) \leq \int_u^\infty e^{-x^2/2}dx \leq \frac{1}{u}e^{-u^2/2} \tag{2.4}$$

$$\frac{2^N}{\sqrt{2\pi}}\int_{u_N(x)}^\infty e^{-z^2/2}dx = e^{-x} \tag{2.5}$$

then (for $x > -\ln N/\ln 2$)

$$u_N(x) = \sqrt{2N\ln 2} + \frac{x}{\sqrt{2N\ln 2}} - \frac{\ln(N\ln 2)+\ln 4\pi}{2\sqrt{2N\ln 2}} + o(1/\sqrt{N}) \tag{2.6}$$

Thus

$$\mathbb{P}\left[\max_{\sigma\in\mathcal{S}_N} X_\sigma \leq u_N(x)\right] = \left(1-2^{-N}e^{-x}\right)^{2^N} \to e^{-e^{-x}} \tag{2.7}$$

In other terms, the random variable $u_N^{-1}\left(\max_{\sigma\in\mathcal{S}_N} X_\sigma\right)$ converges in distribution to a random variable with double-exponential distribution (known as the Gumbel distribution). □

Next we turn to the analysis of the partition function.

$$Z_{\beta,N} \equiv 2^{-N}\sum_{\sigma\in\mathcal{S}_N} e^{\beta\sqrt{N}X_\sigma} \tag{2.8}$$

This will be important, because we will want to consider processes whose invariant measure is given by the so-called Gibbs measure,

$$\mu_{\beta,N}(\sigma) \equiv Z_{\beta,N}^{-1}e^{-\beta H_N(\sigma)}.$$

A first guess would be that a *law of large numbers* might hold, implying that $Z_{\beta,N} \sim \mathbb{E}Z_{\beta,N}$, and hence

$$\lim_{N\uparrow\infty}\Phi_{\beta,N} = \lim_{N\uparrow\infty}\frac{1}{N}\ln\mathbb{E}Z_{\beta,N} = \frac{\beta^2}{2}, \text{a.s.} \tag{2.9}$$

Holds only for small enough values of β!

Theorem 2.0.2 *In the REM,*

$$\lim_{N\uparrow\infty}\mathbb{E}\Phi_{\beta,N} = \begin{cases} \frac{\beta^2}{2}, & \text{for } \beta\leq\beta_c \\ \frac{\beta_c^2}{2}+(\beta-\beta_c)\beta_c, & \text{for } \beta\geq\beta_c \end{cases} \tag{2.10}$$

where $\beta_c = \sqrt{2\ln 2}$.

Proof We use the method of truncated second moments.

We will first derive an upper bound for $\mathbb{E}\Phi_{\beta,N}$. Note first that by Jensen's inequality, $\mathbb{E}\ln Z \leq \ln\mathbb{E}Z$, and thus

$$\mathbb{E}\Phi_{\beta,N} \leq \frac{\beta^2}{2} \tag{2.11}$$

On the other hand we have that

$$\mathbb{E}\frac{d}{d\beta}\Phi_{\beta,N} = N^{-1/2}\mathbb{E}\frac{\mathbb{E}_\sigma X_\sigma e^{\beta\sqrt{N}X_\sigma}}{Z_{\beta,N}} \le N^{-1/2}\mathbb{E}\max_{\sigma\in\mathcal{S}_N} X_\sigma \le \beta\sqrt{2\ln 2}(1+C/N) \tag{2.12}$$

for some constant C. Combining (2.11) and (2.12), we deduce that

$$\mathbb{E}\Phi_{\beta,N} \le \inf_{\beta_0\ge 0}\begin{cases}\frac{\beta^2}{2}, & \text{for } \beta\le\beta_0\\ \frac{\beta_0^2}{2}+(\beta-\beta_0)\sqrt{2\ln 2}(1+C/N), & \text{for } \beta\ge\beta_0\end{cases} \tag{2.13}$$

It is easy to see that the infimum is realized (ignore the C/N correction) for $\beta_0 = \sqrt{2\ln 2}$. This shows that the right-hand side of (2.10) is an upper bound.

It remains to show the corresponding lower bound. Note that, since $\frac{d^2}{d\beta^2}\Phi_{\beta,N} \ge 0$, the slope of $\Phi_{\beta,N}$ is non-decreasing, so that the theorem will be proven if we can show that $\Phi_{\beta,N} \to \beta^2/2$ for all $\beta < \sqrt{2\ln 2}$, i.e. that the law of large numbers holds up to this value of β. A natural idea to prove this is to estimate the variance of the partition function One would compute

$$\begin{aligned}\mathbb{E}Z_{\beta,N}^2 &= \mathbb{E}_\sigma\mathbb{E}_{\sigma'}\mathbb{E}e^{\beta\sqrt{N}(X_\sigma+X_{\sigma'})}\\ &= 2^{-2N}\left(\sum_{\sigma\neq\sigma'} e^{N\beta^2} + \sum_\sigma e^{2N\beta^2}\right)\\ &= e^{N\beta^2}\left[(1-2^{-N}) + 2^{-N}e^{N\beta^2}\right]\end{aligned} \tag{2.14}$$

where all we used is that for $\sigma\neq\sigma'$ X_σ and $X_{\sigma'}$ are independent. The second term in the square brackets is exponentially small if and only if $\beta^2 < \ln 2$. For such values of β we have that

$$\begin{aligned}\mathbb{P}\left[\left|\ln\frac{Z_{\beta,N}}{\mathbb{E}Z_{\beta,N}}\right| > \epsilon N\right] &= \mathbb{P}\left[\frac{Z_{\beta,N}}{\mathbb{E}Z_{\beta,N}} < e^{-\epsilon N} \text{ or } \frac{Z_{\beta,N}}{\mathbb{E}Z_{\beta,N}} > e^{\epsilon N}\right]\\ &\le \mathbb{P}\left[\left(\frac{Z_{\beta,N}}{\mathbb{E}Z_{\beta,N}} - 1\right)^2 > \left(1-e^{-\epsilon N}\right)^2\right]\\ &\le \frac{\mathbb{E}Z_{\beta,N}^2/(\mathbb{E}Z_{\beta,N})^2 - 1}{(1-e^{-\epsilon N})^2}\\ &\le \frac{2^{-N}+2^{-N}e^{N\beta^2}}{(1-e^{-\epsilon N})^2}\end{aligned} \tag{2.15}$$

which is more than enough to get (2.9). But of course this does not correspond to the critical value of β claimed in the proposition!

Instead of the second moment of Z one should compute a truncated version of it, namely, for $c \geq 0$,

$$\widetilde{Z}_{\beta,N}(c) \equiv \mathbb{E}_\sigma e^{\beta\sqrt{N}X_\sigma} \mathbb{1}_{X_\sigma < c\sqrt{N}} \tag{2.16}$$

An elementary computation using (2.4) shows that, if $c > \beta$, then

$$\mathbb{E}\widetilde{Z}_{\beta,N}(c) = e^{\frac{\beta^2 N}{2}} \left(1 - \frac{e^{-N\beta^2/2}}{\sqrt{2\pi N}(c-\beta)}(1 + O(1/N)\right) \tag{2.17}$$

so that such a truncation essentially does not influence the mean partition function. Now compute the mean of the square of the truncated partition function (neglecting irrelevant $O(1/N)$ errors):

$$\mathbb{E}\widetilde{Z}^2_{\beta,N}(c) = (1-2^{-N})[\mathbb{E}\widetilde{Z}_{\beta,N}(c)]^2 + 2^{-N}\mathbb{E}e^{\beta\sqrt{N}2X_\sigma} \mathbb{1}_{X_\sigma < c\sqrt{N}}) \tag{2.18}$$

where

$$\mathbb{E}\, e^{2\beta\sqrt{N}X_\sigma} \mathbb{1}_{X_\sigma < c\sqrt{N}} = \begin{cases} e^{2\beta^2 N}, & \text{if } 2\beta < c \\ 2^{-N} \frac{e^{2c\beta N - \frac{c^2 N}{2}}}{(2\beta - c)\sqrt{2\pi N}}, & \text{otherwise,} \end{cases} \tag{2.19}$$

Combined with (2.17) this implies that, for $c/2 < \beta < c$,

$$\frac{2^{-N}\mathbb{E}\, e^{2\beta\sqrt{N}X_\sigma} \mathbb{1}_{X_\sigma < c\sqrt{N}}}{\left(\mathbb{E}\, \widetilde{Z}_{N,\beta}\right)^2} = \frac{e^{-N(c-\beta)^2 - N(2\ln 2 - c^2)/2}}{(2\beta - c)\sqrt{N}} \tag{2.20}$$

Therefore, for all $c < \sqrt{2\ln 2}$, and all $\beta < c$,

$$\mathbb{E}\left[\frac{\widetilde{Z}_{\beta,N}(c) - \mathbb{E}\widetilde{Z}_{\beta,N}(c)}{\mathbb{E}\widetilde{Z}_{\beta,N}(c)}\right]^2 \leq e^{-Ng(c,\beta)} \tag{2.21}$$

with $g(c,\beta) > 0$. Thus Chebyshev's inequality implies that

$$\mathbb{P}\left[|\widetilde{Z}_{\beta,N}(c) - \mathbb{E}\widetilde{Z}_{\beta,N}(c)| > \delta\mathbb{E}\widetilde{Z}_{\beta,N}(c)\right] \leq \delta^{-2}e^{-Ng(c,\beta)} \tag{2.22}$$

and so, in particular,

$$\lim_{N\uparrow\infty} \frac{1}{N}\mathbb{E}\ln \widetilde{Z}_{\beta,N}(c) = \lim_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{E}\widetilde{Z}_{\beta,N}(c) \tag{2.23}$$

for all $\beta < c < \sqrt{2\ln 2} = \beta_c$. But this implies that for all $\beta < \beta_c$, we can chose c such that

$$\lim_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{E}Z_{\beta,N} \geq \lim_{N\uparrow\infty} \frac{1}{N} \ln \mathbb{E}\widetilde{Z}_{\beta,N}(c) = \frac{\beta^2}{2} \tag{2.24}$$

This proves the theorem. $\square$

Next we will turn to a more refined analysis of what happens if $\beta \geq \sqrt{2\ln 2}$. The key result is the following:

Theorem 2.0.3 *Let $\mathcal{P}$ denotes the Poisson point process on $\mathbb{R}$ with intensity measure $e^{-x}dx$. Then, in the REM, with $\alpha^{-1} = \beta/\sqrt{2\ln 2}$, if $\beta > \sqrt{2\ln 2}$,*

$$e^{-N[\beta\sqrt{2\ln 2}-\ln 2]+\frac{1}{2}\frac{1}{\alpha}[\ln(N\ln 2)+\ln 4\pi]} Z_{\beta,N} \xrightarrow{\mathcal{D}} \int_{-\infty}^{\infty} e^{\alpha z}\mathcal{P}(dz) \tag{2.25}$$

and

$$N\left(\Phi_{\beta,N} - \mathbb{E}\Phi_{\beta,N}\right) \xrightarrow{\mathcal{D}} \ln \int_{-\infty}^{\infty} e^{z/\alpha}\mathcal{P}(dz) - \mathbb{E}\ln \int_{-\infty}^{\infty} e^{z/\alpha}\mathcal{P}(dz). \tag{2.26}$$

Proof Basically, the idea is very simple. We expect that for β large, the partition function will be dominated by the configurations σ corresponding to the largest values of X_σ. Thus we split $Z_{\beta,N}$ carefully into

$$Z_{N,\beta}^{\leq x} \equiv \mathbb{E}_\sigma e^{\beta\sqrt{N}X_\sigma} 1\!\!1_{\{X_\sigma \leq u_N(x)\}} \tag{2.27}$$

and $Z^{>x} \equiv Z_{\beta,N} - Z_{\beta,N}^{\leq x}$. Let us first consider the last summand. It is convenient to rewrite this as (we ignore the sub-leading corrections to $u_N(x)$ and only keep the explicit part of (2.6))

$$\begin{aligned} Z_{\beta,N}^{>x} &= 2^{-N} \sum_{\sigma\in\mathcal{S}_N} e^{\beta\sqrt{N}u_N(u_N^{-1}(X_\sigma))} 1\!\!1_{\{u_N^{-1}(X_\sigma)>x\}} \\ &= e^{N(\beta\sqrt{2\ln 2}-\ln 2)-\frac{1}{2\alpha}[\ln(N\ln 2)+\ln 4\pi]} && (2.28) \\ &\quad \times \sum_{\sigma\in\mathcal{S}_N} e^{\alpha^{-1}u_N^{-1}(X_\sigma)} 1\!\!1_{\{u_N^{-1}(X_\sigma)>x\}} && (2.29) \\ &\equiv \frac{1}{C(\beta,N)} \sum_{\sigma\in\mathcal{S}_N} e^{\alpha^{-1}u_N^{-1}(X_\sigma)} 1\!\!1_{\{u_N^{-1}(X_\sigma)>x\}} && (2.30) \end{aligned}$$

where $C(b,N)$ is defined through the last identity. The key to most of what follows relies on the famous result on the convergence of the extreme value process to a Poisson point process (for a proof see, e.g., [16]):

Theorem 2.0.4 *Let $\mathcal{P}_N$ be point process on $\mathbb{R}$ given by*

$$\mathcal{P}_N \equiv \sum_{\sigma\in\mathcal{S}_N} \delta_{u_N^{-1}(X_\sigma)} \tag{2.31}$$

Then $\mathcal{P}_N$ converges weakly to a Poisson point process on $\mathbb{R}$ with intensity measure $e^{-x}dx$.

Clearly, the weak convergence of $\mathcal{P}_N$ to $\mathcal{P}$ implies convergence in law of the right-hand side of (2.28), provided that $e^{x/\alpha}$ is integrable on $[x,\infty)$ w.r.t. the Poisson point process with intensity e^{-x}. This is, in fact, never a problem: the Poisson point process has almost surely support on a finite set, and therefore $e^{x/\alpha}$ is always a.s. integrable. Note, however, that for $\beta \geq \sqrt{2\ln 2}$ the mean of the integral is infinite, indicating the passage to the low-temperature regime.

Lemma 2.0.5 *Let $Z^{>x}_{\beta,N}, \alpha$ be defined as above, and let $\mathcal{P}$ be the Poisson point process with intensity measure $e^{-z}dz$. Then*

$$C(\beta,N)Z^{>x}_{\beta,N} \stackrel{\mathcal{D}}{\to} \int_x^\infty e^{z/\alpha}\mathcal{P}(dz) \tag{2.32}$$

Next we show that the contribution of the truncated part of the partition function is negligible compared to this contribution. For this it is enough to compute the mean values

$$\begin{aligned}
\mathbb{E}Z^{\leq x}_{\beta,N} &\sim e^{N\beta^2/2} \int_{-\infty}^{u_N(x)-1\beta\sqrt{N}} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \\
&\sim e^{N\beta^2/2} \frac{e^{-(u_N(x)-\beta\sqrt{N})^2/2}}{\sqrt{2\pi}(\beta\sqrt{N}-u_N(x))} \\
&\sim \frac{2^{-N}e^{x(\alpha^{-1}-1)}}{\alpha^{-1}-1} e^{N(\beta\sqrt{2\ln 2}-\ln 2)-\frac{1}{2\alpha}[\ln(N\ln 2)+\ln 4\pi]} \\
&= \frac{e^{x(\alpha^{-1}-1)}}{\alpha^{-1}-1}\frac{1}{C(\beta,N)}
\end{aligned} \tag{2.33}$$

so that

$$C(\beta,N)\mathbb{E}Z^{\leq x}_{\beta,N} \sim \frac{e^{x(\alpha^{-1}-1)}}{\alpha^{-1}-1}$$

which tends to zero as $x \downarrow -\infty$, and so $C(\beta,N)\mathbb{E}Z^{\leq x}_{\beta,N}$ converges to zero in probability. The assertions of Theorem 2.0.3 follow. □

Now note that the right hand side of (2.25) really is a sum, i.e. if we denote by $x_i, i \in \mathbb{N}$ the atoms of the Poisson point process $\mathcal{P}$, then the right hand side of (2.25) can be written as

$$\mathcal{Z}_\alpha \equiv \sum_{i\in\mathbb{N}} e^{x_i/\alpha} \tag{2.34}$$

which can be thought of the partition function of a model with state space $\mathbb{N}$ and Hamiltonian $H(i) = x_i$, where x_i are the atoms of our Poisson process, and temperature α. This model then captures the asymptotics of the random fluctuations of the partition function of the REM. Of course we can associate a Gibbs measure ν_α to this model, via

$$\nu_\alpha(i) \equiv \frac{e^{x_i/\alpha}}{\mathcal{Z}_\alpha}. \tag{2.35}$$

In the next section we will consider a model that effectively can be seen as a Glauber dynamics corresponding to this model. This will be the REM-like trap model of Bouchaud. In Chapter 3 we will than see how the dynamics of the real REM is related to this one.

3

The REM-like trap model

We will now investigate in some detail the simplest of all trap models, called the REM like trap model. The idea is that this should correspond to the dynamics associated to the limiting model associated to the Poisson process that described the asymptotics of the partition function of the REM. Thus, in principle, we would like to define a model with state space $\mathbb{N}$, reversible with respect to the measure ν_α defined in (2.35). It remains to fix the probabilities to go from one state i to another. The natural choice would be the uniform distribution, but of course this makes non sense (immediately). Therefore, one goes through a limiting proceedure: fix an energy $-E$, and consider only the sites i with $x_i > -E$. We know that the number of these sites is Poisson with rate e^E. The values of the corresponding x_i are independent exponential on $(-E, \infty)$. Now condition on the fact that the Poisson variable takes the value, $N \sim e^E$, and shift energies by E, i.e. set $E_i = x_i + E$ (this corresponds to a change of time scale). Then we have N sites with exponential random variables on $[0, \infty)$. It remains to notice that if x is an exponential random variable with mean 1 on $[0, \infty)$, the $e^{x/\alpha}$ has the distribution $\alpha z^{-1-\alpha} dz$.

The resulting model is a trap model with $\mathcal{G} = \mathcal{G}_N$ is the complete graph on N vertices, $a = 0$, and the τ_i have the form $\tau_i = \exp(+E_i/\alpha)$, where E_i are iid exponential random variables. We will assume that the initial distribution at time zero is the uniform distribution.

We will present three ways to analyze this model.

3.1 The renewal theory approach.

We first follow the approach initiated by Bouchaud to study the correlation function $\Pi_N[t_w, t]$ in this simple model. We will always assume that initially, the process starts from the uniform distribution on $\{1, \ldots, N\}$.

Let us begin with the statement of the result.

Proposition 3.1.6 *Define*

$$H_0(w) \equiv \frac{1}{\pi \operatorname{cosec}(\pi\alpha)} \int_w^\infty dx \frac{1}{(1+x)x^\alpha} \tag{3.1}$$

Then, for $\alpha > 0$,

$$\lim_{N\uparrow\infty} \frac{\Pi_N(t_w, t)}{H_0(t_w/t)} = 1, P\text{-}a.s. \tag{3.2}$$

Moreover, the asymptotic behavior of $H_0(w)$ *when* w *tends to zero or* ∞*, respectively, is readily evaluated:*

(i) If $w \downarrow 0$,

$$H_0(w) = 1 - \frac{1}{\pi \operatorname{cosec}(\pi\alpha)} \int_0^w dx \frac{1}{(1+x)x^\alpha} \sim 1 - \frac{w^{1-\alpha}}{(1-\alpha)\pi \operatorname{cosec}(\pi\alpha)} \tag{3.3}$$

(ii) If $w \uparrow \infty$,

$$H_0(w) \sim \frac{1}{\pi \operatorname{cosec}(\pi\alpha)} \int_w^\infty dx \frac{1}{x^{1+\alpha}} = \frac{w^\alpha}{(\alpha)\pi \operatorname{cosec}(\pi\alpha)} \tag{3.4}$$

In the remainder of this subsection we outline the proof of this theorem.

Lemma 3.1.7 *The function* $\Pi_N(t_w, t)$ *satisfies the equations*

$$\Pi_N(t_w, t) = 1 - F_N(t_w + t) + \int_0^t \Pi_N(, t-u) dF_N(u) \tag{3.5}$$

Proof The proof of this lemma is elementary. Just notice that to realize the event defining, the process may either never jump, or it has to make a first jump before at some time $u < t_w$. Since the jump takes the system back to the uniform distribution, we can renew from that time. □

Remember that we study the solution of this equation in the limit when $N \uparrow \infty$. Our first step will be to replace F_N by its limit. Justifying these passages to the limits can be done using concentration of measure techniques, but we will not go into these technical points here. Accepting that this is justified, we get

$$F_\infty(t) \equiv 1 - \alpha \int_1^\infty dx e^{-t/x} x^{-1-\alpha}$$

which is no longer random. Let $\Pi_\infty(s,t)$ denote the unique solution of the equation

$$\Pi_\infty(t_w, t) = 1 - F_\infty(t_w + t) + \int_0^t \Pi_\infty(t_w, t-u) dF_\infty(u) \tag{3.6}$$

Lemma 3.1.8 *For all $t_w, t \geq 0$,*

$$\lim_{N\uparrow\infty} \Pi_N(t_w, t) = \Pi_\infty(t_w, t), \mathbb{P}\text{-}a.s \tag{3.7}$$

The limiting equation (3.6) is solved following standard procedures (see e.g. [15]). One defines the renewal function $M(t)$ that solves the equation

$$M(t) = F_\infty(t) + \int_0^t M(t-u) dF_\infty(u) \tag{3.8}$$

In terms of this function, the solution of (3.6) is then given as

$$\Pi_\infty(t_w, t) = 1 - F_\infty(t_w + t) + \int_0^t (1 - F_\infty(t_w + t - u)) dM(u) \tag{3.9}$$

Setting $f_\infty(t) \equiv F'_\infty(t)$,

$$f_\infty(t) = \alpha \int_1^\infty e^{-t/x} x^{-2-\alpha} dx \tag{3.10}$$

Denote by g^* the Laplace transform of a function g, i.e. $g^*(u) = \int_0^\infty e^{-ut} g(t)$.

$$\begin{aligned} F^*_\infty(u) &= u^{-1} - \alpha \int_1^\infty \frac{dx}{(ux+1)x^\alpha} \\ &= u^{-1} - \alpha u^{-1+\alpha} \int_u^{u\infty} \frac{dx}{(1+x)x^\alpha} \end{aligned} \tag{3.11}$$

In the last expression, the integration is understood to be along a transformed path in the complex plane if u is complex. Performing the change of variable $x = y^{-1} - 1$,

$$\int_u^\infty \frac{dx}{(1+x)x^\alpha} = \int_0^{1/(1+u)} \frac{dy}{(1-y)^\alpha y^{1-\alpha}} = \int u/(1+u)^1 z^{-\alpha} (1-z)^{-\alpha+1}. \tag{3.12}$$

One recognizes the Beta integral

$$\int_0^1 \frac{dy}{(1-y)^{\mu-1}y^{\nu-1}} = \frac{\Gamma(\mu)\Gamma(\nu)}{\Gamma(\mu+\nu)}. \tag{3.13}$$

Thus

$$\int_0^\infty \frac{dx}{(1+x)x^\alpha} = \Gamma(\alpha)\Gamma(1-\alpha) = \frac{\pi}{\sin(\pi\alpha)} = \pi \operatorname{cosec}(\pi\alpha) \tag{3.14}$$

Thus, when $u \to 0$, the integral in (3.11) converges to the constant $\pi \operatorname{cosec}(\pi\alpha)$. Similarly, we have that

$$f_\infty^*(u) = \alpha \int_1^\infty \frac{1}{1+ux} x^{-1-\alpha} dx \tag{3.15}$$

In particular, $f_\infty^*(0) = 1$, and

$$1 - f_\infty^*(u) = \alpha \int_1^\infty \left(1 - \frac{1}{1+ux}\right) x^{-1-\alpha} dx = \alpha u^\alpha \int_u^{u\infty} \frac{dx}{(x+1)x^\alpha} \tag{3.16}$$

Taking the Laplace transform of (3.8) this implies that

$$M^*(u) = \frac{F_\infty^*(u)}{1 - f_\infty^*(u)} = \frac{1}{\alpha u^{1+\alpha} \int_u^{u\infty} \frac{dx}{(1+x)x^\alpha}} - u^{-1} \tag{3.17}$$

and, by classical results on the asymptotics of the inverse Laplace transform (see [Doe], Vol. 2, Section 7), this in turn implies that for $t \uparrow +\infty$,

$$M(t) \sim \frac{t^\alpha}{\pi\alpha\Gamma(\alpha) \operatorname{cosec}(\pi\alpha)} - 1 \tag{3.18}$$

Finally, we can compute the asymptotics of the solution of equation (3.6). Here we will directly make use of the fact that the Laplace transform of $\Pi_\infty(s,t)$ is given explicitly as

$$\Pi_\infty^*(u, t_w) = \frac{\alpha \int_1^\infty e^{-t_w/x} \frac{dx}{(ux+1)x^{1/\alpha}}}{1 - f_\infty^*(u)} \tag{3.19}$$

we have already established the asymptotics of $1 - f_\infty^*(u)$ near $u = 0$. We still need to treat the numerator. It will be convenient to write

$$\begin{aligned} \alpha \int_1^\infty e^{-t_w/x} \frac{dx}{(ux+1)x^\alpha} &= \alpha \int_1^\infty dx \int_{t_w/x}^\infty dv e^{-v} \frac{1}{(ux+1)x^\alpha} \\ &= \alpha \int_0^\infty dv e^{-v} \int_{t_w/v\wedge 1}^\infty dx \frac{1}{(ux+1)x^\alpha} \\ &= \alpha \int_0^\infty dv e^{-v} \int_{t_w/v}^\infty dx \frac{1}{(ux+1)x^\alpha} \\ &\quad -\alpha \int_{t_w}^\infty dv e^{-v} \int_{t_w/v}^1 dx \frac{1}{(ux+1)x^\alpha} \end{aligned} \tag{3.20}$$

Now the first term can be conveniently represented as u^α times an explicit Laplace transform:

$$\alpha \int_0^\infty dv e^{-v} \int_{t_w/v}^\infty dx \frac{1}{(ux+1)x^\alpha} = \alpha u^\alpha \int_0^{\infty/u} dv e^{-uv} \int_{t_w/v}^{u\infty} dx \frac{1}{(x+1)x^\alpha} \tag{3.21}$$

Note that since all integrands vanish at infinity in the right-half plane, $0/u$ and $u\infty$ can be replaced with 0 and ∞, resp., i.e. the integration contours can be deformed to integrations along the real line. We will show that this term is the dominant one.

In fact, combining (3.16) with (3.20) we get from (3.19) that

$$\Pi^*_\infty(u,t_w) = \frac{\int_0^{\infty/u} dv e^{-uv} \int_{t_w/v}^{u\infty} dx \frac{1}{(1+x)x^\alpha}}{\int_u^\infty \frac{dx}{(1+x)x^\alpha}} - \frac{\int_{t_w}^\infty dv e^{-v} \int_{t_w/v}^1 dx \frac{1}{(u+1/x)x^\alpha}}{u^\alpha \int_u^\infty \frac{dx}{(1+x)x^\alpha}} \tag{3.22}$$

Now the integral in the denominator equals

$$\begin{aligned}\int_u^\infty \frac{dx}{(1+x)x^\alpha} &= \int_0^\infty \frac{dx}{(1+x)x^\alpha} - \int_0^u \frac{dx}{(1+x)x^\alpha} \\ &= \pi \operatorname{cosec}(\pi/\alpha) - u^{1-1/\alpha} \sum_{n=0}^\infty (-1)^n \frac{u^n}{n+1-1/\alpha}\end{aligned} \tag{3.23}$$

where the last sum is convergent for $|u| < 1$. Thus the leading singular (at $u=0$) term from the first term in (3.22) is given by

$$\frac{\int_0^\infty dv e^{-uv} \int_{t_w/v}^\infty dx \frac{1}{(1+x)x^\alpha}}{\pi \operatorname{cosec}(\pi\alpha)} \tag{3.24}$$

which obviously is the Laplace transform of the function $H_0(t_w/t)$.

It remains to consider the second term in (3.22). Here the numerator converges to a constant as u tends to zero, in fact, at $u=0$ it equals

$$\int_{t_w}^\infty dv e^{-v} \int_{t_w/v}^1 dx \frac{1}{x^\alpha} = \frac{1}{1-\alpha} \int_{t_w}^\infty dy e^{-y} \left[1 - y^{1-\alpha}\right] \leq const.e^{-t_w} \tag{3.25}$$

Therefore the leading asymptotic of the second term is given by

$$Const.u^{-\alpha}\, e^{-t_w} \tag{3.26}$$

The inverse Laplace transform of the second term has therefore the leading asymptotic behavior

$$H_1(t_w,t) \sim Const.t^{\alpha-1}\, e^{-t_w} \tag{3.27}$$

Note that while the asymptotics in t looks the same as that of the second term of $H_0(t_w/t)$ in the case $t_w/t \downarrow 0$, due to the exponential decay in t_w, this term can be neglected if t_w is large. Thus we have now established the "aging" asymptotics found in Bouchaud.

3.2 The spectral approach

A second way to analyze ageing in this model is via spectral analysis. This looks rather appealing since on may hope that ageing in rather more general situations may be characterized through the spectral properties of the generator of the corresponding Markov chain. This method was developed in a paper with A. Faggionato [14]

Setting $x_i = 1/\tau_i \equiv e^{-E_i/\alpha}$, the infinitesimal generator of the REM-like trap model is easily seen to be given by the following matrix:

$$\mathcal{L}_N \equiv \begin{pmatrix} \frac{(N-1)x_1}{N} & -\frac{x_1}{N} & \dots & -\frac{x_1}{N} \\ -\frac{x_2}{N} & \frac{(N-1)x_2}{N} & \dots & -\frac{x_2}{N} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{x_N}{N} & -\frac{x_N}{N} & \dots & \frac{(N-1)x_N}{N} \end{pmatrix} \tag{3.28}$$

We start by giving a complete description of the eigenvalues and eigenvectors of $\mathcal{L}_N$. Let $\mu = \mu_N$ be the measure on $\mathcal{S}_N$ with $\mu(i) = x_i^{-1} = \tau_i$. Note that $\mathcal{L}_N$ is a symmetric operator on $L^2(\mu)$ and, trivially, $\mathcal{L}_N \mathcal{I} = 0$ where $\mathcal{I}$ is the vector with all entries equal to 1. The following proposition is based on elementary linear algebra:

Proposition 3.2.9 *Let $x_1, x_2, \dots, x_N$ be all distinct. Then, $\mathcal{L}_N$ has N positive simple eigenvalues $0 = \lambda_1 < \lambda_2 \dots < \lambda_N$ such that*

$$\{\lambda_1, \lambda_2, \dots, \lambda_N\} = \{\lambda \in \mathcal{C} : \phi(\lambda) = 0\},$$

where $\phi(\lambda)$ is the meromorphic function

$$\phi(\lambda) \equiv \sum_{j=1}^{N} \frac{\lambda}{x_j - \lambda}, \qquad (\lambda \in \mathcal{C}). \tag{3.29}$$

If the x_i are labeled such that $x_1 < x_2 < \dots < x_N$, then $x_i < \lambda_{i+1} < x_{i+1}$, for $i = 1, \dots, N-1$. Moreover, for any $i = 1, \dots, N$, the vector $\psi^{(i)} \in \mathcal{R}^N$, defined as

$$\psi_j^{(i)} \equiv \frac{x_j}{x_j - \lambda_i}, \qquad \textit{for } j = 1, \dots, N,$$

is an eigenvector of $\mathcal{L}_N$ with eigenvalue λ_i. $\psi^{(1)}, \dots, \psi^{(N)}$ form an orthogonal basis of $L^2(\mu)$.

Since the x_i have a absolutely continuous distribution, we trivially have the

Corollary 3.2.10 *The assertions of Proposition 3.2.9 hold with probability one for all N.*

Proof Let λ be a generic eigenvalue and let us write the corresponding eigenvector, ψ, as $\psi = a(1,\dots,1)^t + w$, where $\sum_{j=1}^N w_j = 0$. Since $(\mathcal{L}_N\psi)_j = x_j w_j$, we have to solve the system

$$x_j w_j = \lambda a + \lambda w_j, \qquad \forall j = 1,\dots,N. \tag{3.30}$$

Since $x_1,\dots,x_N$ are distinct, it must be true that $a \neq 0$ (otherwise we get $\psi = 0$). Without loss of generality, we set $a = 1$. Note that $\lambda \neq x_j$, for $j = 1\dots N$, since otherwise (3.30) would imply that $\lambda = 0 = x_j$. Therefore we get $w_j = \frac{\lambda}{x_j-\lambda}$. Since it must be true that $\sum_{j=1}^N w_j = 0$, we get that λ is an eigenvalue with ψ s.t $\psi_j = \frac{x_j}{x_j-\lambda}$, being the corresponding eigenvector, iff $\phi(\lambda) = 0$. This implies that ϕ has at most N zeros. Since $\phi(0) = 0$, and, for real λ, $\lim_{\lambda\downarrow x_i}\phi(\lambda) = -\infty$, $\lim_{\lambda\uparrow x_i}\phi(\lambda) = \infty$, we get that ϕ has exactly N zeros. From here the assertions of the theorem follow immediately. □

Proposition 3.2.9 has the following simple corollary:

Corollary 3.2.11 *With probability one, the spectral distribution $\sigma_N \equiv \mathrm{Av}_{j=1}^N \delta_{\lambda_j}$ converges weakly to the measure $\alpha x^{\alpha-1}dx$ on $[0,1]$.*

We will now show that Proposition 3.2.9 allows to derive the asymptotics of the autocorrelation functions easily. In fact, it contains far more information on the long time behavior of the systems (see [14]).

Recall that $p_t(i,j)$, the probability to jump from i to j in an interval of time t, can be expressed as $p_t(i,j) = \left(e^{-t\mathcal{L}_N}\right)_{i,j}$. In particular, by writing ν_t for the probability distribution of $Y_N(t)$ and thinking of the Radon derivative $\frac{d\nu_t}{d\mu}$ as column vector,

$$\frac{d\nu_t}{d\mu} = e^{-t\mathcal{L}_N}\frac{d\nu_0}{d\mu},$$

we see that

$$\frac{d\nu_t}{d\mu} = \sum_{k=1}^N \frac{< \frac{d\nu_0}{d\mu}, \psi^{(k)} >}{< \psi^{(k)}, \psi^{(k)} >} e^{-t\lambda_k}\psi^{(k)}. \tag{3.31}$$

The above formulas are true for an arbitrary initial distribution. Taking

ν_0 to be the uniform distribution, by Proposition 3.2.9, we get

$$\frac{d\nu_o}{d\mu} = \sum_{k=1}^{N} \gamma_k \psi^{(k)}, \text{ where } \gamma_k^{-1} \equiv < \psi^{(k)}, \psi^{(k)} > = \sum_{j=1}^{N} \frac{x_j}{(x_j - \lambda_k)^2}.$$

Then, by Proposition 3.2.9 and (3.31),

$$\Pi_N(t, t_w) = \sum_{j=1}^{N} \sum_{k=1}^{N} \frac{\gamma_k e^{-\lambda_k t_w}}{x_j - \lambda_k} e^{-\frac{N-1}{N} x_j t} \tag{3.32}$$

$$\tag{3.33}$$

This formula admits a nice complex integral representation as follows:

Lemma 3.2.12 *Let γ be a positive oriented loop on $\mathcal{C}$ containing in its interior $\lambda_1, \ldots, \lambda_N$. Let g be an holomorphic function on a domain $D \subset \mathcal{C}$ with $\gamma \subset D$. Then, for any $j = 1 \ldots, N$,*

$$\sum_{k=1}^{N} \frac{\gamma_k g(\lambda_k)}{x_j - \lambda_k} = \frac{1}{2\pi i} \int_\gamma \frac{g(\lambda)}{\phi(\lambda)(x_j - \lambda)} d\lambda. \tag{3.34}$$

Proof Let us set $X \equiv \{x_1, \ldots, x_N\}$ and $\Lambda \equiv \{\lambda_1, \lambda_2, \ldots, \lambda_N\}$. Then, $\phi(\lambda)$ is an holomorphic function on $\mathcal{C} \setminus X$, where $\phi'(\lambda) = \sum_{j=1}^{N} \frac{x_j}{(x_j - \lambda)^2}$, and, in particular, $\phi'(\lambda_j) = \gamma_j^{-1}$. Moreover, the function $[\phi(\lambda)(x_j - \lambda)]^{-1}$, a priori defined on $\mathcal{C} \setminus (X \cup \Lambda)$, can be analytically continued to X as a meromorphic function with simple poles only at the points of Λ. Now the conclusion follows from a trivial application of the residue theorem. □

We can obviously use Lemma 3.2.12 to rewrite Equation (3.32) in the form

$$\Pi_N(t, t_w) = \frac{1}{2\pi i} \int_\gamma \frac{e^{-t_w \lambda}}{\lambda} \left(\mathrm{Av}_j \frac{e^{-\frac{N-1}{N} x_j t}}{x_j - \lambda} \Big/ \mathrm{Av}_j \frac{1}{x_j - \lambda} \right) d\lambda \tag{3.35}$$

where Av_j denotes the average over $j = 1, 2, \ldots, N$.

The above integral representation of $\Pi_N(t, t_w)$ has two advantages. First, the appearance of averages allows to compute their limiting behavior as $N \uparrow \infty$ easily by using the ergodicity of the random field $\underline{E}$. Second, by means of the residue theorem, their Laplace transform can be easily computed in order to derive the asymptotic behavior of $\Pi_N(t, t_w)$ for $N, t_w, t \gg 1$.

The next step is now to show that the contour integral representation converges to a nice limiting expression as $N \uparrow \infty$.

Proposition 3.2.13 *Let us define*

$$\Pi(t,t_w) \equiv \frac{1}{2\pi i}\int_\gamma \frac{e^{-t_w\lambda}}{\lambda}\frac{\mathcal{E}_x\left(\frac{e^{-xt}}{\lambda-x}\right)}{\mathcal{E}_x\left(\frac{1}{\lambda-x}\right)}d\lambda, \tag{3.36}$$

where $\mathcal{E}_x$ is the expectation w.r.t. the measure $\alpha x^{\alpha-1}dx$ on $[0,1]$ and γ is any positive oriented complex loop around the interval $[0,1]$. Then,

$$\lim_{N\uparrow\infty}\Pi_N(t,t_w) = \Pi(t,t_w) \qquad \forall t,t_w, \qquad a.s.. \tag{3.37}$$

Proof Recall (3.35) and fix $0<\delta<1/2$. Due to analyticity, we can choose the integration contour, γ, to have distance 1 from the segment $[0,1]$. For each $\lambda\in\gamma$, the random variables $(x_j-\lambda)^{-1}$, $j\in\mathcal{N}$, are i.d.d. and bounded. Therefore, for a suitable positive constant $c>0$,

$$\mathcal{P}\left(\left|\mathrm{Av}_{j=1}^N\frac{1}{x_j-\lambda}-\mathcal{E}_x\left(\frac{1}{\lambda-x}\right)\right|\geq N^{-\frac{1}{2}+\delta}\right)\leq e^{-c\,N^{2\delta}} \qquad \forall\lambda\in\gamma. \tag{3.38}$$

Since for each $x\in[0,1]$ and $\lambda\in\gamma$, $|\frac{\partial}{\partial\lambda}(x-\lambda)^{-1}|\leq 1$, a simple chaining argument allows to deduce from the pointwise estimate (3.38) uniform control in λ. Using the Borel–Cantelli lemma, one can then infer that, a.s.,

$$\sup_{\lambda\in\gamma}\left|\mathrm{Av}_{j=1}^N\frac{1}{x_j-\lambda}-\mathcal{E}_x\left(\frac{1}{\lambda-x}\right)\right|\leq c\,N^{-\frac{1}{2}+\delta}, \qquad \forall N\in\mathcal{N}. \tag{3.39}$$

Similar arguments show that, a.s., given $M\in\mathcal{N}$, there exists a constant, c_M, such that

$$\sup_{M-1\leq t\leq M}\sup_{\lambda\in\gamma}\left|\mathrm{Av}_{j=1}^N\frac{e^{-\frac{N-1}{N}x_jt}}{x_j-\lambda}-\mathcal{E}_x\left(\frac{e^{-x_jt}}{\lambda-x_j}\right)\right|\leq c_M\,N^{-\frac{1}{2}+\delta}, \qquad \forall N\in\mathcal{N}. \tag{3.40}$$

Note that, for each $\lambda\in\gamma$, $\mathrm{Av}_{j=1}^N(x_j-\lambda)^{-1}$ is a convex combination of points of modulus larger or equal than $1/2$, contained in a angular sector with angle non larger than a suitable constant, $c<\pi$. In particular, $\left|\mathrm{Av}_{j=1}^N(x_j-\lambda)^{-1}\right|\geq c'>0$, for all N. From here the assertion of the proposition follows from Lebesgue's dominated convergence theorem. □

It then remains to analyze the complex integrals in the expression for $\Pi(t,t_w)$. This again uses Laplace transforms and is rather standard. I will omit the details. The result, in any case, is the same we have obtained from the renewal approach.

This concludes the second proof. In itself this may not look easier or more instructive, but the added value arises from the fact that many other results can be obtained in the same way. For details, see the paper [14].

3.3 Subordinators

We now come to the last, and maybe most instructive way to prove Proposition 3.1.6. To do this we give present a slightly different way of constructing the process $X_N(t)$. We begin by describing the *trajectories* of our process *disregarding time*. This will be given by discrete time Markov chain, Y_k, $k \in \mathbb{N}$, taking values in $\{1, \dots, N\}$. In our case this is a very trivial process: $Y(k)$ are iid uniform random variables.

Next we construct the *clock process* $S_N(k)$,

$$S_N(k) = \sum_{i=0}^{k-1} e_i \tau_{Y(i)}, \tag{3.41}$$

where e_i are iid exponential r.v.'s with parameter 1. Note that $S_N(k)$ represents the total time the process spends in order to make k steps. Then $X(T)$ is simply constructed as

$$X_N(t) = Y(S_N^{-1}(t)), \tag{3.42}$$

where the right-continuous inverse of an increasing function, ϕ, is defined as

$$\phi^{-1}(t) = \inf \{(u : \phi(u) \geq t\} . \tag{3.43}$$

We are now interested in studying the limit of the clock process as first N and then k go to infinity. More precisely, we are after a result of the form

$$\lim_{n\uparrow\infty} n^{-1/\alpha} S_N([\theta n]) = V_\alpha(\theta), \tag{3.44}$$

with convergence in the sense of weak convergence for the process (indexed by θ) in s suitable topology. We will see that naturally, the limit will be identified with an *α-stable subordinator*.

Stable subordinators. Let us recall some standard terminology and facts. First, a *subordinator* is just an non-decreasing process. A *Lévy-process* is continuous time stochastic process that

(a) has càdlàg paths,
(b) has independent and stationary increments.

A Lévy process, V_α, is called *α-stable*, if for any $t, s \in \mathbb{R}_t$, $V_\alpha(t)$ and $s^{-1/a} V_\alpha(ts)$ have the same law. Recall the special case $\alpha = 2$ which is Brownian motion.

An non-decreasing α-stable Lévy process is called an α-stable subordinator.

The importance of α-stable Lévy processes is that they are the natural candidates for limit of sums of independent random variables.

Due to the assumptions of stationarity and independence of the increments, a Lévy process is fully characterized by the one-dimensional distribution. In the case of the subordinator, the latter is characterized by its Laplace transform.

Theorem 3.3.14 *For each $b \in \mathbb{R}$ and each measure, ν, on $\mathbb{R}\backslash\{0\}$, that satisfies*

$$\int \min(|x|, 1)\nu(dx) < \infty,$$

the function

$$\phi(\theta) \equiv \exp(-\psi(\theta)),$$

where

$$\psi(\theta) \equiv b\theta + \int (e^{-(\theta x)} - 1)\nu(dx), \tag{3.45}$$

is the Laplace transform of a Lévy subordinator. Moreover, the Laplace transform of any Lévy subordinator can be written in this form with uniquely determined (b, ν). The subordinator is stable with index $\alpha \in (0, 1)$, if, for some $K \in \mathbb{R}_+$,

$$\nu(dx) = \frac{K\alpha}{\Gamma(1-\alpha)} x^{-1-\alpha} dx, \tag{3.46}$$

and hence $\psi(\theta) = K\theta^\alpha$.

Since we will more or less see an explicit construction in the sequel, I will not dwell more on generalities about Lévy process.

In any case, if the random variables $\tau_{Y(i)}$, $i \in \mathbb{N}$, were all independent, then $S_N(k)$ would be a sum of independent random variables, and we could simply invoke classic convergence results for iid random variables. Since the random variables $e_i \tau_{Y(i)}$ are positive and in the domain of attraction of an α-stable law, this would give us immediately that (3.44) would hold with V_α an α-stable subordinator.

Now of course the random variables $\tau_{Y(i)}$ are not all independent, since there are only N of them. Thus, a result like (3.44) can only hold in the limit a $N \uparrow \infty$. Indeed we will show that

Theorem 3.3.15 *Let $\alpha > 1$. Then*

$$\lim_{n\uparrow\infty} n^{-\alpha} \lim_{N\uparrow\infty} S_N([\theta n]) = V_\alpha(\theta), \tag{3.47}$$

in distribution in the Skorokhod J_1-topology.

Proof We will first show that the finite dimensional marginals of $n^{-\alpha}S_N([\theta n])$ have the right limits. Let us first consider the set $\mathcal{A}_{n,N} \equiv \{Y : \exists_{0,k<\ell\leq n} : Y(k) = Y(\ell)\}$. We clearly have the estimate

$$\mathbb{P}[A_{n,N}] \leq n^2 N^{-1} \tag{3.48}$$

Thus $\lim_{N\uparrow\infty} S_N([\theta n])$ has the same distribution as

$$\widetilde{S}(\theta n) \equiv \sum_{i=0}^{[\theta n]-1} e_i \tau_i. \tag{3.49}$$

We will now study the convergence of $\widetilde{S}(n\theta)$ as $n \uparrow \infty$. In the process we will construct and study the α-stable subordinator. We shall see that this is closely linked to extreme value theory and Poisson processes.

Let us start by noting that

$$n\mathbb{P}[e_i\tau_i > n^{1/\alpha}c] \to \Gamma(1-\alpha)c^{-\alpha}. \tag{3.50}$$

Now, a standard result from extreme value theory states the following [16]:

Theorem 3.3.16 *Assume that X_i are iid random variables that satisfy*

$$\lim_{n\uparrow\infty} \mathbb{P}[X_i > u_n(c)] = \nu(c). \tag{3.51}$$

where ν is an increasing (respectively decreasing) function. Then, the point process

$$\sum_{i=0}^{n-1} \delta_{(i/n, u_n^{-1}(X_i))} \tag{3.52}$$

converges in distribution to the Poisson point process, $\mathcal{R}$ on $\mathbb{R}_+ \times \mathbb{R}$ with intensity measure $dt \times d\nu(x)$ (respectively $-d\nu$ if ν is decreasing).

Applying this theorem to our case gives us that

Corollary 3.3.17 *The point process*

$$R_n \equiv \sum_{i=0}^{n-1} \delta_{(i/n, n^{-1/\alpha}e_i\tau_i)} \to \mathcal{R} \tag{3.53}$$

converges to the Poisson point process on $(0,1] \times \mathbb{R}_+$ with intensity measure $dt \times \Gamma(1-\alpha)c^{-1-\alpha}dc$.

We will now prof two facts. First, we will show that form the process $\mathcal{R}$ we can construct an α-stable subordinator. Then we will show that our process $\widetilde{S}(\theta n)$ converges to that process.

Proposition 3.3.18 *Let $\mathcal{R}$ be the Poisson process from Theorem 3.3.16 and assume that the measure $d\nu$ has support on $\mathcal{R}$ is such that $\int(x \wedge 1)d\nu(x) < \infty$. Then the integral*

$$V(t) \equiv \int_0^t \int_0^\infty \mathcal{R}(ds, dx)x \tag{3.54}$$

exists and is an increasing process with independent increments, i.e. a Lévy subordinator.

Proof Let us decompose $V(t)$ into two pieces, $V^>(t)$ and $V^<(t)$, where

$$V^>(t) \equiv \int_0^t \int_1^\infty \mathcal{R}(ds, dx)x, \tag{3.55}$$

and

$$V^<(t) \equiv \int_0^t \int_0^1 \mathcal{R}(ds, dx)x. \tag{3.56}$$

Notice that the two processes, if they exist are independent. Moreover, the intensity measure of the set $[0, t] \times [1, \infty)$ is finite by hypothesis. Thus there are only finitely many points of $\mathcal{R}$ on this set, hence $V^>(t)$ is a finite sum and thus almost surely finite. On the other hand, $v^<(t)$ is positive and

$$\mathbb{E}V^<(t) = t \int_0^1 x d\nu(x) < \infty \tag{3.57}$$

also by hypothesis. Thus also $V^<(t)$ is almost surely finite and hence $V(t)$ is almost surely finite and well defined. Since we really should think of it as the distribution function of the measure $\int \mathcal{R}(\cdot, dx)$ on $\mathbb{R}_+$, is is also right-continuous. Since it has independent increments, it satisfies all hypothesis of a Lévy process. Since it is increasing, it is a subordinator. □

For later use we note that it is quite simple to compute the Laplace transform of this process. In fact, Let $V^{(c)}(t)$ be the truncated version

$$V^{(c)}(t) \equiv \int_0^t \int_c^\infty \mathcal{R}(ds, dx)x. \tag{3.58}$$

Set $M \equiv \int_0^t \int_c^\infty ds d\nu(x)$. Then

$$\mathbb{E}e^{-\lambda V^{(c)}(t)} = \sum_{k=0}^{\infty} \frac{M^k}{k!} e^{-M} \left[\frac{t \int_c^{\infty} e^{-\lambda x} d\nu(x)}{M} \right]^k \tag{3.59}$$
$$= e^{-M} \exp\left(t \int_c^{\infty} e^{-\lambda x} d\nu(x) \right)$$
$$= \exp\left(t \int_c^{\infty} \left(e^{-\lambda x} - 1 \right) d\nu(x) \right).$$

Again by our assumptions on ν, the limit $c \downarrow 0$ exists and yields the Laplace transform of the Lévy subordinator in the standard form

$$\mathbb{E}e^{-\lambda V(t)} = \exp\left(t \int_0^{\infty} \left(e^{-\lambda x} - 1 \right) d\nu(x) \right). \tag{3.60}$$

If for $d\nu(x)$ we take a measure $Kx^{-1-\alpha}dx$, then it $\alpha < 1$ the integrability conditions are satisfied and the resulting process is a stable subordinator (with zero drift). The measure ν is called the Lévy measure. This is all for the moment we want to know about Lévy subordinators.

Now let us turn to the proof of the fact that our process $\widetilde{S}$ converges to such an object.

Theorem 3.3.19 *Let $\alpha < 1$. Then*

$$\lim_{n \uparrow \infty} n^{-1/\alpha} \widetilde{S}(\theta n) = V_\alpha(t), \tag{3.61}$$

where V_α is the stable subordinator with zero drift and Lévy measure $\Gamma(1-\alpha)x^{-1-\alpha}dx$. Convergence is in distribution on the Skhorohod space of càdlàgfunctions equipped with the J_1-topology.

Proof We could proof this in two ways: either compute the Laplace transform, or as follows. Set, for fixed $c > 0$,

$$n^{-1/\alpha} \widetilde{S}(\theta n) = n^{-1/\alpha} \sum_{i=0}^{[\theta n]-1} \mathbb{1}_{e_i \tau_i > c n^{1/\alpha}} e_i \tau_i \tag{3.62}$$
$$+ n^{-1/\alpha} \sum_{i=0}^{[\theta n]-1} \mathbb{1}_{e_i \tau_i \leq c n^{1/\alpha}} e_i \tau_i$$
$$\equiv \widetilde{S}_n(\theta) \equiv \widetilde{S}_n^{>}(\theta) + \widetilde{S}_n^{<}(\theta)$$

Now

$$\mathbb{E}\widetilde{S}_n^{<}(\theta)+ = \theta n^{1-1/\alpha}\alpha \int_0^\infty e^{-z} dz \int_1^{cn^{-1/\alpha}/z} x^{-\alpha} dx \tag{3.63}$$
$$= \frac{\theta\alpha}{1-\alpha}\left(c^{1-\alpha}/\Gamma(\alpha) - n^{1-1/\alpha}\right) \sim \theta c^{1-\alpha},$$

which tends to zero as $c \downarrow 0$. On the other hand, (I) is a function of the point process R_n:

$$\widetilde{S}_n^{>}(\theta) = \int_0^\theta \int_c^\infty R_n(ds, dx)x. \tag{3.64}$$

This converges for any positive c, as $n \uparrow \infty$, to $\int_0^\theta \int_c^\infty \mathcal{R}(ds, dx)x$, and finally, as we have seen, also as $c \downarrow 0$. Since in this limit $\widetilde{S}_n^{<}(\theta)$ tends to zero, we have proven the assertion.

The advantage of the prove is that it gives the convergence in a strong topology, the so-called J_1-topology. The J_1-topology is the topology given by the J_1-metric: for $f, g \in D$

$$d_{J_1}(f, g) = \inf_{\lambda\in\Lambda} \{\|f \circ \lambda - g\|_\infty \vee \|\lambda - e\|_\infty\}, \tag{3.65}$$

where Λ is the set of strictly increasing functions mapping $[0, T]$ onto itself such that both λ and its inverse are continuous, and e is the identity map on $[0, T]$.

We will need a criterion for tightness of probability measures on D. To this end we define several moduli of continuity,

$$w_f(\delta) = \sup\left\{ \min\left(|f(t) - f(t_1)|, |f(t_2) - f(t)|\right) : t_1 \le t \le t_2 \le T, t_2 - t_1 \le \delta\right\},$$
$$v_f(t, \delta) = \sup\left\{|f(t_1) - f(t_2)| : t_1, t_2 \in [0, T] \cup (t - \delta, t + \delta)\right\}. \tag{3.66}$$

The following result is a restatement of Theorem 12.12.3 of [17] and Theorem 15.3 of [9].

Theorem 3.3.20 *The sequence of probability measures $\{P_n\}$ is tight in the J_1-topology if*

(i) For each positive ε there exist c such that

$$P_n[f : \|f\|_\infty > c] \le \varepsilon, \qquad n \ge 1. \tag{3.67}$$

(ii) For each $\varepsilon > 0$ and $\eta > 0$, there exist a δ, $0 < \delta < T$, and an integer n_0 such that

$$P_n[f : w_f(\delta) \ge \eta] \le \varepsilon, \qquad n \ge n_0, \tag{3.68}$$

and

$$P_n[f : v_f(0,\delta) \geq \eta] \leq \varepsilon \text{ and } P_n[f : v_f(T,\delta) \geq \eta] \leq \varepsilon, \qquad n \geq n_0. \tag{3.69}$$

□

Let us check that these criteria are verified in our case. (i) is already checked. Condition (iii) amounts to checking that there is no jump at 0 and at T. In fact, using that all our processes are increasing,

$$\begin{aligned}
\mathbb{P}\left[v_{\widetilde{S}_n}(0,\delta) > \eta\right] &= \mathbb{P}\left[\widetilde{S}_n(\delta) > \eta\right] \\
&\leq \mathbb{P}\left[\widetilde{S}_n^{<}(\delta) > \eta/2\right] + \mathbb{P}\left[\widetilde{S}_n^{>}(\delta) > \eta/2\right] \\
&\leq 2\mathbb{E}\widetilde{S}_n^{<}(\delta)/\eta + \delta n \mathbb{P}\left[e_i\tau_i > cn^{-1/\alpha}\right] \\
&\leq 2\delta c^{1-\alpha}/\eta + \Gamma(1-\alpha)\delta c^{-\alpha}.
\end{aligned} \tag{3.70}$$

Clearly, for any $\eta > 0$ and $\epsilon > 0$, the right-hand side of (3.70) can be made smaller than ϵ by an appropriate choice of δ and c.

The task to check (ii) is not much harder. We may check this condition again for $\widetilde{S}_n^{<}$ and $\widetilde{S}_n^{>}$ separately. For the former, we need a second moment estimate,

$$\mathbb{E}\widetilde{S}_n^{<}(\delta)^2 \leq \text{const.}\delta^2 c^{2-\alpha},$$

and then a standard partitioning argument tells us that

$$\begin{aligned}
\mathbb{P}[w_{\widetilde{S}_n^{<}}(\delta) > \eta/2] &\leq \mathbb{P}\left[\exists_{k \leq T/\delta} : \widetilde{S}_n^{<}((k+1)\delta) - \widetilde{S}_n^{<}(k\delta) \geq \eta/2\right] \\
&\leq \text{const.} T\delta c^{2-\alpha}/\eta.
\end{aligned} \tag{3.71}$$

For $\widetilde{S}_n^{>}0$, the event $\{w_{\widetilde{S}_n^{<}}(\delta) > \eta/2\}$ can only occur if two atoms of R_n have distance smaller than 2δ. The probability of this to happen is controlled by

$$T\delta n^2 \mathbb{P}[e_i\tau_i > n^{1/\alpha}c]^2 \leq T\delta\Gamma(1-\alpha)^2 c^{-2\alpha}. \tag{3.72}$$

Again, both (3.71) and (3.72) can be made smaller than ϵ by suitable choice of c and δ, no matter what η is. This proves the theorem. □

The main advantage of having convergence in the J_1-topology is that it ensures convergence of the jumps: If the limiting subordinator has a jump of given size, then the approximants had jumps converging to the same size, and it cannot be the case that there were many small jumps of the approximants that merged together to produce that of the limit.

But the jumps of the clock process S_N are closely linked to the correlation function. Indeed, if This implies in particular that $X_N(s)$ remains constant on the interval $t_w, t_w + t$, if and only if the clock process jumps over this interval, i.e. if (t_w, t_w+t) in not in the range of S_N. Combining these observations, we get the following fact:

Lemma 3.3.21 *The correlation function* Π_N *satisfies*

$$\lim_{t_w\uparrow\infty}\lim_{N\uparrow\infty}\Pi_N(t_w,\theta t_w)=\mathbb{P}\left[(1,1+\theta)\notin \mathit{range}(V_\alpha)\right]. \tag{3.73}$$

Of course, the probability on the right is well-known (see e.g. the book by Bertoin [8]) and given by the expression that we already know.

4

From the REM to the REM-like trap model

One of the central questions is of course how simple model that exhibit aging can be derived from the more realistic models. The first step in this direction is to see how the REM-like trap model can be understood as a simplification of its namesake, the "real REM".

4.1 Dynamics of the REM

Recall that the Random Energy model is defined by a assigning to each vertex of the hypercube, $\sigma \in S_M \equiv \{-1,1\}^N$ an energy

$$H_N(\sigma) \equiv \sqrt{N} X_\sigma, \tag{4.1}$$

where X_σ, $\sigma \in \mathcal{S}_N$, are iid standard Gaussian random variables. Glauber dynamics of this model is then a continuous time Markov process on $\mathcal{S}_N$ whose rates, $p_N(\sigma,\sigma')$ are non-vanishing when σ and σ' differ in at most one coordinate and that are reversible with respect to the measures $e^{-\beta H_N(\sigma)}$. A popular choice for such rates is the so-called *Metropolis algorithm*, where

$$p_N(\sigma,\sigma') = \exp\left(-\beta[H_N(\sigma') - H_N(\sigma)]_+\right),$$

if σ' is obtained from σ by "flipping" one spin. We are not able to treat this choice (yet). Instead, we will consider another choice for transition rates that is sometimes called *random time change dynamics* (RTCD).

It will be very convenient for us to define this dynamics as follows. We denote by $Y_N(k) \in \mathcal{S}_N$, $k \in \mathbb{N}$, the simple unbiased random walk (SRW) on $\mathcal{S}_N$ started at some fixed point of $\mathcal{S}_N$, say at $\{1,\dots,1\}$. For $\beta > 0$ we define the clock-process by

$$S_N(k) = \sum_{i=0}^{k-1} e_i \exp\left\{\beta\sqrt{N} H_N\big(Y_N(i)\big)\right\}, \tag{4.2}$$

where $\{e_i, i \in \mathbb{N}\}$ is a sequence of mean-one i.i.d. exponential random variables. We denote by $\mathcal{Y}$ the σ-algebra generated by the SRW random variables $Y_N(k), k \in \mathbb{N}, N \in \mathbb{N}$. The σ-algebra generated by the random variables $e_i, i \in \mathbb{N}$ will be denoted by $\mathcal{E}$. Then the process

$$\sigma_N(t) \equiv Y_N(S_N^{-1}(t)) \tag{4.3}$$

is a continuous time Markov process on $\mathcal{S}_N$ that is reversible with respect to the measure $\mu_{\beta,N}$; its generator is given by

$$L_N(\sigma,\tau) = \begin{cases} N^{-1}e^{-\beta\sqrt{N}H_N(\sigma)}, & \text{if } \operatorname{dist}(\sigma,\tau)=1, \\ -e^{-\beta\sqrt{N}H_N(\sigma)}, & \text{if } \sigma=\tau, \\ 0, & \text{otherwise;} \end{cases} \tag{4.4}$$

here $\operatorname{dist}(\cdot,\cdot)$ is the graph distance on the hypercube,

$$\operatorname{dist}(\sigma,\tau) = \frac{1}{2}\sum_{i=1}^{N} |\sigma_i - \tau_i| \tag{4.5}$$

One important point now is that we must choose *time-scales* when studying this dynamics. In the trap model we took the limit $N \uparrow \infty$ first and then let time go to infinity. We could also have chosen diagonal limits (see [14, 3]), but here this will be quite more relevant and interesting.

4.2 Random walk on the extremes

The first results on the REM were obtained in two papers [5, 6]. The idea in these papers was as follows. It is well known (see e.g. [13]) that the equilibrium measure of the REM at temperatures below the critical temperature is concentrated on a essentially finite set of individual spin configurations, whose renormalized energies are well approximated by a Poisson point process with intensity measure $e^{-E}dE$. This means in particular that the process spends eventually all of its time in these favored configurations. Thus the dynamics of the process at sufficiently large times should be described by a process on such a set of states, called the "top",

$$T_E \equiv \{\sigma \in \mathcal{S}_N : -H_N(\sigma) \geq \sqrt{N}u_N(-E)\}, \tag{4.6}$$

where

$$u_N(x) \equiv \sqrt{2N\ln 2} + \frac{x}{\sqrt{2N\ln 2}} - \frac{1}{2}\frac{\ln(N\ln 2) + \ln 4\pi}{\sqrt{2N\ln 2}} \tag{4.7}$$

The natural candidate for this dynamics would then be the REM-like trap model with $N = |T_E|$. Why? First, the waiting time in a extremal state of the REM has mean

$$\hat{\tau}_i = C_{\beta,N} e^{\beta/\sqrt{2\ln 2} E_i}$$

where E_i are exponential random variables of mean 1. This means that the rescaled waiting times

$$\tau_i = \hat{\tau}_i / C(\beta, N)$$

are distributed according to the law $\alpha x^{-1-\alpha} dx$ where $\alpha \equiv \sqrt{2\ln 2}/\beta$. To complete the picture, one would need to prove two more things:

(i) The process jumps from one state in T_E to another with uniform probability.
(ii) The times between consecutive visits of different states in T_E are asymptotically exponentially distributed.

In [5, 6] these things were proven, using essentially a very elaborate version of the renewal approach. This was technically very elaborate. Here we will not go at all into the details, but rather present a new approach based on the study of the clock process.

4.3 Extremes on the random walk

Following the discussion of subordinators in the REM-like trap model, we see that there should be way to study aging that is completely orthogonal to that of [5, 6]. Instead of studying the process on the extreme states, we should study the extreme states on the random walk Y_N! Once we adopt this point of view, thing fall nicely into place.

Let us begin with a heuristic explanation of what is going on.

We first fix a scale, K (that later should depend on the volume, N and the inverse temperature temperature, β. We are interested in the time the process takes to make a number of steps of order K. Thus, we consider the trajectory of the radom walk Y_N of length K.

We are interested in the distribution of the elapsed time along the trajectory, i.e. the rescaled clock process

$$S_N(sK) \equiv \sum_{i=0}^{[sK]-1} e_i e^{-\beta H_N(Y_i)}.$$

Aging should be expected when this time is distributed in a very irregular way on the trajectory, notably when after rescaling the clock converges to a stable subordinator.

Let us now introduce the Gaussian process

$$Z_N(i) \equiv X_{Y_N(i)}, \tag{4.8}$$

indexed by $i \in \mathbb{N}$. Clearly it has the covariance

$$\mathrm{cov}(Z_N(i), Z_N(j)) = \mathbb{1}_{Y_N(i)=Y_N(j)}. \tag{4.9}$$

Let us first make the naive assumption that the random variables $Z_N(i)$ are an independent family for $i = 0, \ldots, K$; this were justified if the random walk was self-avoiding. We will later see that this is not such a bad assumption if K is not too large. Under this assumption, things will be quite easy. To simplify matters, take $K = 2^n$. As is well known, the maximum of K independent standard Gaussian random variables is of order $\sqrt{2\ln K}$, and the point process

$$\sum_{i=0}^{K-1} \delta_{u_n^{-1}(Z_N(i))}$$

converges (as $K \uparrow \infty$) to a Poisson point process with intensity $e^{-x}dx$. Thus the time spent in one of the extreme states is of order

$$e^{\sqrt{N}\beta\sqrt{2\ln K}} = e^{\beta\sqrt{Nn}\sqrt{2\ln 2}}.$$

The question is whether it is this time is bigger or smaller than the time spent in the remaining states. This is checked as in the trap model: one splits the sum into two parts, the terms for which $Z_N(i) \geq u_n(-E)$, and those which are smaller. Then the mean of the contributions from the second gives at most

$$K\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u_n(-E)} e^{-z^2/2} e^{\beta\sqrt{N}z}$$
$$\sim \begin{cases} Ke^{N\beta^2/2}, & \text{if } u_n(-E) > \beta\sqrt{N}, \\ e^{\beta\sqrt{N}u_n(-E)}, & \text{otherwise} \end{cases}$$

In the second case, which corresponds to

$$\beta > \sqrt{\frac{n}{N}2\ln 2} \equiv \varrho, \tag{4.10}$$

this contribution vanishes compared to that from the extremes, while in the other case a law of large numbers holds. Let us assume that n is proportional to N such that $\varrho \in (0,1)$ is fixed.

Now note finally that

$$e^{\beta\sqrt{N}Z_N(i)} = e^{\beta\sqrt{N}u_n\left(u_n^{-1}(Z_N(i))\right)} = e^{\beta\sqrt{N}u_n(0)}e^{\frac{\beta}{\varrho}u_n^{-1}(Z_N(i))}.$$

Thus, if (4.10) is holds, we should expect that

$$e^{-\beta\sqrt{N}u_n(0)}\sum_{i=0}^{s(K-1)} e_i e^{\beta\sqrt{N}Z_N(i)} \sim \sum_{i=0}^{s(K-1)} e_i e^{\frac{\beta}{\varrho}u_n^{-1}(Z_N(i))} \sim V_\alpha(s),$$

where $\alpha \equiv \varrho/\beta$, and V_α an α-stable subordinator.

We see that then only assumption that is not justified in the above discussion is the independence of the process $Z_N(i)$.

To do this, we need some properties of the simple random walk on the hypercube.

4.3.1 Random walk on the hypercube

We want to show that the simple unbiased random walk on $\mathcal{S}_N$ started in some point σ behaves as follows:

(i) The first step takes the walk to a distance 1 from the starting point.
(ii) With probability of order $1 - 1/N$, the walk than reaches a distance $N/2$ before it returns to the starting point. This takes at least time $N/2$.
(iii) From distance $N/2$, the probability that the walk reaches the starting point before returning to distance $N/2$ is 2^{-N}. Thus, the probability that the walk returns to the origin before L unsuccessful trials (returns to distance $N/2$), is smaller than $L2^{-N}$.
(iv) After time $KN^2 \ln N$, the walk is exponentially close to equilibrium, i.e. for any $\sigma, \sigma' \in \mathcal{S}_N$

$$\left| \frac{\mathbb{P}_\sigma[Y_N(k) = \sigma' \cup Y_N(k+1) = \sigma']}{2} - 2^{-N} \right| < 2^{-8N}. \tag{4.11}$$

Statements (i) is trivial. Statement (ii) follows form Lemma 8.3 in [12]. Statement (iii) follows from Lemma 8.4 of [12] and reversibility (see also the proof of that lemma). Statement (iv) is well know and a proof can be found in [1].

Thus, as long as the $T \leq o(1) \times 2^N$, then the probability that the walk returns ever to a point it has once visited is of order $1/N + o(1)$.

Let us fix this observation:

Proposition 4.3.22 *Let $L \leq 2^{aN}$ with $a < 1$. Let Y_N be the SRW on $\mathcal{S}_N$, and let*

$$R_N(L) \equiv \{\sigma \in \mathcal{S}_N : \exists k \leq L : Y_N(k) = \sigma\} \tag{4.12}$$

denote the range of the random walk of length L. Then

$$\mathbb{P}\left[L(1 - N^{-1/2}) \leq |R_N(L)| \leq L\right] \geq 1 - cN^{-1/2}. \tag{4.13}$$

Proof We use the facts stated above. We will characterize the range as follows as a union of disjoint points:

$$R_N(L) = \bigcup_{\substack{k=0 \\ Y_\ell \neq Y_k, \forall \ell > k}}^{L} \{Y_N(k)\}, \tag{4.14}$$

i.e. we collect each last visit of a point σ. This gives for the cardinality (we simplify in the sequel $Y_N(k) \equiv Y_k$)

$$|R_N(L)| = \sum_{k=0}^{L} \mathbb{1}_{\{Y_\ell \neq Y_k, \forall \ell > k\}}. \tag{4.15}$$

Hence

$$\mathbb{E}|R_N(L)| = L\,\mathbb{P}\left[Y_\ell \neq Y_k, \forall \ell > k\right] \geq L(1 - cN^{-1}). \tag{4.16}$$

Hence the claimed result follows from Chebeychev's inequality. □

Remark 4.3.1 The estimate on the probability in (4.12) can be improved to allow to show that the event considered holds with probability one for all but finitely many N. To do this we have to use a second moment estimate and some de-correlation of the variables $\mathbb{1}_{\{Y_\ell \neq Y_k, \forall \ell > k\}}$ and $\mathbb{1}_{\{Y_\ell \neq Y_k, \forall \ell > m\}}$, if $m - k \geq N^2$, say. This follows easily from the fact that the random walk reaches equilibrium on a time scale of order $N \ln N$. I leave the details to the reader.

4.3.2 The subordinator on the SRW trajectory

We have now enough information on the SRW trajectories to conclude that the self-intersections do change the extremal properties of the process.

In fact, the argument is simple: We have to show that with probability tending to one, an iid process on the missing $N^{-1/2}L$ sites will not reach the level of the maximum over L sites: Clearly,

$$\mathbb{P}\left[\max_{i=1}^{LN^{-1/2}} X_i \geq u_L(E)\right] \leq N^{-1/2} L \mathbb{P}\left[X_i \geq u_L(E)\right] = N^{-1/2} e^{-E} \downarrow 0, \tag{4.17}$$

for any E. Thus, we can ignore the exceptional sites, as they do not contribute to the sum for any value of the truncation parameter E.

This allows us to conclude deduce the following main theorem.

Theorem 4.3.23 *Let $a < 1$ be fixed. Then,*

$$e^{-\beta\sqrt{N}u_{aN}(0)} S_N(s2^{aN}) \to V_\alpha(s), \tag{4.18}$$

in distribution (i.e. the law of the process defined on the Skorokhod space equipped with the J_1-topology converges), for almost all realizations on the SRW Y. Here $\alpha = \sqrt{2a\ln 2}/\beta$.

As a consequence, one can prove the same convergence result for the correlation functions as in the REM-like trap model.

Bibliography

[1] G. Ben Arous, A. Bovier, and J. Černý. Universality of the rem for dynamics of mean-field spin glasses. preprint, arXiv math.PR 0706.2135, 2007.

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

[3] 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.

[4] G. Ben Arous and J. Černý. The arcsine law as a universal aging scheme for trap models. to appear in Communications on Pure and Applied Mathematics, 2007.

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

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

[7] Gérard Ben Arous, Jiří Černý, and Thomas Mountford. Aging in two-dimensional Bouchaud's model. *Probab. Theory Related Fields*, 134(1):1–43, 2006.

[8] Jean Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

[9] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.

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

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

[12] Anton Bovier. Metastability and ageing in stochastic dynamics. In *Dynamics and randomness II*, volume 10 of *Nonlinear Phenom. Complex Systems*, pages 17–79. Kluwer Acad. Publ., Dordrecht, 2004.

[13] Anton Bovier. *Statistical mechanics of disordered systems*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2006.

[14] Anton Bovier and Alessandra Faggionato. Spectral characterization of aging: the REM-like trap model. *Ann. Appl. Probab.*, 15(3):1997–2037, 2005.

[15] William Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.

[16] 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.

[17] W. Whitt. *Stochastic-process limits*. Springer Series in Operations Research. Springer-Verlag, New York, 2002.